Expanding market participation among smallholder livestock producers

A collection of studies employing Gibbs sampling and data from the Ethiopian highlands, 1998–2001

Socio-economics and Policy Research Working Paper 48

G. Holloway and S. Ehui

International Livestock Research Institute (ILRI)
P.O. Box 30709, Nairobi, Kenya

Working Papers Editorial Committee

Authors' affiliations

Garth Holloway, Livestock Policy Analysis Programme, International Livestock Research Institute (ILRI), P.O. Box 5689, Addis Ababa, Ethiopia, and University of Reading, UK
Simeon Ehui, Livestock Policy Analysis Programme, ILRI, P.O. Box 5689, Addis Ababa, Ethiopia

---

Table of Contents

Abstract

This compendium reproduces results from several, independent research projects undertaken at the Livestock Policy Analysis Programme (LPAP) of the International Livestock Research Institute (ILRI), Addis Ababa, Ethiopia. This reproduction brings together separate pieces of research that relate to the same goals, namely, market expansion, food security, poverty alleviation and hunger prevention. It is to showcase the power of Markov chain Monte Carlo (MCMC) methods, particularly Gibbs sampling, in providing direct answers to policy questions. It is hoped that the empirical research showcased in this compendium will spur other researchers to apply MCMC methods and the Bayesian paradigm to the heterogeneous research projects and policy questions that employed applied empirical research encounters in less developed regions.

Keywords: Compendium, LPAP, ILRI, MCMC, Bayesian paradigm, applied empirical research, less developed areas.

Journal of Economic Literature Classifications: O11, C34, O13

Acknowledgements

1   Introduction

Projected increases in consumption of animal products in the developing world can improve incomes of poor farmers and food processors. Indeed the expected increase in demand for livestock products brought about by increased incomes, urbanisation and population growth presents new and expanding market opportunities for smallholder livestock producers in the developing world (Delgado 1999). However, inappropriate policies and misallocation of investment resources could skew the distribution of the benefits and opportunities away from the smallholders who would potentially gain the most from these market opportunities. In this context, a search for policies designed to effect benefits to smallholders seems appropriate.

A major constraint to increasing the welfare of smallholders is their inability to access markets. Enhancing the ability of poor smallholder farmers to reach markets, and actively engage in them, is one of the most pressing development challenges. Remoteness results in reduced farm-gate prices and returns to labour and capital and increased input costs. This, in turn, reduces incentives to participate in economic transactions and results in subsistence rather than market-oriented production systems. Sparsely populated rural areas, remoteness from towns and high transport costs are physical barriers in accessing markets. Transaction costs such as lack of information about markets, lack of negotiating skills, and lack of collective organisation are other impediments to market access. The question of how to expand the market participation of smallholder livestock producers is a major challenge facing many governments and non-governmental organisations (NGOs) in developing countries.

In recent years several studies have been conducted using data from the Ethiopian highlands. The objectives of these studies are to (1) identify the resources that can promote entry and sustain milk market development; (2) identify the levels of these resources that are required for entry into the market; and (3) identify the minimum efficient scale of operations that is required for entry.

Small-scale dairy production is an important source of cash income for subsistence farmers in the East African Highlands. Dairy products are a traditional consumption item with strong demand and the temperate climate allows the crossbreeding of local cows with European dairy breeds to raise productivity.

Despite this potential, smallholder participation in market-led dairy development has not been widespread in sub-Saharan Africa (SSA) outside of Kenya. Even in regions with favourable climates for livestock development, such as the Ethiopian highlands, participation in fluid milk markets by rural smallholders is limited. Changes in sectoral and macro-economic policies are frequently necessary, but not sufficient, to provide the requisite incentives for smallholders to participate in markets. Barriers to smallholder participation in dairy production range from the availability and cost of animals to the labour needed to bring products to market. Small-scale milk producers face many hidden costs that make it difficult for them to gain access to markets and productive assets. Among the barriers that may be influenced by policy are transactions costs—the pecuniary and non-pecuniary costs associated with arranging and carrying out an exchange of goods or services.

The existence of relatively high marketing costs for fluid milk in Africa, the prevalence of thinness in fluid milk markets, and the risks attached to marketing perishables in the tropics suggest that transactions costs play a central role in dairy production and marketing. Under such conditions, producers marketing co-operatives that effectively reduce transactions costs may enhance participation. Hence, it is vital to know what governments can do to better

support these organisations and their emergence, and determine whether alternative institutions should be encouraged.

This compendium reproduces results from several research projects undertaken at the Livestock Policy Analysis Programme (LPAP) of the International Livestock Research Institute (ILRI). Collecting these projects in a single volume brings together separate pieces of research that relate to the same goals, namely, market expansion, food security, poverty alleviation and hunger prevention, and to showcase the power of Markov chain Monte Carlo (MCMC) methods (particularly Gibbs sampling) and the Bayesian paradigm, in providing direct answers to policy questions.

It is hoped that some of the methods showcased in this compendium will spur other researchers to apply the techniques in this volume to the heterogeneous research projects and policy questions that empirical research often encounters in developing countries.

2 Markov chain Monte Carlo methods and the Gibbs sampler

Like some of the innovations studied as the basis for interrelating these separate works, occasionally there arises an innovation that is so rich and powerful that it seems surprising that it is not more frequently exploited. In modern mathematical statistics—and in the history of mathematical statistics prior to Bayes theorem—there probably has not been an innovation that has had such a lasting impact on the way scientists view problems in empirical science and the stock of methods collected for solving these problems. The advent of Markov chain Monte Carlo (MCMC) methods is one such innovation. The purpose of this chapter is to introduce briefly MCMC to the reader in a user-friendly manner.

2.1 Background

Although there was a significant lag before Markov chain Monte Carlo (MCMC) methods became noticed by the profession—the original papers dating to Metropolis and Hastings in the 1950s—MCMC has had a famous impact since the seminal statistical contributions appeared (Gelfand and Smith 1990). Since the early 1990's there have been significant advances in the biological sciences, the humanities, and particularly the medical sciences. However, the economics and agricultural economics sciences have been less fervent in their acceptance of these new techniques. Whether this state of affairs arises due to a disproportionately fewer number of Bayesians working in these fields is open to debate, but this situation is clearly changing. Moreover, the advent of MCMC has stimulated considerable new entry into applied Bayesian statistics.

This collection of papers is related by its application of MCMC to solving two important problems for economic development policy. These problems are the identification of the resources and their quantities that are necessary to effect entry among representative non-participating households at two sites close to Addis Ababa from which data were collected in the 1997 production year.

Broadly defined, the project's objectives were to identify the factors that precipitate the emergence of new milk markets when the presence of relatively high transactions costs were considered a major impediment. This objective is important. One of the major impediments to economic development throughout SSA is a lack of density of market participation (Stiglitz 1989). Over time, and with varying personnel involved, these objectives became more refined and focused on the end products that are contained in this collection.

But the final product can never be better than the quality of the original outlay, nor can it improve upon the necessary inputs in a new venture. The essential inputs into any empirical

exercise are the data. The data used in these various projects are very rich and it is important to keep this in mind as one reads through the diverse set of applications contained in this compendium. The data are due, primarily, to the efforts laid out by Charles Nicholson in the 1997 survey period while he was a postdoctoral fellow at the International Livestock Research Institute (ILRI), supported by a grant from The Rockefeller Foundation. As well as showcasing some of the important contributions of MCMC, the works collected here showcase the quality of the data collected by an ILRI scientist.

2.2 Objectives

To derive policy prescriptions for identifying opportunities for expanding market participation, the independent works collected here provide classic examples of the scientific value of innovation, specifically, the routine application of MCMC to quality data. We introduce MCMC to the reader through selecting the crucial component of all the models developed and expanded upon in this collection and demonstrate its operation with a limited available information. The crucial component in all the projects is the normal-linear model. Each of the models further developed here are simple extensions in the number of unknown quantities, in the forms of the distributions that they entail, and in the complexity of relationships between the components that the normal-linear model underlies. Hence, providing an example of the technique's operation in this context to satisfy the reader with improved sets of information (most of the empirical models here are based on data sets ranging in size between 204 to 1428 observations) can do much better. The success of these procedures is based on the crucial issue of 'convergence to the true distribution' that the next section demonstrates in the simplest framework possible.

2.3 Bayes' rule, conditioning and the Gibbs sampler

Consider the problem of locating the mean of a normal distribution, from data $y \equiv (y_1, y_2, ..., y_N)'$ that are independent and identically distributed as Normal (μ, σ). This is a two-parameter problem in the unknown quantities $\theta \equiv (\mu, \sigma)'$. The conventional Bayesian approach to this problem is to set-up a prior probability distribution for the unknown quantities, $\pi(\theta)$, combine this prior with the observed data likelihood, $\ell(\theta|y)$, and, through Bayes rule, derive the joint posterior distribution for the unknown quantities of interest, after observing the data

$$\pi(\theta|y) \propto \ell(\theta|y)\pi(\theta) \qquad (1)$$

where the symbol 'μ' means 'is proportional to'. In other words, net of an unimportant constant (that simply scales the posterior probability density function so that it has mass equal to one and thereby constitutes a true density) the posterior measure on the left-hand side is a joint probability density function for the unknown quantities θ and is the *target* density in the exercise. This is simply Bayes theorem.

MCMC pertains to the analysis of the joint posterior density and, particularly, the derivation of the marginal probability distribution functions, $\pi(\mu|y)$ and $\pi(\sigma|y)$, which are the end products of any Bayesian investigation.

When the functional form of $\pi(\theta|y)$ permits integration of each of the components of the joint density, marginal probability distributions can be derived easily. But when is intractable, meaning that the necessary integrations cannot be carried out in closed form, a number of numerical avenues opens for empirical analysis. One of these alternative approaches is MCMC and, two of its special cases—data augmentation and Gibbs sampling—provide the basis for all of the estimation conducted in this collection. Here, we restrict attention to the Gibbs sampler.

The Gibbs sampler becomes a candidate for evaluating the joint posterior distribution when each of the full, conditional distributions comprising the joint posterior have well-known forms that are easy to sample from.

This point is worth re-emphasising. The application of the Gibbs sampler requires two conditions to be met. First, we require that the marginal probability density functions for each of the component distributions are not available in closed form; without this condition there would be no need to make any numerical approximation to the posterior. Second, we require that each of the full conditional distributions comprising the joint posterior have well-known forms. Each of the problems that follow this chapter are linked in their fulfillment of these two criteria and, hence, we can apply the Gibbs sampler in order to derive inferences about unknown quantities of interest.

2.4 Application to normal means

The problem of locating the mean among normal data is a problem with a tractable posterior for which no MCMC approximation is required, but, in view of its familiarity, it is useful for demonstration.

In the normal-data example, the component conditional distributions do have well known forms. A standard, non-informative prior, $\pi(\theta) \mu \sigma^{-1}$, leads to a joint posterior which, in turn, results in component conditional distributions that have, respectively, normal and inverse-gamma forms. Precisely, the posterior distribution for the mean, μ, conditional on the standard deviation, σ, is normal and the posterior distribution for the standard deviation conditional on the mean has an inverted-gamma form. Under relatively mild regularity conditions (Gelfand and Smith 1990) that are satisfied by each of the models developed in this compendium, the draws that alternate in sequence between μ, conditional on , on the one hand; and conditional on μ, on the other, form a Markov chain with highly desirable convergence properties. Specifically, under the stated regularity conditions these chains converge to the target probability distributions that we seek.

It is essentially these observations, and generalisations of them that are employed repeatedly throughout this paper.

The interested reader is directed to some slightly stronger (perhaps, more persuasive) mathematical arguments in Casella and George (1993), and in Chib and Greenberg (1995). Below, we give two examples of a special case of the normal-linear model, which is used repeatedly as a basis for investigation in each of the discrete- and truncated-distribution problems arising in examining market participation.

2.5 Demonstrations

Suppose that the data vector has length $N = 10$, and the mean and variance are, respectively zero and one, so that the data are independent and identically distributed standard normal. Figure 1 presents examples of convergence in distribution by presenting the results of the first 50 draws in the Gibbs sequence, when the sequence is given the highly unrealistic starting values $(\mu^0, \sigma^0) = (1000, 1000)$. Note that we only have 10 observations from which to draw inferences. However, the convergence in distribution is quite striking. The draws oscillate for the first few iterations until the target distributions are located and thereupon simulate draws from the target conditional distributions, namely a normal distribution, $N(\sum_i y_i /N, 1)$ for the mean, with posterior mean equal to the data mean, and an inverted-gamma distribution for the variance parameter with parameters, $v = N$ and $s^2 = \sum y_i^2/N$.

Figure 1. *Convergence in the Gibbs sample.*

There are three features worth re-emphasising. First, we obtain convergence with very limited information; here we have only 10 sample observations. Second, we obtain convergence even with very unrealistic starting values. Third, convergence to the target densities is extremely rapid.

In the previous example, the marginal distributions of interest can be obtained exactly and hence, no Gibbs sampling is actually required to simulate a draw from the target density. Now, consider an identical set-up but with the addition that data are censored at the a priori unknown mean. In this case, due to the evaluation of integrals implied by the censoring, Gibbs sampling is required to simulate a draw from the joint posterior. However, the same early convergence in distributions emerges (Figure 2).

Figure 2. *Convergence in the Gibbs sample.*

This example is important, for three reasons. First, it re-emphasises the important point that convergence in distribution occurs quite quickly—even in a limited information environment. Second, it confirms the assertion that the convergence in the standard normal-means model was not due in any way to the particular simplicity of that model. Third, the example shows, in perhaps the simplest setting, that the Gibbs sampling procedure works to good effect when the data in question are censored. This latter point is important when it is recognised that the bulk of the models visited in this collection contain censored data, in particular, censored observations on household marketable surplus.

In summary, these two examples are not intended to provide overwhelming evidence of the

usefulness of the procedures encountered in the remainder of this collection, nor its power in small sample sizes. Rather, the exercises are intended to give a flavour of the power of the technique under relatively unfavourable sampling circumstances.

The normal-inverted-gamma form, it is worth stressing, is an obvious example to choose because it also provides the basis for all of the applications that follow. Most of these applications possess somewhat more complicated posterior forms. However, with the exception of only two of them, most of these forms appear frequently in the literature and have been studied, repeatedly, in a similar context to the examples just presented. If these simple examples are not persuasive, we hope that the many applications that we now visit will convince the reader of the power of the methodology in analysing important empirical problems with considerable policy importance.

2.6 Overview

The remainder of this work is organised as follows. Section 3 introduces the data used in the various applications, provides the motivation for their collection and presents summary of statistical reports. Section 4 applies a standard probit procedure to the binary participation data. Section 5 applies a single-equation Tobit procedure to the marketable surplus data. Sections 6 to 9 consider various extensions of the basic Tobit and probit set-ups. Section 6 considers the impact of production decisions; Section 7 considers a count-data problem in crossbred cow adoption; Section 8 introduces fixed costs; and Section 9 considers two-step participation and selling decisions. Conclusions are offered in Section10.

3 Introduction to the Nicholson data

Because each of the applications that follow rely on a data set collected at ILRI in 1997, it seems appropriate to introduce these data, the reasons why they were collected and the topic of primary interest at the outset of each of the independent projects that this report summarises.

3.1 Background

Small-scale dairy production is an important source of cash income for subsistence farmers in the East African highlands. Dairy products are a traditional consumption item with strong demand, and the temperate climate allows crossbreeding of local cows with European dairy breeds to raise productivity. Particularly where infrastructure and expertise in dairy processing exist, such markets allow smallholders to participate in the agro-industrial sub-sector and potentially in regional export markets and beyond. Moreover, growth in dairy demand in sub-Saharan Africa (SSA) is projected to increase over the next 20 years due to expected population and income growth. Milk production and dairy product consumption are expected to grow in the region of 3.8 to 4% annually between 1993 and 2020 (Delgado 1999). Increased domestic dairy production has the potential in much of SSA to generate additional income and employment and thereby improve the welfare of rural populations (Staal et al. 1997). However, there are concerns that the benefits of this expected growth might bypass resource-poor livestock producers unless specific policy actions are taken.

Barriers to smallholder participation in dairy production range from the availability and cost of animals to the labour needed to bring products to market. Despite the potential, smallholder participation in market-led dairy development has not been widespread in SSA outside of Kenya. Even in regions with favourable climates for livestock development, such as the Ethiopian highlands, participation in fluid milk markets by rural smallholders has been limited. Changes in sectoral and macro-economic policies are frequently necessary, but not sufficient, to provide the requisite incentives for smallholders to participate in markets.

Small-scale milk producers face many hidden costs that make it difficult for them to gain access to markets and productive assets (Staal et al. 1997). Among the barriers that may be influenced by policy are transactions costs—the pecuniary and non-pecuniary costs associated with arranging and carrying out an exchange of goods or services. The existence of relatively high marketing costs for fluid milk in Africa, the prevalence of thinness in fluid milk markets and the risks attached to marketing perishables in the tropics suggest that transactions costs play a central role in dairy production and marketing. Under such

conditions, producers' marketing co-operatives that effectively reduce transactions costs may enhance participation. Hence, it is vital to know what governments can do to better support these organisations and their emergence, and determine whether alternative institutions should be encouraged.

3.2 Transactions costs, co-operatives and milk market development

Transaction costs are the embodiment of barriers to access to market participation by resource-poor smallholders. They include the costs of searching for a partner with whom to exchange, screening potential trading partners to ascertain their trustworthiness, bargaining with potential trading partners (and officials) to reach an agreement, transferring the product, monitoring the agreement to see that its conditions are fulfilled, and enforcing the exchange agreement.

The nature of milk and its derivatives in part explains the high transactions costs associated with exchanges of fluid milk. Raw milk is highly perishable and, thus, requires rapid transportation to consumption centres or processing into less perishable forms. Further, bulking of milk from multiple suppliers increases the potential level of losses due to spoilage. These losses limit marketing options for small and remote dairy producers, raise transport costs, and imply greater losses due to spoilage than commodities such as grains. Because milk production typically is a year-round activity, dairy producers often must be concerned with maintaining outlets for their production.

The search for stable market outlets by producers is complicated by significant seasonal variation in milk production and dairy product consumption (Debrah and Berhanu 1991; Jaffee 1994). In part due to high perishability, but also due to natural variation, milk quality is variable. Some of its properties (e.g. bacterial counts) are also not easily ascertained. Although not a perfect proxy, we conjecture that distance between production and purchasing points is highly correlated with quality, which declines rapidly after milking. The lack of easily measurable quality standards may also allow agents purchasing raw milk from producers to reject milk without just cause when they have contracted to purchase more milk than can be profitably sold.

Differential transactions costs among households stem from asymmetries in access to assets, information, services and remunerative markets (Delgado 1999). Handling these access problems requires institutional innovation. First, the problem of resource-poor smallholders is often so great that a net transfer (such as a heifer) is necessary to induce entry. Second, technical and market information for new commercial items is more likely to be useful to individuals with higher levels of schooling, greater work experience, better access to management and technical advice, and better knowledge of market opportunities. Smallholders may require particular support in information and management. Third, access to services is often unequally distributed within communities. Poor infrastructure, low population density, and low effective demand make necessary institutions for risk-sharing and economies of scale in provision of agricultural services, especially in remoter areas. Fourth, better access to remunerative markets for high-value-to-weight items is necessary for promoting growth of smallholder agriculture.

3.3 Co-operatives as catalysts

A common form of collective action to address access problems of this type is a participatory, farmer-led co-operative that handles input purchasing and distribution and output marketing, usually after some form of bulking or processing. Farmers gain the benefit of assured supplies of the right inputs at the right time, frequently, credit against output deliveries, and an assured market for the output at a price that is not always known in advance, but applied equally to all

farmers in given location and time period. Extension is sometimes part of the services provided, typically at higher rates (and quality) than state extension services. Co-operatives, by providing bulking and bargaining services, increase outlet market access and help farmers avoid the hazards of being encumbered with a perishable crop with no rural demand. In short, participatory co-operatives are very helpful in overcoming access barriers to assets, information, services, and the markets within which smallholders wish to produce high-value items (Jaffee 1994).

Like contract farming, producer co-operatives can offer processors/marketers the advantage of an assured supply of the commodity at known intervals at a fixed price and a controlled quality. They can also provide the option of making collateralised loans to farmers. For processors or marketers, such arrangements eliminate the principal agent issues faced by collectives and out grower schemes in monitoring effort by the individual producer. The schemes also provide better relations with local communities than large-scale farms, avoiding the expense and risk of investing in such enterprises, sharing production risk with the farmer, and helping ensure that farmers provide produce of a consistent quality (Grosh 1994; Delgado 1999).

Producers' co-operatives are unlike contract farming schemes, however, with respect to negotiations among different partners. If the issue in contract farming revolves around the power of farmers to negotiate with processors in producers' co-operatives, the issue in the co-operatives themselves is the power of members, collectively, to hold management accountable. Producers' co-operatives in Africa have had a generally unhappy history, because of difficulties in holding management accountable to the members (i.e. shirking), leading to inappropriate political activities or financial irregularities in management (de Janvry et al. 1993; Akwabi-Ameyaw 1997), and also due to over-ambitious investment in scale and enterprises beyond management's capability. The degree of moral hazard seems to be greater if co-operatives are general in their orientations rather than created for specific purposes, such as farmer-run local milk marketing co-operatives in Uganda and Kenya (Staal et al. 1997). In Ethiopia, on the other hand, the perception exists that there may be enormous potential for their role, in concert with production innovations, as market precipitators (Nicholson 1997).

3.4 Experience in Ethiopia

The traditional system of milk production in Ethiopia, comprising small rural and peri-urban farmers, uses local breeds, which produce about 400–680 kg of milk/cow per lactation period (Debrah and Berhanu 1991). More recently, intensive systems as diverse as state enterprises and small and large private farms use exotic breeds and their crosses, which have the potential to produce 1120–2500 litres over a 279-day lactation (Debrah and Berhanu 1991). Fresh milk marketing is channelled through both formal and informal outlets, with informal markets supplying some 85% of total fresh milk in the Addis Ababa area (Staal 1995). The major formal outlets are dominated by a government enterprise called the Dairy Development Enterprise (DDE), which has established numerous collection centres that buy milk at a uniform government controlled price that requires no minimum delivery. In 1992/93, the DDE supplied 12% of total fresh milk sales in Addis Ababa (Staal 1995). The DDE is concerned primarily with fluid milk marketing, although it does make some cheese and yoghurt in its Addis Ababa processing facilities.

The informal fresh milk market involves direct delivery of raw milk by producers to consumers in the immediate neighbourhood and sales to itinerant traders or individuals in nearby towns. Milk is transported to towns on foot, by donkey, by horse or public transport and frequently commands a higher price than in the originating locale (Debrah and Berhanu 1991). In Ethiopia, fresh milk sales by smallholder farmers are important only when they are close to formal milk marketing facilities such as government enterprises or milk groups. Results from a

sample of farmers in northern Shewa in 1986 estimated that 96% of the marketable milk was sold to the DDE (Debrah and Berhanu 1991). Farmers far from such formal marketing outlets prefer to produce other dairy products instead, such as cooking butter and cottage cheese. In fact, the vast majority of milk produced outside urban centres in Ethiopia is processed into products by the farm household, and sold to traders or other households in local markets.

The other principal outlets for milk are 'milk groups', which are milk marketing co-operatives recently established by the Ethiopian Ministry of Agriculture's Smallholder Dairy Development Project (SDDP) with the support of the Finnish International Development Association (FINNIDA). The milk groups buy milk from both members and non-members, process it, and sell the derivative products to traders and local consumers. Although the milk groups sometimes sell fluid milk products such as sour milk, skim milk, or buttermilk, most of their revenue is generated by sales of processed dairy products, butter and cottage cheese (Nicholson et al. 2000). The groups do not presently represent a significant source of fresh milk for either rural or urban markets.

3.5 Survey design

The SDDP milk groups purchase raw milk from farmers, then use hand-operated equipment to process the milk into butter, local cottage-type cheese (*ayib*), and yoghurt-like sour milk (*ergo*). These dairy products are sold to local households, to traders who market them in turn to major urban centres, and local restauranteurs. Typically, the value added from processing the fluid milk into products (less funds retained for maintenance of the groups' facilities) is returned as a semi-annual, lump-sum payment to group members and others who have supplied the group during the period since the previous payment.

At the time of data collection four of these milk groups existed, two in the Shewa region north of Addis Ababa and two in the Arsi region near the regional centre, Asella. The activities of these groups are focused exclusively on the processing and selling of dairy products. They provide no additional services (i.e. no credit, feeds, veterinary services etc.) to farmers or to buyers and, therefore, represent the simpler end of the continuum of activities that co-operative organisations might undertake.

Although the number of farmers and the amount of milk received at each group is not a large proportion of regional totals, the formation of these groups has created a new outlet for sales of fluid milk by producers. Prior to the formation of the groups, the households processed nearly all locally produced milk into butter and *ayib* (a local form of cottage cheese). Even now, most milk produced in these areas is marketed as home- processed dairy products and sold to traders or other households in local markets. Thus, the milk groups can be considered organisational innovations that increase the number of marketing options available to smallholder dairy farmers and mitigate some of the principal transactions costs that retard entry.

3.6 Data collection procedures

In order to respond to questions surrounding the possible impacts of 'other factors' on entry, data were collected from four rural communities called peasant associations (PAs) (which are state-designated partitions of rural districts) near two of the four milk groups formed by the SDDP. Preliminary surveys were undertaken in December 1996 and January 1997 to ascertain the extent of crossbred cow ownership. On the basis of the preliminary surveys, the Mirti and Ashebaka PAs in the area of the Lemu Ariya milk group were selected from Arsi region, and the Ilu-Kura and Archo PAs were selected near the Edoro milk group in Shewa region. One PA in each region was close enough to the milk group that co-operative selling occurred; the other was distant enough that sales were precluded. None of the households in

the Ashebaka and Archo PAs participated in the milk groups, whereas a proportion of the households in Mirti and Ilu-Kura PAs delivered milk to the groups.

A census of households in these four PAs was conducted for the purpose of developing a sampling frame. Using the census results, a sample of 36 households was selected in each of the PAs, stratified by whether the household owned crossbred cows, participated in the group, and their distance to the group or to another local markets where dairy products could be sold. During June 1997, baseline surveys of household characteristics and current cattle management practices were administered to 144 households. From June 1997 to October 1997, data on milk allocation and marketing, significant events occurring in the cattle herd (births, deaths, purchases, sales, illness etc.), and cow feeding practices were collected every two to three weeks.

In the empirical applications that follow we focus on the 68 households in the Mirti and Ilu-Kura PAs for which samples were observed on milk sales in the seven days prior to three respective visits, yielding a total of $1428 = 68 \times 7 \times 3$ observations. Importantly, only 15% of the observations correspond to participating households. Table 1 summarises the data by market participation status.

Table 1. *Selected characteristics of survey households, by market participation status.*

	Sample means (standard errors)	
	Sold to the milk group	Did not sell to the milk group
Number of crossbred cows	1.41	0.49
t = 15.32	(0.99)	(0.69)
Number of local cows	1.26	1.42
t = −1.81	(1.03)	(1.12)
Time to the milk group, minutes	35.16	45.53
t = −4.37	(18.76)	(29.94)
Farm experience of household head, years	23.20	24.79
t = −1.22	(12.58)	(16.21)
Formal schooling of household head, years	1.96	1.90
t = 0.22	(4.01)	(3.24)
Visits by an extension agent during past year	3.19	0.78
t = 14.74	(3.59)	(1.66)

Note: t statistics (1426 degrees of freedom) reported for difference between means.

4   Probit analysis of the participation decision

The natural vehicle for analysis in preliminary investigations of household-panel data is the probit model. Having described the data, the motivation for the survey and the survey collection procedures, this chapter is concerned with the motivation and application of MCMC to probit estimation.

4.1 Motivation

Motivation for application of the probit model follows. Let i = 1, 2, ..., N, denote the households in question. Each household compares the level of utility derived from market participation, y_i^*, against its reservation utility attainable without market participation, vi^*. Here, we use an asterisk (*) to denote the fact that both levels of utility are latent random variables.

Assuming that differences between utilities are determined by characteristics, we assume that these characteristics are specific to each household, $x_i \equiv (x_{i1}, x_{i2}, ..., x_{iq})$. Without loss of generality, we set $vi^* = 0$ and denote the difference between the incurred and reserve utility levels yi^*, and their relationship to the characteristics by the function. The condition characterising the discrete choice about whether to participate in the market can then be written as:

$$y_i^\wedge = f_i(x_i) \qquad (2)$$

with participation when yi > 0 and non-participation otherwise. We define the indicator variable $\delta_i = 1$ when $y_i > 0$ and the household participates in the market, and $\delta_i = 0$ under non-participation. This is the standard 'index-utility' representation of the probit model and is the natural first-step in assessing household-panel data in a market-entry situation.

4.2 Statistical implementation

Statistical implementation of this simple framework follows closely the ideas outlined in Albert and Chib (1993). A linear version of the participation equation (2) has the form

$$z_i = x_i \beta + u_i \qquad (3)$$

where $z_i > 0$ if $\delta_i = 1$ and $\delta_i = 0$ otherwise; and where β is a vector of unknown coefficients controlling the relationship between household-specific characteristics and market participation, and u_i is a random error. The econometrician observes data $\delta_i = 1$ if the latent random variable $z_i > 0$ and $\delta_i = 0$ otherwise; and observes the vector of household-specific covariates, x_i. The objective is to draw inferences about β and any other structural parameters by combining the observed and latent data. To do this, we assume that the participation variable, z_i, has a normal distribution with mean the product of the conditioning data and the unknown coefficient matrix $x_i\beta$ and variance equal to one. The restriction on the variance is imposed for identification purposes.

4.3 Specification and estimation

The estimation procedure can be introduced by looking at the complete-data model, which we denote

$$z = x\beta + u \tag{4}$$

where $z \equiv (z_1, z_2, ..., z_N)$ is the latent data; $x \equiv (x_1, x_2, ..., x_N)$, $x_1 \equiv (x_{11}, x_{12}, ..., x_{1q})$, $x_2 \equiv (x_{21}, x_{22}, ..., x_{2q})$, ..., $x_N \equiv (x_{N1}, x_{N2}, ..., x_{Nq})$, $x \equiv (x_1, x_2, ..., x_N)$ are observations on the covariates; $\beta_1, \beta_2, ..., \beta_q$ is the parameter depicting the effects of changes in the covariates on the latent data; and the error vector $u \equiv (u_1, u_2, ..., u_N)$ is assumed to have the normal distribution $N(O_N, I_N)$, where O_N denotes the N-dimensional null vector and IN is the $N \times N$ identity matrix.

With this notation at hand we use a conventional non-informative prior for the unknown parameters, namely $\pi(\beta) \mu 1$. (Recall that the covariance is restricted to take unit value for identification purposes.) Even with the unit restriction on the error variance, the model in its current setting is still intractable, due to the evaluation of integrals implied by the probit set-up. The key step in overcoming this impediment—as ably demonstrated in Albert and Chib (1993) —is to augment the observed data likelihood with the latent data and derive estimates of these latent data as part of the estimation exercise. Accordingly, with the prior now specified to include these latent data, $\pi(\beta, z) \mu 1$, the complete conditional distributions characterising the joint posterior distribution for the parameters β and the latent data z have simple forms. In particular, in terms of the current notation, these conditional distributions are

$$z|\beta \sim \text{Truncated-normal } (Ez, Vz)$$
$$\beta|z \sim \text{normal } (E\beta, V\beta) \tag{5}$$

where $Ez \equiv x\beta$, $Vz \equiv I_N$, $E\beta \equiv (x'x)^{-1}x'z$ and $V\beta \equiv (x'x)^{-1}$. The crucial observation is that these two distributions are easy to sample from. Consequently, simulations from the joint posterior can be undertaken through the following, simple algorithm:

Step 1: Select starting values $z^{(s)}$.

Step 2: Draw $\beta^{(s+1)}$ from the multivariate-normal $(E\beta^{(s+1)}, V\beta^{(s+1)}$ distribution,

   where $E\beta^{(s+1)}$ and $V\beta^{(s+1)}$ denote conditioning on $z^{(s)}$ from Step 1.

Step 3: Draw $z^{(s+1)}$ from the multivariate-normal $(Ez^{(s+1)}, Vz^{(s+1)})$ distribution (6)

   where $Ez^{(s+1)}$ and $Vz^{(s+1)}$ denote conditioning on $\beta^{(s+1)}$ from Step 2.

Step 4: Repeat steps 1–3 many times, S^1, until convergence is attained.

Step 5: Repeat steps 1–3 many times, S^2, and collect samples $\{\beta^{(s)}\ s = 1, 2, ..., S^2\}$

and $\{z^{(s)}\ s = 1, 2, ..., S^2\}$.

The draws in the last step can be used to compute summary statistics (means, medians, standard deviations) or plot histograms of any summary measure of interest. In the results reported below, the algorithm is run for a 'burn-in phase' of $S^1 = 2000$ observations followed by a 'collection phase' of $S^2 = 2000$ observations.

4.4 Estimating distance to market

While the impact of the covariates x on the latent participation variable z are important themselves, more interest resides in computing a measure of additional resources required for each of the non-participating households to enter the market. We call these measures 'distances to market'. Specifically, these distance measures are estimates of the additional levels of the regressors that make each non-participating household in the sample become active in the market. This question is answered directly as a by-product of data augmentation, showcasing the power of Markov chain Monte Carlo methods in policy formation.

Recall that, in each round of the algorithm in (6) we compute an estimate of the latent participation variable. For households that do not participate in the market this quantity has a negative value. This (negative-valued) quantity has important implications for policy. A household with a larger negative-valued latent variable is further from the market than one that has a smaller-valued latent quantity. But these estimates can, in turn, be transformed into a meaningful distance measure across each of the covariates in the model using some simple algebra.

Suppose, in the context of equations (4) that we wish to measure distance in terms of independent variable 'k', then all we need do is solve (setting the left-hand-side to zero) the probit equation in terms of the value of 'covariate k' and then subtract from it the household's observed level of the resource in question. The quantity that results is fundamental for policy because it provides an estimate of resource deficiency in the household and, hence, provides an estimate of the additional amount of the resource that is required to engender positive marketable surplus. It follows that these quantities are the ones that precipitate entry into the market, dilute the density of non-participation and, therefore, overcome a main impediment to economic development. They are the values:

$$\hat{x}_{ki} = \frac{\beta_0 + \sum_{j=1}^{q} \beta_j x_{ji} + u_i}{-\beta_1} - x_{ki}, \ k = 1, 2,...,q;\ i \in c \qquad (7)$$

where (the censor set) c denotes the set of households that do not participate in the market. Formally, $c \equiv \{i|\delta_i = 0\}$.

Note that quantities (7) are available across *each* non-participating household in the censor set. Therefore, further enhancing their appeal as policy measures, they can be used to provide precise measures of the levels of *each* covariate for *each* household. The question remaining (that is particularly relevant in the context of Bayesian inference) is the existence of a posterior distribution of each of these distance estimates and the existence of moments and other measures of central tendency that can be used to characterise these distributions— especially their locations and their scales.

Because the distance measures contain quantities that are either observed or are easily simulated as a by-product of the Gibbs sample, the natural inclination is to use the formulae, together with the outputs $\{\beta^{(s)}\ s = 1, 2, ..., S^2\}$ and $\{z^{(s)}\ s = 1, 2, ..., S^2\}$ in (6) to compute quantities $\{\hat{x}_{ki}(s)\ s = 1, 2, ..., S^2\}$ from which means, standard errors and histograms can be constructed. However, these measures can only be meaningful when the posterior distribution is proper (that is, the distribution integrates to a finite measure) and there exist posterior moments to which the sample estimates correspond. Each of the quantities in (7) contains a quotient that is (conditionally distributed as) a ratio of normal random variables and, thus, the proposal to use the output of the sample to compute $\{\hat{x}_{ki}^{(s)}\ s = 1, 2, ..., S^2\}$ requires that the distribution of these ratios of normal random variates be proper and that their moments exist.

Findings by early contributors in the field (Merrill 1928; Geary 1930; Fieller 1932; Hinkley 1969) generate two relevant conclusions. First, the distributions of the quotients are proper but, second, moments may not exist. In loose terms, the requirements for the existence of moments depend on a 'relative-variance condition', namely that the means of the quantities in the denominators of the quotients on the right sides of (7) are 'large' in relation to their corresponding standard deviations. When this condition is met, moments exist and it is appropriate to characterise the locations and scales of the distributions through sample means and variances. When the moments do not exist (but the distributions are proper), it is inappropriate to use mean and variance estimates, but appropriate to use other descriptive measures such as histograms, posterior modes or, perhaps, medians of the sample estimates. Importantly, when the relative variance condition is met, the exact distribution of the ratio of normals is shown to be approximately normal. Consequently, some idea of the appropriateness of the various measures can be deduced by comparing the locations of modal estimates with estimates of the means and medians computed from the sample. When the relative-variance condition is met and the normal distribution provides a good approximation to the true distribution, the locations of the separate estimates should be similar. Unfortunately, due to the complex form of the posterior (Hinkley 1969, p. 636, equations 1 and 2) and the number of non-participating households (179 in total), computing posterior modes by a Monte Carlo variant of the EM algorithm (Dempster et al. 1977) as proposed by Chib (1996) is infeasible. Thus, we settle on four measures of central tendency, namely the mean obtained from the output of the Gibbs sample, the median of the Gibbs sample, a posterior-means estimate obtained by replacing the equation coefficients in (7) by their posterior means and the conditional means computed from the mixtures:

$$E\left\{\hat{x}_{ki}\right\} = S_{2}^{-1}\sum^{S_j} E\left\{\hat{x}_{ki}\Big|z^{(i)},\beta^{(i)}\right\}, \quad K=1,2,...,q; \quad i \in c \qquad (8)$$

where the expectations on the right-hand sides are taken with respect to the conditional distributions , $\pi(x_{ki}|z\beta), k = 1, 2, .., q; i \in c$. Some algebra reveals that, as long as the numerator can be safely assumed to be non-zero, these latter distributions are, themselves, conditionally normal, implying that the desired expectations do, in fact, exist. This point is important due to the fact that the measures in (7) provide more accurate estimates than the means obtained directly from the output of the Gibbs sample—a feature of the Gibbs sample predicated on the Rao-Blackwell theorem and illustrated, lucidly, by Gelfand and Smith (1990).

4.5 Regression results

Table 2 reports the results of the probit regression on the (68 households × 3 visits × 7 days milk sales =) 1428 observations. Column one reports definitions and column 2 reports posterior means of the Gibbs sample with implied asymptotic t statistics in parentheses. All but one of the covariates—years of farm experience—are significant at the conventional 5%

significance level, and most of the covariate parameter estimates have marginal significance levels beyond 1%. In addition, reports of the signs of the predicted values of the estimated model suggest that only a small proportion of the observations lie outside their negative (positive) ranges for the non-participating (participating) households.

Table 2. *Probit-equation regression estimates.*

Regressor	Estimate (implied t-statistics)
Number of crossbred cows	0.7184
	(11.1314)
Number of local cows	0.2609
	(5.3243)
Time to the milk group, minutes	−0.0131
	(−5.6077)
Farm experience of household head, years	0.0022
	(0.4294)
Formal schooling of household head, years	0.0701
	(3.7501)
Extension agent visits during the past year	0.2148
	(10.2652)
Constant	−2.2100
	(−10.6799)
	Summary statistics
	Participants
Positive predicted values	63
Negative predicted values	105
	Non-participants
Positive predicted values	14
Negative predicted values	1246

These results suggest that the parsimonious formulation adopted here, with entry postulated to depend on animal assets (local and crossbred animals), knowledge assets (education and visits by extension agents) and location (distance to walk to the milk group), is a good approximation to the actual decision-process affecting entry decisions. Hence, the simple probit model seems suited to gauge an indication of the types of policies that could lead to participation among the non-participating households.

The results in general, but more especially those with respect to crossbreed cows, extension services and local breed cows raise interesting questions about the design of appropriate policies to effect participation, their relative potencies and the relative costs of implementing them, which we consider, below.

4.6 Distance to market estimates

In considering participation policy, we confine attention to the number of crossbred milking cows in the household, the number of local breed milking cows and the number of visits by extension agents that the household experienced during the 12 months preceding the survey. The focus is restricted primarily due to space limitations, but these four quantities are, perhaps, the most interesting ones due to the fact that they may be readily changed in the

short term. In reporting the results we rearrange the (1248) observations corresponding to the non-participating households so that the first observation in the set corresponds to the household that is 'nearest' to the market and the last observation is the one that is 'farthest' from the market; where 'near' and 'far' are defined with reference to the units of measurement of the covariate in question (Table 3).

Table 3. *Distance to market estimates.*

Regressor	Estimate (implied t-statistic)
Number of crossbred cows	2.4758
	(1.0852)
Number of local cows	6.8261
	(1.0852)
Extension agent visits during the past year	8.2819
	(3.6544)

With the distance estimates reported in ascending order the three graphs have the following conventions: Households with positive requirements are distant from the market, households with zero requirements are located at the market perimeter and households with negative requirements are within the market boundary. Preliminary plots of the four measures of distance (the Gibbs-sample means, the Gibbs-sample medians, the posterior-means estimates, and the conditional means estimates obtained by the Rao-Blackwell theorem) reveal that each of the estimates are virtually indistinguishable from each other. This observation suggests that the 'relative variance condition' is met so that the posterior distributions are 'almost normal'. Hence, either the mean or the median estimates should suffice as accurate estimates of the distance quantities. Figure 3 reports estimates of crossbred cow requirements. With the Gibbs-sample medians as reference points, there are only three households that are (resource-sufficient) within the market boundary; each of the remaining households has a deficiency of crossbreed cows. This observation is important because it identifies crossbred cow use as an (almost) homogeneously deficient factor across non-participants. Across the entire set of censored observations the median requirement for entry is an addition of 2.48 crossbred cows; the maximum additional requirement (the household farthest from the market) is 5.07; and the minimum requirement is –1.32, which is the household with the greatest 'excess'.

Figure 3. *Crossbred cow distance to market estimates*.

Turning to local cow requirements (Figure 4), we focus attentions again on the Gibbs-sample medians. Average household ownership of indigenous milking cows at the Ilu-Kura and Mirti

sites amount to 1.49 and 1.31 animals, respectively. The maximum median requirement is 13.95 animals and the minimum requirement is –3.56 animals—three of the households have an excess of local breed cows. The median requirement across the non-participating households is 6.82 animals.

Figure 4. *Local breed cow distance to market estimates*.

Results for the number of visits by extension agents are reported in Figure 5. The Gibbs-sample medians are much closer than in Figure 4, but we will use the median estimates as the reference points. Average number of visits at the Ilu-Kura site amount to 1.82/household per year and at the Mirti site amount to 0.36/household per year. From Figure 5, we can deduce that the household closest to the market has an excess of 4.41 visits, and the household farthest from the market requires an additional 16.99 visits before it would enter the market. Hence, the distribution of requirements across the households is more varied than the animal inputs requirements. The estimated median additional requirement in the censor set is 8.28 visits/household per year, which reflects a substantial increase over current levels. Whether this strategy represents a practical alternative remains to be seen. Further work is needed to establish the best form of extension services to provide and determine whether their provision within groups of farmers, rather than individually, is useful. Only then can the precise costs involved in administering extension services be ascertained and its potential as a viable, market-precipitating policy be established.

Figure 5. *Visits by agents distance to market estimates*.

4.7 Conclusions

Collectively the results demonstrate three conclusions. First, standard probit analysis of the participation data provides a useful and informative vehicle for deriving policy estimates. Second, useful quantities for policy analysis are derived simply and robustly as a by-product of the data augmentation step in a Gibbs sampling. Third, the results suggest that on average 2.48 crossbred cows, 6.28 local breed cows and 8.68 visits by extension agents/household per year are the primary measures upon which extension agents and policy planners should focus attentions.

5 Tobit estimation of milk sales decisions

Although probit estimation of the desired 'distance' measures seems to be a fruitful avenue for initial investigation, it suffers a number of limitations. One limitation that seems important here in the context of examining marketable surplus data from the households is that the probit model ignores potentially important information contained in the sales data. In this section we explore the uses of the milk sales data for deriving inferences about entry levels and critical levels of the three key covariates—crossbreed animals, local breed animals and visits by extension agents—as the key precipitators for promoting entry among the subsistence households.

The Tobit procedure is motivated in three steps. First, household maximisation is formalised. Second, relaxing the non-negativity restriction on marketable surplus, a set of latent values are implied for the non-participating households. Third, because we observe the value zero for these households rather than the latent quantities, the data are censored and Tobit estimation is relevant. Here we present the main features of the estimation procedure; details of the procedure are presented in Chib (1992).

5.1 Conceptual model

Let $\Phi_i(\cdot)$ denote the level of a maximand (a quantity which is to be maximised) of interest in household i (say, the level of expected utility); let $\varphi_i(\cdot)$ denote its first-order partial derivative with respect to variable, v_i (the level of marketable surplus from the household); and let $x_i \equiv (x_{1i}, x_{2i}, \ldots x_{qi})$ denote the vector of covariates in question. Across each of the households i = 1, 2, ..., N, we are concerned with the problem:

$$\max_{v_i} \Phi_i(v_i, x_i) \qquad \text{subject to} \qquad v_i \geq 0 \qquad (9)$$

the derivative condition on the objective function

$$\varphi_i(v_i, x_i) \leq 0 \qquad (10)$$

the non-negativity restriction on marketable surplus:

$$v_i \geq 0 \qquad (11)$$

and the complementary-slackness condition:

$$\varphi_i(v_i, x_i) v_i = 0 \qquad (12)$$

Ignoring the restriction in (11) for the moment and assuming strict equality in (10), a first-order MacLaurin-series expansion in the left-hand side yields:

$$\varphi_i + \varphi_i v_i + \sum_{l=1}^{q} \varphi_{xli} x_{li} = 0 \qquad (13)$$

where the function φ_i and the partial derivatives φ_i and φ_{xki}, $k = 1, 2 \dots, q$, are evaluated at the point $v_i = 0$, $x_i = 0$. Accordingly, we have a (locally) valid expression relating the household's choice of v_i to the levels of the covariates, x_{ki}, $k = 1, 2, ..., q$, in the linear equation:

$$v_i = \beta_0 + \sum_{l=1}^{q} \beta_l x_{li}, \qquad i = 1,2,...,N \qquad (14)$$

where $\beta_0 \equiv -\varphi_i \varphi_{vi} -1$ and $\beta_k \equiv -\varphi_{xki} \varphi_{vi} -1$, $k=1, 2 \dots, m$. But, when v_i is negative we actually observe zero and, therefore, the relevant statistical framework is the censored regression model:

$$z_i = \beta_0 + \sum_{l=1}^{q} \beta_l x_{li} + \varepsilon_i \qquad i = 1,2,...,N \qquad (15)$$

where $\varepsilon_i \sim N(0,\sigma^2)$ and we observe $y_i = \max\{z_i,0\}$.

Once again, although interest resides with the parameters in (15), fundamental concern lies with the levels of the covariates that are required for participation in the market, that is, the measures beyond which positive marketable surplus is implied for the non-participants in the (censor) set $c \equiv \{i|z_i \le 0\}$. Just as we did for the probit specification in (7), we can derive estimates of the quantities of interest through a transformation of the estimation equation. In terms of the model in (15), we have 'distance' estimates as follows:

$$\hat{x}_{ki} = \frac{\beta_0 + \sum_{j=1}^{q} \beta_j x_{ji} + \varepsilon_i}{-\beta_k} - x_{ki}, \ k=1,2,...,q, \ i \in c \qquad (16)$$

The covariates upon which we focus attentions include those considered in the participation exercise in the previous chapter, namely a modern production practice (crossbred cow use), a traditional production practice (indigenous cow use), three intellectual-capital-forming variables (experience, education, extension), and the provision of infrastructure (as measured by time to transport milk to market).

5.2 The Tobit algorithm

In presenting the estimation algorithm for the Tobit model, we note that the Tobit set-up is very similar to the probit model. But this assertion is also true of each of the models to follow. Hence, to conserve notation and remove attention from mathematical detail, we will henceforth note only the dependence of the full conditional densities on latent data and parameters and use the symbols z and θ, respectively, to signify these groups of unknown quantities in the joint posterior of interest.

In the case of the Tobit model, the latent data are the vector arrangement of non-positive quantities z_i, for each household $i \, \varepsilon \, c \, \{i|y_i = 0\}$. The parameters, on the other hand, are $\theta \equiv (\beta, \sigma)'$, with the first vector, β, specifying a column vector of coefficients and the second parameter, σ, specifying the standard deviation of the error variance.

Here the latent data, arranged in the column vector z are important for two reasons. First, as

demonstrated incisively by Chib (1992), by augmenting the joint posterior density with these unknown quantities, integrations implied by the Tobit truncation are no longer required, lending the posterior to evaluation simply and, straightforwardly, through data augmentation and Gibbs sampling. Second, because these latent quantities are restricted to lie in the negative segment of the real line, a negative level of marketable surplus is implied. Occasionally, and presently, with a slight abuse of notation, we will use the symbol z to denote both observable and latent data and occasionally, as previously in the case of probit estimation, we will use z to denote only latent quantities. The specification in question should be clear from its context.

With these conventions at hand, the problem of deriving inferences for both z and q can be considered in terms of the regression model:

$$z = x\beta + u \tag{17}$$

where $z \equiv (z_1, z_2, ..., z_N)'$ are the observed and latent data on marketed surplus; $x \equiv (x_1', x_2', ..., x_N')$, $x_1 \equiv (x_{11}, x_{12}, ..., x_{1q})$, $x_2 \equiv (x_{21}, x_{22}, ..., x_{2q})$, ..., $x_N \equiv (x_{N1}, x_{N2}, ..., x_{Nq})$ are observations on the covariates; $\beta \equiv (\beta_1, \beta_2, ..., \beta_q)'$ are the parameters depicting the effects of changes in the covariates on marketed surplus; and the error vector $u \equiv (u_1, u_2, ..., u_N)'$ is assumed to have the normal distribution $N(0_N, \sigma^2 I_N)$, where 0_N denotes the N-dimensional null vector and I_N is the $N \times N$ identity matrix. Compared to the set-up in the previous section, we have introduced one additional parameter, namely the error variance σ, and so the estimation algorithm contains one additional step.

We use a conventional non-informative prior for the unknown parameters, namely $\pi (\beta, \sigma, z)$ $\mu \sigma^{-1}$. It follows, as demonstrated in Chib (1992), that the introduction of the latent data generates a posterior which, while intractable, has component conditional distributions with well-known forms.

In particular, in terms of the current notation, these conditional distributions can be written as:

$$
\begin{aligned}
z|\beta,\sigma &\sim \text{truncated-normal } (Ez, Vz) \\
\beta|\sigma,z &\sim \text{normal } (E\beta, V\beta) \\
\sigma|\beta,z &\sim \text{inverse-gamma}(v, s^2),
\end{aligned} \tag{18}
$$

where $E_z \equiv x\beta$, $V_z \equiv \sigma^2 I_N$; $E\beta \equiv (x'x)^{-1}x'z$, $V\beta \equiv \sigma^2 (x'x)^{-1}$; $v \equiv N + 1 - K$ and $s^2 \equiv (z - x\beta)'(z - x\beta)/v$. Note, once again, that it is simple to sample from multivariate-normal, truncated-normal and inverted-gamma distributions. Consequently, simulations from the joint posterior can be undertaken by sampling sequentially from the respective distributions. The algorithm is very similar to the probit algorithm presented in chapter 4, with the alteration that one additional step must be included in order to simulate the draw from the inverse-gamma distribution, and one additional starting value must be inserted into the first step in the algorithm.

Once again, the resulting outputs, $\{\sigma^{(s)}\ s = 1, 2, ..., S^2\}$ $\{\beta^{(s)}\ s = 1, 2, ..., S^2\}$ and $\{z^{(s)}\ s = 1, 2, ..., S^2\}$, where S^2 is a number of reasonable magnitude, can be used to plot densities or draw inferences about the likelihood that any of the unknown quantities lie within a specified interval. In the results reported below, the algorithm is run for a 'burn-in phase' of $S^1 = 2000$ observations followed by a 'collection phase' of $S^2 = 2000$ observations.

The key aspect of the Tobit algorithm that is worth re-emphasising is its production of the latent quantities for the marketable surpluses for the non-participating households. By definition of the Tobit specification, these quantities are negative random variables that specify

the amount of marketable surplus (in this case daily sales of fluid milk) by which the households in question are deficient. Unlike the latent specification in the probit model, the dependent variable in (17) takes on positive and zero values. When a zero value is observed, we assume this to imply that the household in question, rather than possessing an excess of the marketable product, actually has a demand for the commodity (that is, a negative supply). Hence, sales quantities are left-censored at zero. This simple observation is developed further in Figure 6.

Figure 6 depicts the utility-maximising household-supply decision. Utility (which is, of course, latent or unobservable) is depicted on the vertical axis and the potential sales quantity is depicted on the horizontal axis. For two households (households i and j) one household maximises utility by producing a positive sales quantity (q_j) whereas the second finds utility maximised in the negative segment of the real line over the supply quantity (q_i). Unlike the first household, the second household's implicit supply quantity is unobserved and latent. In Figure 6 the quantity z_i is used to represent this latent value. This value is very important for policy purposes because it provides a simple and highly intuitive quantity with which to measure a household's distance from market (d_i). As such, the values $z_i = \delta_i$, for i ϵ (the censor set) $c \equiv \{$ $i|q_i = 0 \}$ are an important part of the estimation exercise. In the section that follows we show how they can be used to simplify the estimation problems arising due to censoring in the sales data and latency arising in the probit regression. Hence, these latent data represent another metric by which we may measure a deficiency in the non-participating households.

Figure 6. *Rationalising distance to market in the Tobit regression*.

5.3 Results

Table 4 reports results of the estimation. All but one of the covariates (experience) is significant at the 5% level. Thus, each of the other covariates has a significant impact on marketable surplus and, therefore, entry into the milk market. Focusing on the parameter estimates themselves, the addition of one crossbred cow raises surplus by about 4.4 litres of milk per day and the addition of one local cow increases surplus by about 1.8 litres—a clear and obvious difference between the modern and the traditional production techniques. Distance to market on the other hand causes surplus to decline, and we estimate that for each

one-hour reduction in return time to walk to the milk group, marketable surplus increases by about 3.5 litres. Of the capital-forming variables, (experience, education and extension) education and visits by an extension agent are significant, but surplus is unresponsive to farm experience. The estimates of the responses to education and extension are, perhaps, more important for our study because these variables are potentially more likely to be directly affected by policy. For each additional year of formal schooling of the farm decision-maker, daily marketable surplus increases by about 0.30 litres and, for each additional visit by an extension agent, increases by almost 1.0 litre. The summary statistics suggest a reasonable amount of fit given the high proportion of censoring in the sample—approximately 85% are non-participants.

Table 4. *Marketable surplus Tobit-equation estimates.*

Regressor	Estimate (standard error)
Number of crossbred cows	4.43
	(0.38)
Number of local cows	1.81
	(0.26)
Time to the milk group, minutes	–0.06
	(0.01)
Farm experience of household head, years	0.0027
	(0.0233)
Formal schooling of household head, years	0.28
	(0.10)
Extension agent visits during the past year	0.94
	(0.11)
Constant	–12.40
	(1.39)
Square root of the variance	27.47
	(3.98)
	Summary statistics
	Uncensored observations
R^2	0.35
Positive predicted values	63
Negative predicted values	105
	Censored observations
R^2	0.98
Positive predicted values	21
Negative predicted values	1239

Turning to the distance measures, Table 5 reports point estimates of the 'distance' statistics (equation 16). In order to effect entry, the representative non-participant must increase surplus by 9.8 litres of milk per day. Such an increase, it appears, could be effected by a variety of techniques, including additions to the milking herd of 2.2 crossbred animals or, instead, by 6.4 local cows, a feasible but nonetheless substantial increase in productive assets. Of the remaining covariates for which the distance estimates are significant, entry could also be effected by reducing transport time by almost two hours or by increasing the frequency of extension visits to around 10/household per year.

Table 5. *Distance estimates.*

	Estimate (standard error)
Marketable surplus	−9.81
	(5.63)
Number of crossbred cows	2.52
	(0.13)
Number of indigenous cows	6.45
	(0.67)
Time to the milk group, minutes	−114.26
	(33.50)
Farm experience of household head, years	−757.12
	(58289.48)
Formal schooling of household head, years	45.26
	(444.96)
Extension agent visits during the past year	10.43
	(0.91)

5.4 Conclusions

The results of the current investigation, as well as those of the probit specification in the previous section, suggest a clear message: Institutional innovations by themselves are insufficient to catalyse entry; they must be accompanied by a mix of other inputs including infrastructure, knowledge and asset accumulation in the household. Although it is not surprising that milk groups increase the participation of smallholders in fluid milk markets in Ethiopia's highlands, our empirical results provide insights about how to promote further market participation by smallholder producers. Locating groups so as to minimise the time required to market milk increases the number of participating producers and the level of marketable surplus. Given the difficulty and cost of providing crossbred animals (as experienced by such heifer-loan schemes as Heifer Project International in other parts of Africa), investment in infrastructure such as the milk groups provides a low cost mechanism for increasing smallholder participation and furthering the integration of traditional producers into agro-industrial systems. These results are likely to hold relevance for other perishable and time-constrained agricultural products, such as winter vegetables, cut flowers etc. and, perhaps, for a wider and broader set of circumstances.

6 Simultaneous estimation of production and selling decisions

The results of the two previous exercises are striking in their confirmation of similar ranges of resource deficiencies across the households. But both of these approaches ignore potentially important information, such as the correlation across production and selling decisions that the household production theory tells us is important and is relevant for policy investigations concerning subsistence producers. In this section we extend the frameworks of the previous sections in order to investigate this issue.

6.1 Conceptual framework

In linking household production theory to the two-equation Tobit model, two avenues beckon. One is the development of an explicit household production model that motivates Becker-type 'cost-of-time' arguments as principal impediments to participation, develops formally an explicit transactions-costs function that depends on household characteristics and investigates identification and estimation in the context of the data. An alternative, less explicit formulation that proves adequate for our purposes, considers maximisation in the context of choice across the level of milk output and the level of marketable surplus and derives a statistical model with informal, albeit less explicit transactions-costs underpinnings.

Henceforth, we will use the conventions of upper-case bold font denoting matrices and vectors. The development of the two-equation model is, in some senses, very similar to the single-equation approach in the previous section. But the introduction of production effects does introduce a few complications.

Let $\boldsymbol{\pi}_i(\cdot)$ denote the level of a maximand of interest in household i (for example, the level of expected utility); let $\pi_{si}(\cdot)$ and $\pi_{pi}(\cdot)$ denote its first-order partial derivatives with respect to the levels of marketable surplus and household production in the choice vector $y_i \equiv (y_{si}, y_{pi})$; and let $xi \equiv (x_{i0}, x_{i1}, x_{i2}, \dots, x_{iq})$ denote the covariates in question. Across the households i = 1, 2, …, N, we are concerned with the problem:

$$\max_{y_{si}, y_{pi}} \Pi_i(y_i \mid x_i) \qquad \text{subject to } 0 \le y_{si} \le y_{pi} \qquad (19)$$

which states simply that households choose production and marketable surpluses subject to both quantities being positive and marketable surplus being less than the amount of home production (none of the households in the sample purchase milk or dairy products). The Lagrangean for this problem is:

$$\ell(y_i, \lambda | x_i) \equiv \Pi_i(y_i | x_i) + \lambda(y_{pi} - y_{si})$$

(20)

and the corresponding Kuhn-Tucker conditions are:

$$\partial\ell(\cdot)/\partial y_{pi} \leq 0, \quad y_{pi} \geq 0, \quad y_{pi}\partial\ell(\cdot)/\partial y_{pi} = 0$$

(21)

$$\partial\ell(\cdot)/\partial y_{si} \leq 0, \quad y_{si} \geq 0, \quad y_{si}\partial\ell(\cdot)/\partial y_{si} = 0$$

(22)

$$\partial\ell(\cdot)/\partial\lambda \geq 0, \quad \lambda \geq 0, \quad \lambda\partial\ell(\cdot)/\partial\lambda = 0$$

(23)

One narrowing of focus, and consequent simplification, arises from looking at the production and marketable surplus data: All of the households in the sample have strictly positive production and also, at every observation, production exceeds sales. Hence, we assume an interior solution with respect to output, but not for marketable surplus; assume that production levels are non-binding and, therefore, focus attentions on the complementary-slackness condition, equation (23). This condition is fundamental to the theory, the procedures that follow and, indeed, the entire contribution. Links to the Tobit regression become clear when it is recognised that this complementary-slackness condition implies a censored regression on marketable surplus, with left-censoring when the household chooses not to market product. Estimation of this complementary-slackness condition is intractable, so, let us relax the non-negativity constraint, momentarily. In other words, instead of conditions (21)–(23), let us consider the two first-order conditions and the third:

$$\partial\ell(\cdot)/\partial y_{ji} = 0, j = s, p; \qquad i = 1,2,...,N$$

(24)

where, for some of the households in the sample, a negative level of marketable 'surplus' is implied. Once again, a Taylor series expansion at the point $(y_i, x_i)' = 0$ in the left sides of (24) yields:

$$\begin{pmatrix} \alpha_{ss} & \alpha_{sp} \\ \alpha_{ps} & \alpha_{pp} \end{pmatrix}\begin{pmatrix} y_{si} \\ y_{pi} \end{pmatrix} + \begin{pmatrix} \gamma_{s0} & \gamma_{s1} & \cdots & \gamma_{sm} \\ \gamma_{p0} & \gamma_{p1} & \cdots & \gamma_{pm} \end{pmatrix}\begin{pmatrix} x_{0i} \\ x_{1i} \\ \cdots \\ x_{mi} \end{pmatrix} = 0 \quad (25)$$

or, from solving the system:

$$y_{ji} = \beta_{j0} + \sum^q \beta_{jk}x_{ki}, j = s, p; \quad i = 1,2,...,N$$

(26)

where, the a's denote partial derivatives of (24) with respect to the endogenous variables; the g's denote partial derivatives with respect to the characteristics; for each k = 1, 2, ..., q, $\beta_{sk} \equiv -(y_{sk}\alpha_{pp} - y_{pk}\alpha_{sp})/(\alpha_{ss}\alpha_{pp} - \alpha_{ps}\alpha_{sp}), \beta_{pk} \equiv -(y_{pk}\alpha_{ss} - y_{sk}\alpha_{ps})/(\alpha_{ss}\alpha_{pp} - \alpha_{ps}\alpha_{sp})$; and we set $x_{0i} = 1$ for all i = 1, 2, ..., N.

The marketable surplus (equation 26) implies a Tobit regression when it is recognised that some of the quantities on the left-hand side will be negative (recall that we relaxed the non-negativity restriction in the process of deriving equation 26), in which case, for all $i \in c \equiv \{ i | y_{si} = 0 \}$ we observe $y_{si} = 0$. Thus, the data are left-censored at zero and Tobit estimation is relevant.

6.2 Two-equation Tobit model

We are concerned with the model:

$$z_{ji} = \beta_{j0} + \sum_{k=1}^{m} \beta_{jk} x_{ki} + u_{ji}, \qquad j = s, p; \ i = 1,2,...,N \qquad (27)$$

where $(u_{si}, u_{pi}) \sim N(0, \Sigma)$ and, for each $i = 1, 2 \dots N$, we observe $y_{si} = \max\{z_{si}, 0\}$. One justification for the distributional assumption is that, without a more precise rationale, the (omitted) remainder terms in the Taylor-series approximation are likely to be normally distributed. Thus, the extension to Chib (1992) is the addition of a second equation with, possibly, non-negligible covariance through $\Sigma(2 \times 2)$. This extension is more than academic due to the presumption that output and marketable surplus equation errors are positively correlated. Hence, the desire to account for all of the relevant information affecting inferences mandates that this covariance structure be accounted for. Omitting correlation in (27) could lead to miss-specification bias.

The system can be written compactly as:

$$Z = XB + U \qquad (28)$$

where $Z_{(N \times 2)} \equiv (z_s, z_p)$, $z_s \equiv (z_{s1}, z_{s2}, \dots, z_{sN})'$, $z_p \equiv (z_{p1}, z_{p2}, \dots, z_{pN})'$ are the latent and observed data on marketable surplus and the observed (uncensored) output data; $X_{(N \times (m + 1))} \equiv (x_0, x_1,\dots, x_m)$, $x_0 \equiv (1, 1, \dots, 1)'$, $x_1 \equiv (x_{01}, x_{11}, x_{12}, \dots, x_{1N})'$, $x2 \equiv (x_{02}, x_{12}, x_{22}, \dots, x_{N2})'$, $\dots, x_q \equiv (x_{0q}, x_{1q}, x_{2q}, \dots, x_{Nq})'$ are observations on the regressors; $B_{((m + 1) \times 2)}$ (B_s, B_p), $B_s \equiv (B_{0s}, B_{1s}, \dots, B_{qs})'$, $B_p \equiv (B_{0p}, B_{1p}, \dots, B_{qp})'$ are coefficients of the regressors in the two equations; $U_{(N \times 2)} \sim N(0, \Sigma Ä I_N)$; and we reorder the observations so that the first N_1 observations in Z_s refer to the households with positive marketable surplus and the last $N_2 = N - N_1$ observations are derived from the censor set, c. Following Chib (1992), developments simplify considerably by augmenting the likelihood with the latent data.

The essential observations (Chib 1992, p. 88, equation 16) are three. First, when the observed data $Y(N \times 2) \equiv (y_s, y_p)$, $y_s \equiv (y_{1s}, y_{2s}, \dots, y_{Ns})'$, $y_p \equiv (y_{1p}, y_{2p}, \dots, y_{Np})'$ are augmented by these unknown or latent components of Z, the posterior distributions conditioned by (Y, Z), on the one hand, and Z, on the other hand, are identical asymptotically. Second, as opposed to Y, because no censoring is involved in Z, the latter matrix generates a tractable form for the posterior. Third, although N_2 elements of Z are unknown, these unknown components have a well-known distribution which is trivial to sample from. Consequently, the distributions of the parameters and the latent data can be combined, stepwise, to generate a Markov chain, facilitating posterior inference and policy analysis.

6.3 Estimation

To formalise things somewhat, note that with the data reordering now in place, the partitions in $z_s \equiv (z_{s1}, z_{s2})'$ refer respectively to the blocks of observed and censored data; $z_{s1} \equiv y_{s1}$; $z_{s2} \leq 0$

$= y_{s2}$, the latter, of course, denote the censored observations on marketable surplus; and we have $z_p \equiv y_p$ and, thus, $z_{p1} \equiv y_{p1}$ and $z_{p2} \equiv y_{p2}$. Acknowledging these partitions we rewrite (28) as:

$$\begin{pmatrix} z_{s1} & z_{p1} \\ z_{s2} & z_{p2} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} (B_s, B_p) + \begin{pmatrix} u_{s1} & u_{p1} \\ u_{s2} & u_{p2} \end{pmatrix}; \tag{29}$$

partition the covariance matrix accordingly:

$$\Sigma = \begin{pmatrix} \sum_{ss} & \sum_{sp} \\ \sum_{ps} & \sum_{pp} \end{pmatrix} \tag{30}$$

and note that the situation is only slightly more complicated than in Chib (1992) due to the fact that we are working with multivariate data. It follows, exploiting standard developments of the multivariate-normal model that, under the diffuse prior $\pi(B, S, z_{s2}) \mu| \Sigma |^{-(m + 2)/2}$ (Drèze and Richard 1983, pp. 539–544; Zellner 1996, pp. 224–247) the complete posterior distribution (for the combined latent data and the parameters) has component, conditional distributions:

$$z_{s2} | B, \Sigma \quad \sim \quad \text{Truncated-Normal}(E_{zs2}, V_{zs2})$$

$$B | \Sigma, Z \quad \sim \quad \text{Normal}(E_B, V_B) \tag{31}$$

$$\Sigma | Z, B \quad \sim \quad \text{Inverted-Wishart}(W, v)$$

where, in terms of equations (28)–(30), $E_{zs2} \equiv X_2 B_s + (z_{p2} - X_2 B_p) \Sigma_{pp}^{-1} \Sigma_{sp}$; $V_{zs2} \equiv \Sigma_{ss} - \Sigma_{sp} \Sigma_{pp-1} \Sigma_{sp}$; $E_B \equiv (X'X)^{-1} X'Z$; $V_B \equiv \Sigma \ddot{A} (X'X)^{-1}$; $W \equiv Z'(I_N - X(X'X)^{-1}X')Z$; $v \circ N - (m + 1) + 2 + 1$ (Zellner 1996, pp. 224–227, 380–382). The essential feature of (31) that is important for estimation is that each of the relevant distributions is easy to sample from. Consequently, draws from the posterior distribution are simulated through an algorithm that draws sequentially from the latent data z and the parameter matrix $\theta \equiv (B, \Sigma)'$ and, once again, the outputs { $z_{s2}(s)$ s = 1, 2, ..., S_2 }, { $B^{(s)}$ s = 1, 2, ..., S_2 } and { $\Sigma^{(s)}$ s = 1, 2, ..., S_2 } can be used to derive inferences about any unknown quantity of interest. The draw from the truncated normal is implemented efficiently by adjusting a draw from the standard uniform distribution using the probability integral transform (Mood et al. 1974, p. 202). In each of the models discussed below, the algorithm is run for a convergence phase of 5000 iterations. Convergence is ascertained by commencing the iterations at different starting values, plotting posterior densities of interest at various intervals and establishing the number of iterations at which the plotted distributions are essentially indistinguishable. Experiments suggest that convergence is achieved after about 500 iterations. After the burn-in phase of 5000 iterations, a Gibbs sample of 5000 observations is collected. Each estimation took approximately 60 minutes of real time on a DELL™ Optiplex G1 running a Pentium II processor at 133 megahertz with 128 megabytes of RAM. Commands are executed in MATLAB™ version 5.1.0.421 and are available upon request.

6.4 Locating key regressors

In this exercise we combine the production and sales data. However, unlike the sales data, which are available seven days prior to each of the three visits to the households, the production data are only recorded at the particular day that the household is surveyed. Therefore, the total number of observations in the sample is reduced to 204 (= 68 households× 3 visits).

We begin with a subset of 14 regressors, namely 3 measures of modern production; 3 measures of traditional production; the distance variable and 3 measures of transport readiness; 3 measures of knowledge formation; a cumulative measure of previous market involvement; and 1 measure of the household demographic variables. Respectively, the included variables are number of crossbreed milking cows; a binary variable representing ownership of farm equipment; a binary variable denoting credit use; number of indigenous milking cows; land area in pasture; land area in crops; distance to the milk group; ownership of transport equipment; the number of children (usually between 5 and 15 years to handle transport equipment); years of farm experience; years of formal schooling; numbers of visits by an extension agent discussing improved production and marketing techniques; the sum of six binary variables characterising previous market involvement (specifically, whether the household received income from selling (a) grains, (b) fodder, (c) live animals, (d) labour, (e) crafts or (f) other goods and services); and a binary variable denoting the gender of the household head. Given these 14 sources of variation, the model is estimated with constant terms relevant to each of the two study sites—Mirti and Ilu-Kura—yielding a total of 30 equation coefficients and 3 covariance terms. In this specification, multicollinearity appears to be a severe problem as determined by the condition number of the normalised covariate matrix (Greene 2000 p. 280, footnote 4). Indeed, most of the coefficient estimates in this initial specification were insignificant and their values were deemed particularly sensitive to the exclusion of any of the remaining regressors. Some further investigation suggested that the binary variables were particularly troublesome. In a subsequent formulation with the binary regressors excluded, the effects of traditional inputs, pastureland and cropland; the effects of the two transport variables; and the effects of the knowledge accumulator, experience, were insignificant in both equations. Thus, the preferred model is a rather parsimonious specification with seven regressors in each equation, including the number of crossbreed cows being milked; the number of indigenous breed cows being milked; the return distance (in minutes) to transport bucketed, fluid milk to the milk group; the number of years of formal schooling by the household head; the number of visits by an extension agent in the year prior to the survey discussing production and marketing activities; and two site specific dummy variables.

6.5 Results

Table 6 reports posterior means of the regression coefficients and the covariance terms from the 5000 observation Gibbs sample. The limits of the 90% highest-posterior-density regions are reported in parentheses to the left and right of the posterior means (lower limit to the left, upper limit to the right). Only two of the regression coefficients have 90% highest-posterior densities that contain zero. They are, respectively, the effects of distance to the milk group on milk output and the Ilu-Kura site-specific dummy variables in the milk output equation. Each of the remaining coefficients is significant at a 90% or higher level. In addition, each of the covariance terms is estimated to a considerable degree of precision. The mean error variance in the sales equation (18.23) is considerably higher than the mean error variance in the output equation (3.57) and, as we expected, there is a strong, significantly positive correlation between the marketable surplus and output equation errors. Focusing on the point estimates of the regression coefficients, the model predicts that the addition of one crossbred cow causes marketable surplus to rise by about 2.56 litres of milk per day and causes output to rise by about 3.02 litres per day. The addition of one local cow causes output to rise by 1.24 litres and causes sales to rise by 1.10 litres. These results are plausible and suggest that consumption demands for milk in the household are responsive to production shifts. Specifically, with output increments observed to be lower than the corresponding marketable surplus expansion, consumption (which equals production minus marketings) must decline in response to the increases in cow numbers. Distance to the milk group is not a strong determinant of household output, but its impact on sales is considerable. Each one-minute

increase in return time to transport bucketed-fluid milk to the milk group causes marketable surplus to contract by about 0.06 litres. Whether this reduction is important as a policy measure depends ultimately on the costs of reducing transport time and the method—whether infrastructural (improving roads and public access), vehicular (from capital acquisition) or spatial (from the creation of additional milk groups). Turning to the knowledge accumulation variables, education has an important impact on sales, but a lower, less significant impact on output. Number of visits by extension agents, on the other hand, is highly significant in both the sales and output equations. The model predicts that each additional extension agent visit causes daily output and sales to rise by about 0.16 and 0.54 litres, respectively. In addition, these estimates are noticeably precise.

Table 6. *Marketable surplus and milk output equation estimates.*

Covariate	Equation estimates					
	Marketable surplus			Milk output		
Number of crossbreed cows	(1.42)	2.56	(3.51)	(2.72)	3.02	(3.26)
Number of indigenous cows	(0.14)	1.10	(1.88)	(1.02)	1.24	(1.40)
Visits by an extension agent	(0.20)	0.54	(0.83)	(0.04)	0.16	(0.24)
Distance to the milk group	(−0.10)	−0.06	(−0.03)	(−0.01)	0.00	(0.01)
Years of formal schooling	(0.14)	0.44	(0.71)	(0.00)	0.07	(0.13)
Ilu-Kura	(−9.98)	−6.10	(−3.45)	(−0.81)	−0.09	(0.47)
Mirti	(−4.66)	−9.99	(−6.80)	(−1.67)	−1.03	(−0.53)
	Covariance					
Marketable surplus	(9.58)	18.23	(27.71)	(2.26)	3.65	(4.93)
Milk output	(symmetric)			(3.02)	3.57	(4.06)
			Uncensored observations			
R^2		0.24			0.79	
Positive predicted values		11			25	
Negative predicted values		14			0	
			Censored observations			
R^2		0.71			0.50	
Positive predicted values		3			177	
Negative predicted values		176			2	

In addition to the locations of the posterior means relative to their sample variances, one can gauge the explanatory power of the model from traditional R^2 measures and, in this particular setting, by examining the number of positive predictions in the endogenous variables across the uncensored observations and the number of negative predictions across the censored observations. The summary statistics in the lower part of Table 6 indicate that about 24% of the variation in sales and about 79% of the variation in output is explained by the seven covariates. In addition, of the 25 positive sales values, and the corresponding 25 observations

on output, the model predicts, respectively, that 11 of the sales values and 25 of the output observations are positive. Although these rates of accuracy are not high, they are to be expected in a model with such a high degree of censoring (about 88%). Turning to the censored observations in the lower part of Table 6, we construct an R^2 measure by using the mean values of the draws of the latent variables in the Gibbs sample. Because these values are drawn conditionally on the values of the regression and covariance parameters we should expect a reasonable rate of prediction across the latent data. In fact, about 71% of their variation is explained by the seven covariates. In addition, in the proportion of the output data corresponding to the censored sales observations the model explains about 50% of the variation. Turning to prediction across the 179 censored observations, only 3 of the predicted sales values are positive, the remaining 176 observations are predicted to be negative. In the output equation the model predicts that 177 of the observations on output (99% of the censored sample) are positive and that 2 observations (1%) are negative. These results are satisfactory for a parsimonious specification in which there is a high degree of censoring, and one should be reasonably comfortable with the policy implications drawn from it. The results in general, but more especially those with respect to crossbreed cows, extension services and local breed cows raise interesting questions about the design of appropriate policies to effect participation, their relative potencies and the relative costs of implementing them, which we consider, below.

6.6 Distance to market estimates

In considering participation policy, we confine attentions, once again, to the four quantities that are potentially most interesting, namely the level of latent surplus milk by which the household is deficient, its level of crossbred milking cows, its level of local breed milking cows and the number of visits by an extension agent it absorbed in the 12 months preceding the survey (Figures 7–10). The focus is restricted primarily due to space limitations, but these four quantities are, perhaps, the most interesting ones due to the fact that they may be readily changeable in the short term. In reporting the results we rearrange the 179 censored observations so that the first observation in the censor set is the one that is 'nearest' to the market and the last observation is the one that is 'farthest' from the market, where 'near' and 'far' are defined with reference to the units of measurement of the covariate or latent variable in question. With the distance estimates reported in ascending order the three graphs have the following conventions: households with positive requirements are distant from the market, households with zero requirements are located at the market perimeter and households with negative requirements are within the market boundary. Preliminary plots of the four measures of distance (the Gibbs-sample means, the Gibbs-sample medians, the posterior-means estimates, and the conditional means estimates obtained by Rao-Blackwell theorem) reveal that the Gibbs-sample means and the conditional means estimates are virtually indistinguishable from each other, that the posterior means and Gibbs-sample medians estimates are virtually indistinguishable, but that the two groups of measures diverge somewhat depending on the estimated quantity in the denominator of the distance measure (see equation 7). Hence, Figures 7–10 are plotted with respect to both the Gibbs-sample means and the Gibbs-sample medians.

Figure 7 reports estimates of the latent milk marketable surplus requirements that would cause non-participants to enter the market. The Gibbs-sample means and the Gibbs-sample median values are approximately the same for each household, as they should be because the conditional distribution of the surplus value is normal. Using the Gibbs-sample means as points of reference, the minimum requirement (the household closest to the market) is 1.8 litres of milk per day and the maximum requirement (the household that is farthest from the market) is 15.5 litres. To place these requirements in better perspective, the maximum daily sales and production figures in the sample are 10.2 and 18.5 litres of milk per day, respectively. Thus, many of the sales targets estimates reported in Figure 7 appear credible

and provide a useful reference point against which to consider methods by which these targets can be achieved.

Figure 7. *Distance to market estimates: Milk.*

Figure 8 reports estimates of crossbreed cow requirements. Both the mean and median estimates are very similar across the households—a result arising due to the fact that the 90% confidence interval for crossbred cows in the marketable surplus equation (equation 27) is relatively distant from zero. With the Gibbs-sample means as reference points, there are only three households that are within the market boundary (resource-sufficient); each of the remaining households has a deficiency of crossbreed cows. This observation is important because it identifies crossbreed cow use as an (almost) homogeneously deficient factor across non-participants. Across the entire set of censored observations the mean requirement for entry is an addition of 2.86 crossbreed cows; the maximum additional requirement (the household farthest from the market) is 6.18; and the minimum requirement is –1.12, which is the household with the greatest 'excess'. These requirements should be compared to the mean levels of ownership across the two sample sites, 0.79 animals at Ilu-Kura and 0.71 at Mirti. The censored-sample mean estimate of crossbreed cow requirements represents a relative change in factor intensity at the two sites of 262 and 303 percent, respectively. These quantities represent large, but not infeasible increases in factor intensity.

Figure 8. *Distance to market estimates: Crossbred cows*.

Turning to local cow requirements (Figure 9), we observe that the mean and median Gibbs sample reports are divergent. This occurrence arises due to the fact that the 90% confidence interval for local cow use in the marketable surplus equation (equation 27) lies relatively close to zero. Hence, the normal approximation is unlikely to be accurate in this particular case, and we focus attentions on the Gibbs-sample medians in Figure 9 rather than the sample means. Average household ownership of indigenous milking cows at the Ilu-Kura and Mirti sites amount to 1.49 and 1.31 animals, respectively. The maximum median requirement is 12.67 animals and the minimum requirement is –2.59—three of the households have an excess of local breed cows. The median requirement across the non-participating households, 5.78 animals, reflects relative increases of 288 to 341 percent over the empirically observed holdings. Once again, these requirements reflect substantial increases in input use.

Figure 9. *Distance to market estimates: Local breed cows*.

Results for extension (Figure 10) yield high variability across the sample. The Gibbs-sample means and Gibbs-sample medians are much closer than in Figure 9, but we will use the median estimates as the reference points. Average number of visits amount to 1.82 per household per year at Ilu-Kura and 0.36 at Mirti. From Figure 10, the closest household has an excess of 5.61 visits, and the household farthest from the market requires an additional 26.77 visits before it would enter the market. Hence, the distribution of requirements across the households is more varied than the animal inputs requirements. The estimated median additional requirement in the censor set is 12.33 visits per household per year, which reflects a substantial increase over current levels. Whether this strategy represents a practical alternative remains to be seen. Further work is needed to establish the best form of extension services to provide and determine whether their provision within groups of farmers, rather than individually, is useful. Only then can the precise costs involved in administering extension services be ascertained and its potential as a viable, market-precipitating policy be established.

Figure 10. *Distance to market estimates: visits by extension agents*.

Making valid cost comparisons is hampered by the extreme variability in prices of livestock across locations, transaction dates and quality realisations (crossbreed cow prices increase monotonically with potential milk yield) and the lack of sufficient estimates of the per unit costs of extension visits. Deriving estimates of visits by an extension agent is compounded by the fact that a major cost to the public sector derives from the fixed costs of educating extension agents. Focusing attentions on the animal inputs, representative price bands do exist for the Addis Ababa market (Abebe Misgina, personal communication). Current rates for local breed heifers vary between 800 to 1200 Ethiopian Birr (EB) (US$ 1 = EB 8.32 at May 2001) and for crossbreed heifers between EB 3000 to 5000. Using the midpoints of these ranges, the (pecuniary, variable) costs involved in effecting participation for the representative household using the two, alternative strategies (crossbreed versus local breed animals) amount to EB 11,440 and 5780, respectively. Hence, the local breed strategy appears to dominate, at least on a per unit cost basis. Further inquiry into the precise costs of the three strategies, the interrelationships that exist between them—including the nature of any positive or negative externalities that might exist—and their measurement, leaves considerable scope for additional empirical inquiry.

6.7 Conclusions

We have identified factors and estimated quantities that effect participation in a sample of peri-urban dairy farmers in the Ethiopian highlands. That this sample may be reflective of a wide and broader set of circumstances makes our results interesting in the context of milk market expansion in SSA. Some idiosyncrasies of milk marketing prohibit generalisation, but other features suggest that the procedures developed here are potentially applicable to other situations in which transactions costs prohibit market emergence. At least in the smaller context of the survey region our results are useful for understanding the structure of milk markets; their emergence in the highlands; the institutions and production innovations that create them; and the policies we should enact in order to sustain them. The findings suggest a clear message. The formation of marketing institutions like co-operatives, and the enactment of production innovations like the introduction of crossbreed cows do much to sow the seeds from which formerly latent markets may propagate. But their creation needs, it seems, a healthy stimulus to the stock of human capital, a widening of the knowledge base and dissemination through extension. The formation of milk groups, like the ones inaugurated by the Finnish International Development Association (FINNIDA) in the Ethiopian highlands appear, to the casual observer, to have been successful. But their successes must be evaluated relative to the alternative uses of the resources required to create them. Ultimately, the economic benefits of market creation must be offset against its costs but, at least for the time being, a sample of formerly subsistence farmers have gained considerably since the project's inception. Whether they will be successful in attracting new entrants remains to be seen, but for now, the milk groups evidence the fact that, when production innovations are given a little encouragement through extension, there is much to be gained. In the meantime, work continues on extending the methodology, particularly the use of data augmentation, to a broader set of data, a wider class of markets, and a different set of contexts within which presently latent markets may eventually emerge.

7   Crossbred cow adoption

The results of the last section are important in motivating the importance of including additional, relevant information in the decision model. In the previous chapter we consider the importance of allowing for correlation in the production and marketing decisions and found that, when these pieces of information were included, both the distance-to-market estimates and the equation-system parameter estimates are affected. One piece of additional information that is ignored in the estimations thus far that could credibly affect some of the distance-to-market estimates is the effect on production and sales of the decision to adopt a crossbred animal. A maximum ownership of four animals at the sample sites is obviously scarce and important. It seems natural, therefore, to consider the consequences of adoption of crossbred animals in the related decision-making process. Inclusion of these inputs is likely to affect the related decisions about how much milk to produce and how much of the produced milk to sell. The rationale is simple. Errors in the equations predicting adoption of crossbred animals, production and selling decisions are likely to be correlated. But the institutional motivation for modelling the adoption decision is deeper than this.

7.1 Motivation

Two, recent, hitherto unrelated innovations are fundamental to the expansion of milk markets in SSA. One is the crossbreeding of exotic dairy cattle with grade, indigenous stock and the second is the formation of marketing co-operatives in peri-urban settings. This chapter formulates a multivariate, count-data model to answer the question: What kinds and levels of additional inputs are required to increase milk market participation? This chapter addresses this question through routine application of Gibbs sampling and data augmentation. A trivariate model of crossbred cow adoption (the count data), household milk production (a Gaussian component) and marketable surplus (a Tobit, censored regression component) is used to measure non-participating households' distances from market and make recommendations about the levels of inputs needed to induce entry by non-participants.

In the highlands, perishability, and high search and transportation costs relative to return suggest that milk marketing co-operatives may be promising catalysts for market development. We focus on this study for one main policy reason. The outlook for expansion of dairy production and marketing in SSA is strong given the expectation of continuing strong demand driven by urbanisation and income growth. However, despite this positive outlook, growth in dairy production and marketing by smallholder producers (1–2 cows) has shown wide variation across regions. It is indeed striking that market-oriented smallholder dairy development has not been widespread in SSA outside Kenya, despite exposure to similar

changes in unfavourable national and international policies.

7.2 Dairy innovations in the Ethiopian highlands

The Finnish International Development Association (FINNIDA) undertook and funded an especially relevant smallholder dairy development project in the Ethiopian highlands in 1995–97. The project had two focuses. One, targeted principally at production management practices, was the introduction of grade imported cattle that were crossed with indigenous, local stock. The other, targeted at marketing strategies, was the construction and inauguration of several co-operative dairy enterprises. The aim of the project was to provide the necessary inputs that would catalyse growth and diversification of smallholder activities, particularly the processing of milk into the derivative products such as butter, the local cottage-type cheese (*ayib*) and, although considerably less significant, a local yoghurt-type sour liquid (*irgo*). The first co-operatives were formed in 1995, some two years after the termination of parastatal marketing co-operatives. There are now four. Researchers at the International Livestock Research Institute (ILRI) (formerly the International Livestock Centre for Africa, ILCA) identified the formation of these co-operatives as catalysts for long-term sustainability and growth in the local dairy sector. ILRI researchers targeted two milk groups areas as the basis for studying factors that promote participation. The two target areas are the Ilu-Kura and Mirti peasant associations, respectively, in the northern Shewa and Arsi regions.

An important focal point of livestock research in these areas is the Asella research station. Established in 1967–68, it has been the main source of information on crossbred cattle management practices in the highlands. Crossbreeding of imported, high-yielding animals with grade, indigenous stock began in 1968 and, since then, an extensive set of records on the performance of various crosses have been maintained. Perhaps the most striking feature of the crossbred animals is their enormous potential for increasing per capita production. Results from a pooled set of within-station trials over the period 1968–77 are illustrative. Over a variety of seasons, calving dates and feed regimes, crosses between the local Arsi and Zebu animals with introduced stock produced upwards of 100% increases in output, measured in terms of annual, fat-corrected milk yields (kg) per unit metabolic weight of milking cows (kg) (Kiwuwa et al. 1983). The station and latter on the field trials have confirmed, to varying degrees, the production advantages of crossbreeding. In short, hybrid-vigour has been an important innovation in local herds, sparked enormous increases in per capita milk production, and led consequently to increases in the potential to generate marketable surpluses and enter local milk markets. However, the technological advantage of crossbreeding has brought with it some serious impediments retarding its widespread adoption by local farmers, including, most significantly, susceptibility of crossbred cattle to local diseases (such as anthrax, rinderpest and blackleg) which have forced adopters to change management practices and inevitably has lowered adoption rates. Thus, the actual impacts of crossbreeding on milk market participation are unclear, raising considerable scope for empirical inquiry.

7.3 A single-equation approach to the count data

The data are the panel of observations on the sample of 68 households visited 3 times in 1997 (204 observations that were analysed in the preceding chapter). Primary interest centres on the number of crossbred cows milked in each household and time period. A simple, hierarchical model relating the observed counts to covariates and a random effect per household was investigated. However, because all but one of the covariates remain constant across the time periods, many of the within-household, cross-product matrices are ill-conditioned, leading to convergence problems in the estimation algorithm. For this reason, in this initial exploration we adopt a simplified model in which each observation is assumed to be generated independently from a common distribution. In other words, we assume that, across households, h = 1, 2, …, H (= 68), and time periods, t = 1, 2, …, T (= 3), the counts, y_i, i = 1,

2, … n (= 204) are independent draws from a common distribution. Whether our results are highly sensitive to this assumption remains to be seen. This question remains an important issue for future work for which the results that follow can be considered a starting point. Limitations aside, the 204 observations generate count frequencies of 116, 59, 24 and 5 corresponding to y_i = 0, 1, 2, and 3, respectively, and generate a sample mean (0.60) and a sample variance (0.63) that are close enough so that concerns about over-dispersion appear unwarranted.

A single-equation approach to the data that is readily extendable to multivariate settings evolves from ideas in Albert and Chib (1993). Assume that for each observation ('household'), there exists a latent Gaussian random variable, z_i, representing the (continuous) desired level of the capital stock. Because the observed level of the stock arrives in discrete amounts, j = 0, 1, 2, …, it is natural to constrain the desired level to the domain $j < z_i \leq j + 1$ and model the probability that 'household' i adopts j (a discrete number of) units of the capital stock as being equivalent to the probability that the latent variable z_i is contained within the limits j and j + 1. We assume, in turn, that the desired level of the stock depends on a set of household- and time-specific covariates, leading to the regression relationship:

$$z_i = x_i\beta + u_i, \qquad i = 1,2,...,N \qquad (32)$$

where z_i denotes the latent (desired) level of the capital stock, $x_i \equiv (x_{i1}, x_{i2}, ..., x_{iq})$ are the covariates, $\beta \equiv (\beta_1, \beta_2, ..., \beta_q)'$ denotes the effects of the covariates on the desired level of adoption, and u_i is a random error which we assume is normally distributed with mean 0 and variance σ^2. Defining, $y \equiv (y_1, y_2, ..., y_n)'$ as the vector of observed counts and $x \equiv (x_1', x_2', ..., x_N')'$ as the observations on the covariates, the joint posterior for the unknown quantities σ, β and z has the component, conditional distributions:

$$\sigma | \beta, z, \alpha, y \quad \sim \quad \text{inverted-gama}(v_\sigma, s_\sigma^2)$$

$$\beta | z, \alpha, \sigma, y \quad \sim \quad \text{Normal}(\hat\beta, \text{cov}\,\hat\beta) \qquad (33)$$

$$z | \alpha, \sigma, \beta, y \quad \sim \quad \text{Truncated-normal}(\hat z, \text{cov}\,\hat z)$$

where the definitions appearing in the right-hand sides are, respectively, $v_\sigma = n$, and where each element of $s^2_\sigma \equiv (z-x'(\,z-x/,v_\sigma, \hat\beta \equiv (x'x)^{-1}x'z \text{ cov}\,\hat\beta \equiv \sigma^2\,(x'x)^{-1}, z \equiv x\beta, \text{cov } z \equiv \sigma^2 I_t$ and where each element of $z \equiv (z, z_2,...z_n$ is drawn such that $j < z_i \leq j + 1$ as $y_i = j$. Note that it is straightforward to simulate from inverted-gamma, normal and truncated-normal distributions. Consequently, a Gibbs sampling algorithm can be used to simulate from the joint posterior p (σ, β, z|y) and calculate posterior quantities of interest.

Estimates of the truncated Gaussian approach to the count data are presented in Table 7, column 2, which reports the posterior means of the regression coefficients and, in parentheses, the means relative to their estimated standard errors from the Gibbs sample. The algorithm was implemented with a burn-in phase of 5000 iterations followed by a collection phase of 5000 iterations. Commands are executed in MATLAB™ version 5.1 and absorbed 10 minutes of real time on a Dell workstation with a Pentium II, 133 mhz co-processor and 160 mb of RAM. The definitions of the covariates in column 1 of Table 7 are: 'Gender' ≡ gender of the household head (= 1 if male); 'Distance' ≡ return distance to the milk group in minutes; 'Credit' ≡ credit use (= 1 if used credit); 'Local'≡ number of local breed cows currently being milked; 'Extension' ≡ number of visits by an extension agent in the previous

year discussing production and marketing activities; 'Periodi,' i = 1, 2, 3 ≡ time-specific dummy variable (= 1 if from period i).

Table 7. *Single- and multiple-equations, crossbred cow adoption estimates.*

	Cow numbers			
Effect	Single-equation estimate	Multiple-equation estimates	Milk output	Milk sales
Gender	0.3005	0.2984	0.9658	1.4736
	(2.3457)	(2.3207)	(1.8870)	(0.8504)
Distance	–0.0011	–0.0011	–0.0003	–0.0630
	(–0.6428)	(–0.6133)	(–0.0404)	(–2.3346)
Credit	–0.2597	–0.2502	–1.0322	2.8230
	(–2.2719)	(–2.1678)	(–2.2457)	(1.7480)
Local	–0.2239	–0.2219	0.5249	0.7633
	(–4.8159)	(–4.6764)	(2.7866)	(1.1956)
Extension	0.0657	0.0653	0.4107	0.9090
	(2.6823)	(2.6257)	(4.2162)	(3.2919)
Period1	1.2246	1.2257	1.0850	–10.1126
	(7.0712)	(7.2158)	(1.6194)	(–3.0531)
Period2	1.2675	1.2714	1.6553	–7.7757
	(7.5214)	(7.3856)	(2.4381)	(–2.6286)
Period3	1.2235	1.2298	2.1778	–6.9460
	(7.0143)	(7.0355)	(3.1391)	(–2.4466)
	Variance	Covariance		
Cows	0.4295	0.4386	1.5386	1.5678
	(8.3870)	(8.3122)	(8.3642)	(3.2931)
Output			8.2035	9.1609
			(9.9864)	(4.8287)
Sales				27.7353
				(2.3433)

Although our main interest in the model lies in its use as a tool for incorporating other household decisions, several features of the single-equation results are noteworthy. First, households with male heads are more likely to adopt crossbred animals. Second, there is significant negative dependence on prior credit use, implying that prior capital commitments constrain households' abilities to purchase crossbreed animals. Third, there is a significant negative dependence on the number of local breed cows, implying that the households in question view local and crossbreed cows as substitutes in the milk production process. Fourth, there is evidence that extension visits promote adoption.

7.4 Simultaneous adoption, production and sales decisions

The main criticism of the single-equation model is its lack of account for other decision variables affecting household behaviour. For example, the classical household production model (Singh et al. 1986) assumes that households jointly make decisions about three quantities: how much time to devote to production activities, how much of the produce to sell, and how much of the produce to consume. Under reasonable assumptions, and with the aid of a few simple transformations, the time allocation can be subsumed within the production

decision.

Assuming that the household does not waste produced goods, the dependence between produced, consumed and sold quantities allows us to ignore one variable, leaving us with endogenous choice over two quantities. In the current setting it is natural to consider production and sales quantities, but especially their relationship to the level of adoption. The logic is simple. Errors in equations depicting the number of crossbred cows milked, the output that they produce and the level of milk output that is sold will be correlated. This simple fact is the main motivation for developing the Gaussian, latent-variable estimator for crossbred cow adoption. Now, it is relatively straightforward to include adoption decisions in a multivariate system. This task is simplified by making use of standard developments in the multivariate-normal linear model (Zellner 1996, pp. 224–227). The difference between the traditional multivariate model (Zellner 1996, equation 8.1) and the present set-up is the presence of latent (or missing data) in the crossbred cow adoption equations and in the milk sales equation corresponding to zero sales, since observations on the non-participating households are censored at zero. These features of the problem are easily accommodated by embedding ideas in Albert and Chib (1993) and in Chib (1993) in the system:

$$Y = XB + U \tag{34}$$

where $Y \equiv (y_1, y_2, \ldots, y_m)$, $y_1 \equiv (y_{11}, y_{21}, \ldots, y_{N1})'$, $y_2 \equiv (y_{12}, y_{22}, \ldots, y_{N2})'$, ..., $ym \equiv (y_{1m}, y_{2m}, \ldots, y_{Nm})'$ are observations, respectively, on crossbred cows, daily milk output (litres) and daily milk sales (litres) to the milk co-operative; $X \equiv (x_1, x_2, \ldots, x_q)$, $x_1 \equiv (x_{11}, x_{21}, \ldots, x_{N1})'$, $x_2 \equiv (x_{12}, x_{22}, \ldots, x_{N2})'$, ..., $x_q \equiv (x_{1q}, x_{2q}, \ldots, x_{nq})$ are observations on the covariates; $B \equiv (\beta_1, \beta_2, \ldots, \beta_m)$, $\beta_1 \equiv (\beta_{11}, \beta_{21}, \ldots, \beta_{q1})'$, $\beta_2 \equiv (\beta_{12}, \beta_{22}, \ldots, \beta_{q2})$, ..., $\beta_m \equiv (\beta_{1m}, \beta_{2m}, \ldots, \beta_{qm})'$ denote the effects of the covariates on the adoption, production and sales decisions; and $U \equiv (u_1, u_2, \ldots, u_m)$, $u_1 \equiv (u_{11}, u_{21}, \ldots, u_{N1})'$, $u_2 \equiv (u_{12}, u_{22}, \ldots, u_{N2})$, ..., $u_m \equiv (u_{1m}, u_{2m}, \ldots, u_{Nm})$ is a matrix of errors with rows assumed to be normally distributed with mean a $1 \times m$ null matrix and covariance Σ, an $m \times m$ positive definite symmetric matrix. Considerable interest centres on the empirical magnitude of Σ, reflecting correlation in the errors surrounding the adoption, production and sales decisions.

As before, we assume that the observed counts, y_1, correspond to a vector of latent variables z_1 representing the household's propensity to adopt the crossbred animal. In addition, we introduce a second, latent vector, z_{3c}, corresponding to the (negative) sales quantities implied for the censored components of the sales variable, y_3. (In our sample, 179 of the total 204 observations are censored.) It follows immediately that the joint posterior for the unknown quantities, B, Σ, z_1 and z_{3c}, has component conditional distributions:

$$\Sigma | B, z_1, z_{3c}, Y \quad \sim \quad \text{inverted-Wishart}(v_\Sigma, W)$$

$$B | z_1, z_{3c}, \Sigma, Y \quad \sim \quad \text{normal}(\hat{B}, \text{cov } \hat{B})$$

$$z_1 | z_{3c}, \Sigma, B, Y \quad \sim \quad \text{truncated-normal}(\hat{z}_1, \text{cov } \hat{z}_1) \tag{35}$$

$$z_{3c} | z_1, \Sigma, B, Y \quad \sim \quad \text{truncated-normal}(\hat{z}_{3c}, \text{cov } \hat{z}_{3c})$$

where the definitions appearing in the right-hand sides are, respectively, $v_\Sigma \equiv n$, $W \equiv (Y - XB)'(Y - XB)$; $\hat{B} \equiv (X'X)^{-1}$, $X'Y$, cov $\hat{B} \equiv \Sigma \ddot{A}(X'X)^{-1}$; $\hat{z}_1 \equiv XB_1$, B_1 denotes the first column of B, cov

$\hat{z}_1 \equiv (\Sigma - \Sigma_{1(23)}\Sigma_{(23)(23)} - 1 \ \Sigma_{(23)1} \ddot{A} I_n$, Σ_{11} denotes the element in the first row and first column of Σ $\Sigma_{1(23)}(=\Sigma_{(23)1})$ is the first row, second and third column submatrix of Σ and $\Sigma_{(23)(23)}$, and is

the second and third row, second and third column submatrix of Σ; $\hat{z}_3 \equiv X_c B_3, X_c$ denotes the submatrix of X corresponding to the censored observations; cov $z_3 \equiv \Sigma_{33} - \Sigma_{3(12)} \Sigma_{(12)(12)} - 1$ $\Sigma_{(12)3})$ $\ddot{A} I_c$, $B_3 \Sigma_{33}$ denotes the third column of B, Σ_{33} denotes the element in the third row and third column of Σ, $\Sigma_{3(12)}(=\Sigma_{(12)3}$ denotes the third row, second and third column submatrix of Σ, and $\Sigma_{(12)(12)}$ denotes the first and second row, first and second column submatrix of Σ. Note, once again, that it is straightforward to simulate from inverted-Wishart, normal and truncated-normal distributions, with truncation in z_1 corresponding to the counts (viz. j < zi ≤ j + 1 as yi = j) and truncation in z_{3c} corresponding to the censoring (viz. $z_{3i} \le 0$ as $y_{3i} = 0$). Consequently, a Gibbs sampling algorithm can be used to simulate from the joint posterior $p(\Sigma,B,z_1,z_{3c}|y)$ and calculate posterior quantities of interest. The algorithm is implemented by sampling sequentially from (35) with a burn-in phase of 5000 observations and a collection phase of 5000 observations and absorbed 11 minutes of real time under the same conditions as those applied for the single-equation model in the previous section.

Columns 3–5 of Table 7, present the results of the algorithm applied to the Ethiopian data. In the multivariate setting, gender, credit use, local cow ownership and visits by an extension agent remain significant determinants of crossbred cow adoption, with the covariates having quite similar impacts to those obtained from single-equation estimation. In addition, the own-equation variance of the multiple-equation model is quite similar to the variance in the single equation model. At first glance, these results would seem to suggest that there is little to be gained from inclusion of additional household decisions in the adoption model, but the covariance estimates in the lower part of Table 7 suggest otherwise. There are strong, positive relationships between the adoption and production, adoption and selling, and production and selling equations. Hence, inferences are surely strengthened by the inclusion of this additional, important information. Other features of the multiple-equation results include a slightly positive response to gender, a strongly negative response to credit use, and strongly positive responses to local cow ownership and visits by an extension agent in the milk output equation, whereas distance to market and visits by an extension agent appear to be the only strongly significant determinants of sales activities. Collectively, the results from the two systems highlight the importance of including additional household decisions in modelling the impacts of the covariates on sales decisions and raise considerable scope for further empirical inquiry.

7.5 Computing distances to market

An important policy matter is the quantity of a relevant resource that precipitates market entry by a formerly subsistence household. As before, in chapter 4–6, this task is made easy by noting that the sought quantities are the minimum levels of resources that make the (negative, latent) surplus values across the censored observations become non-negative. These quantities are readily computable for the non-participating households as a by-product of the data augmentation step in the Gibbs sampling algorithm. Focusing on the two most influential determinants of sales, we are interested in the average level of milk delivery by which the non-participating households are deficient, and the minimum distance to market and visits by extension agents values that reduce this deficiency to zero. The average non-participant is deficient in sales of about 8.6 (s.d. = 2.7) litres of milk per day. Holding other factors constant, an 8.6 litre daily increase in sales could be effected in one of two ways, namely an improvement in infrastructure ensuring return-travel times (to transport bucketed fluid milk to the milk co-operative) below 97 (s.d. = 49.2) minutes, or by offering extension visits in excess of 9.6 (s.d. = 3.1) per household per year. If greater milk market participation is an objective,

then future work should aim at assessing the cost-effectiveness of these two competing strategies, with an aim to rank the policies in terms of an appropriate welfare criterion.

7.6 Conclusions

This chapter has applied the MCMC techniques of Gibbs sampling and data augmentation to a sample of count data pertaining to the adoption of crossbred cows by formerly subsistence households. An intimate relationship exists between the decision to adopt a crossbred animal, the decision to produce surplus milk, and the decision to market quantities of that milk which is produced but not consumed internally by members of the household. We show how routine application of Gibbs sampling and data augmentation overcome potential estimation problems. A latent, categorical-variable formulation that works well in the single-equation setting, is readily extendable to multivariate situations and sheds light on important policy questions surrounding the adoption and market participation decisions. Future work should focus attentions on visits by extension agents (around nine/household per year), and improvements in infrastructure targeting travel times (in the order of 97 minutes) as appropriate policy targets; should aim to relax the inter-household commonality assumption; and should aim, more generally, to extend the ideas presented in this chapter to other, policy-relevant settings.

8   Estimation when fixed costs cannot be ignored

Until now we have not said anything about fixed costs, yet these costs are known to have important implications for market entry. The same is true in the household setting and this chapter contains a contribution around the idea that these costs are important, should be accounted for and, indeed, where possible, should be estimated as part of the empirical investigation.

8.1 Motivation

The main component of this chapter is, once again, the methodology proposed in the seminal paper on Gibbs sampling Tobit regressions (Chib 1992). In that paper the censoring point is assumed to be known with certainty to be zero. This assumption is the one that is most often applied. But this assumption implies that the minimum economic quantity of marketable milk surplus is zero. Theory predicts otherwise.

When subsistence households make path-altering decisions such as the decision to enter a market they incur fixed costs. Usually, but not always, this cost is a cost of time. Because fixed costs are prevalent there is a non-negligible but positive level of surplus that must accrue before the household actually participates in the market. Estimation of this quantity is important for policy analysis in the developing countries settings where transactions costs are high and inhibit market participation. In this chapter we show how Gibbs sampling a non-zero censored Tobit regression leads to precise estimates of the surpluses that must be accrued and to estimate the levels of the household resources that are required for market participation.

In our application to peri-urban milk market participation in the Ethiopian highlands, we focus on seven covariates, namely, the level of a traditional production practice (the number of indigenous milking cows); the level of a modern production practice (the number of crossbreed milking cows); the level of the intellectual-capital stock (measured in terms of years of formal education of the household head and the number of visits by an extension agent discussing production and marketing activities during the twelve months prior to the survey date); the physical distance that the household resides from the market (measured in terms of return time to transport bucketed fluid milk to market); and two site specific dummy variables representing the two locations at which the data were collected. As the previous chapters have shown, these set of covariates are the ones bearing the strongest relationships to marketable surplus accrual in the household. For each non-participating household, and for the average across all non-participants, we estimate the levels of these covariates that precipitate entry into the market.

8.2 Theory

The development of a Tobit regression follows, essentially, the same steps as those outlined in equations (9)–(14) and the notion that non-participating households 'distance' from the market is motivated in three steps. First, we develop a Kuhn-Tucker representation of the household's decision problem; second, we analyse the impact of relaxing a non-negativity restriction on marketable surplus; third, we formalise a Tobit regression that follows naturally from the first two steps. The difference here, however, is due to an alleged barrier impeding entry, leading to censoring of the Tobit regression at a point different from the traditional assumptions of zero.

Developments about the unknown-point-of-censoring model are now easily derived using the known-censoring formulation (equations 9–14) as a starting point.

With θ the assumed, random censoring point, the household's problem is to select marketable surplus, v_i, to solve:

$$\max_{v_i} \quad \Phi_i(v_i | x_i) \quad \text{subject to} \quad v_i \geq \theta. \quad (36)$$

This formulation gives rise to a corresponding Lagrangean,

$$\Phi_i(v_i | x_i) + \lambda(v_i - \theta) \quad (37)$$

and the associated Kuhn-Tucker conditions for a maximum,

$$\varphi_i(v_i | x_i) + \lambda = 0; \quad v_i - \theta \geq 0; \quad \lambda \leq 0; \quad \lambda(v_i - \theta) = 0 \quad (38)$$

As we have done previously, assume for the moment that the constraint is non-binding ($\lambda = 0$) and focus on the interior solution $\varphi_i(vi | xi) = 0$, recognising that some households will have marketable surplus below the censoring point (that is, for some i, $v_i \leq \theta$).

8.3 Statistical model

Developments analogous to those for the standard Tobit model lead to the statistical model

$$z_i = \beta_0 + \sum_{l=1}^{4} \beta_l x_{li} + \varepsilon_i \quad i = 1,2,...,N \quad (39)$$

where $\varepsilon_i \sim N(0, \sigma^2)$ and we observe $y_i = \max\{z_i, 0\}$.

With θ unknown, the non-informative prior $\pi(\beta, \sigma, z_2, \theta) \mu\sigma^{-1}$ can be combined with ideas in Albert and Chib (1993, p. 671, equation 3) to write the posterior density in the form:

$$\pi(\beta, \sigma, z_2, \theta | y) \propto \sigma^{-1} \prod_{i=1}^{N} N(z_i | x_i \beta, \sigma^2) \times \left[I(z_i > \theta)I(y_i = z_i) + I(z_i \leq \theta)I(y_i = 0) \right] \quad (40)$$

where I(·) denotes the indicator function (Mood et al. 1974 p. 20). Following developments in Albert and Chib (1993, p. 671, equations 4–6), the likelihood can be augmented with the latent data in order to form a Gibbs sampling algorithm for the unobserved σ, β, z_i ε c, and the unknown censoring point, θ. Some simple manipulations reveal that s has an inverse-gamma distribution, that β is multivariate normal, and that each z_i ($y_i = 0$) is normal with mean x_i β and

2

variance σ , truncated to the left by the condition $z_i \leq \theta$. Regarding θ, it follows from equation (40) that (when σ, β and the z_i's are known) the censoring point has the (fully) conditional distribution:

$$\pi(\theta|\beta, \sigma, z, y) \propto \prod_{i=1}^{N}\left[I(x_i > \theta)I(y_i = x_i) + I(x_i \leq \theta)I(y_i = 0)\right] \qquad (41)$$

which is uniform on the interval [max{z_i, i ε c}, min{z_i, i ε c}]. (See Albert and Chib 1993, equation 18, for a similar development in the context of an ordered probit specification.) Consequently, the algorithm just outlined is adjusted by including an additional step to estimate θ. The distance estimates are calculated, as before, by inserting into the algorithm the additional computation:

$$\overline{x_{ki}}^{(i+1)} = \theta^{(i)} - \frac{\beta_0^{(i+1)} + \sum_{j=1}^{m}\beta_j^{(i+1)}x_{ji} + \varepsilon^{(i+1)}}{-\beta^{(i+1)}} - x_{ki}, k = 1, 2, ..., m, i \in c \quad (42)$$

where $\varepsilon^{(s+1)}$ denotes a draw from N(0, $\sigma^{(s+1)}$). The algorithm in this section and the algorithm in chapter 5, together with the two measures of distance, (16) and (42), provide comparative measures with which to assess the robustness of the known, censoring assumption.

8.4 Results

Table 8 reports Gibbs sampling, data augmentation estimates of the Tobit regression based on a non-informative prior. Results from the zero-censoring formulation are presented in column 2 and results from the random censoring formulation are presented in column 3. Blockwise, the results consist of the parameter estimates, summary statistics and the distance-to-market estimates, respectively. The parameter estimates report means with 95% highest posterior density regions in parentheses; the distance estimates, report medians of the Gibbs sample.

Table 8. *Response of marketable surplus (litre of milk per day) to covariates.*

	Model	
Covariate	Zero censoring	Random censoring
Point of censoring		0.98
		(0.93, 1.00)
Sigma	4.98	3.93
	(4.49, 5.50)	(3.51, 4.34)
Distance	−0.07	−0.05
	(−0.09, −0.05)	(−0.07, −0.04)
Education	0.61	0.45
	(0.43, 0.79)	(0.32, 0.59)
Crossbred	4.13	3.45
	(3.61, 4.70)	(3.05, 3.91)
Local	1.85	1.54
	(1.38, 2.31)	(1.21, 1.92)
Extension	0.65	0.50
	(0.49, 0.83)	(0.37, 0.64)
Ilu-Kura	−9.55	−6.87

	(−11.32, −7.94)	(−8.19, −5.68)
Mirti	−14.90	−10.98
	(−16.99, −12.78)	(−12.51, −9.34)
	Censored observations	
R^2	0.93	0.94
Positive predictions	7.00	7.00
Negative predictions	1253.00	1253.00
	Uncensored observations	
R^2	0.29	0.32
Positive predictions	28.00	28.00
Negative predictions	140.00	140.00
	Distance estimates	
Milk	9.53	6.65
	(1.48, 19.86)	(0.21, 14.84)
Distance	−88.67	−74.01
	(−245.82, 87.23)	(−230.78, 87.85)
Education	−0.53	1.08
	(−16.28, 21.62)	(−15.20, 23.32)
Crossbred	−0.08	0.14
	(−2.40, 2.99)	(−2.03, 2.90)
Local	−0.17	0.31
	(−5.33, 6.92)	(−4.52, 6.69)
Extension	−0.49	0.97
	(−15.51, 19.76)	(−14.25, 20.79)

Note: Figures in parentheses are 95% confidence intervals (highest-posterior-density regions).

An important focal point between the two regressions is the true point of censoring in the marketable surplus data. This quantity is vitally important in policy calculations because it corresponds to a minimum efficient scale of operations for non-participants contemplating entry into milk markets.

The mean estimate of the censoring point, 0.98 (95% highest posterior density regions [0.93, 1.00]), leads to five observations that are important for policy. First, extension agents and policy planners aiming for increased market participation should target marketable surplus levels in non-participating households in the order of 0.98 litres of milk per household per day. Second (perhaps unsurprisingly, but in order to put this quantity in better perspective), 0.98 litres of milk per day lies close to the average level of milk sales among the market participating households at the two survey sites. Third, in view of this second observation, the estimated censoring point seems highly plausible. Fourth, with a 95% highest posterior density regions of [0.93, 1.00] the reported estimate is quite precise and leads, in turn, to precise policy calculations. Fifth, with 95% confidence, the true value of the censoring point is significantly different from zero—the censoring point that is typically assumed in the traditional formulation.

Focusing on the coefficient estimates, each of the parameters under the two formulations is significantly different from zero. Marketable surplus is declining with respect to increases in the amount of time that it takes to walk milk to the co-operative and is increasing with respect to the number of years of formal schooling of the household head, the number of crossbreed

cows that the household milks, the number of local breed cows that it milks and the number of extension visits availed to it in the 12 months prior to the survey. In addition, the estimates of the coefficients of the site-specific dummy variables are quite different. *Ceteris paribus*, milk marketing potential appears to be significantly higher at the Ilu-Kura site. Whether this difference is due to climatic or to management factors remains an interesting, open question.

Comparing results across the two formulations, the differences in magnitudes of the parameter estimates are striking. Without exception, the absolute value of each of the regression estimates is *greater* in the zero-censoring formulation. Hence, in the (two-space) graphs of marketable surplus against the covariates, the Tobit regression is noticeably 'steeper' under the zero-censoring formulation. This point is important in itself, but also has a significant bearing on the distance-to-market estimates. In interpreting these estimates, the requisite calculations (equations 16 and 42) imply the following conventions. Households with a positive requirements level are resource deficient (lie outside the market boundary); households with a zero requirement lie on the market boundary; and households with a negative requirement are resource sufficient (lie within the market radius).

Whereas the degree of explanatory power in both specifications is quite comparable (although some slight improvement is apparent in the random-censoring formulation), the distance-to-market estimates appear to be markedly affected by the allowance of a random censoring point. First, and most significantly, the median levels of *additional* marketable surplus required for entry are 9.53 and 6.65 (litres per household per day) under the zero-censoring and the random-censoring formulations, respectively. This difference is substantial and prima facie suggests that fewer additional resources may be required when a departure in the censoring point (the minimum efficient scale of production) from zero is assumed. But this conclusion is false. In fact, due to the greater 'slopes' of the Tobit regression under the zero-censoring formulation, estimates of each median resource requirement are actually lower under zero-censoring. More importantly, with the exception of the distance (time to walk to market) covariate, all median requirements estimates are *negative* in the case of zero censoring and *positive* under random censoring. This observation adds weight to the view that imposing a zero point of censoring on the marketable surplus data is a stringent restriction that may lead to significant bias in regression and corresponding distance estimates and lead to less plausible estimates of policy-important quantities. Turning to the median reports themselves, the median additional levels of resources required for entry under the random censoring formulation are 1.08 years of education, 0.14 crossbred cows, 0.31 local breed cows, and 0.97 extension visits. These are small, but plausible levels of additional resources for the 'representative household'.

One additional feature of the results that is not drawn out by these statistics is the wide range of estimates across the 1260 non-participating households. This feature is illustrated in greater detail by plotting the resource requirements. These plots are reported in Figures 11–16. In each case the households are ranked by their distance to market and then ordered so that the household that is farthest from the market (the household with the greatest resource deficiency) is the last household reported (household #1260) and the household with the greatest abundance of the resource is the first household in the ranking (household #1).

Figure 11 reports Gibbs-sample median requirements of additional milk across the households. By virtue of the fact that each household is a non-participant, each of the requirements estimates is positive. The reports range from a low of 1.5 litres of milk per day to a high of 18.6 and a significant divergence between the estimates from the two formulations is apparent—imposing zero censoring on the data leads to significantly greater estimates of additional milk requirements throughout the entire sample.
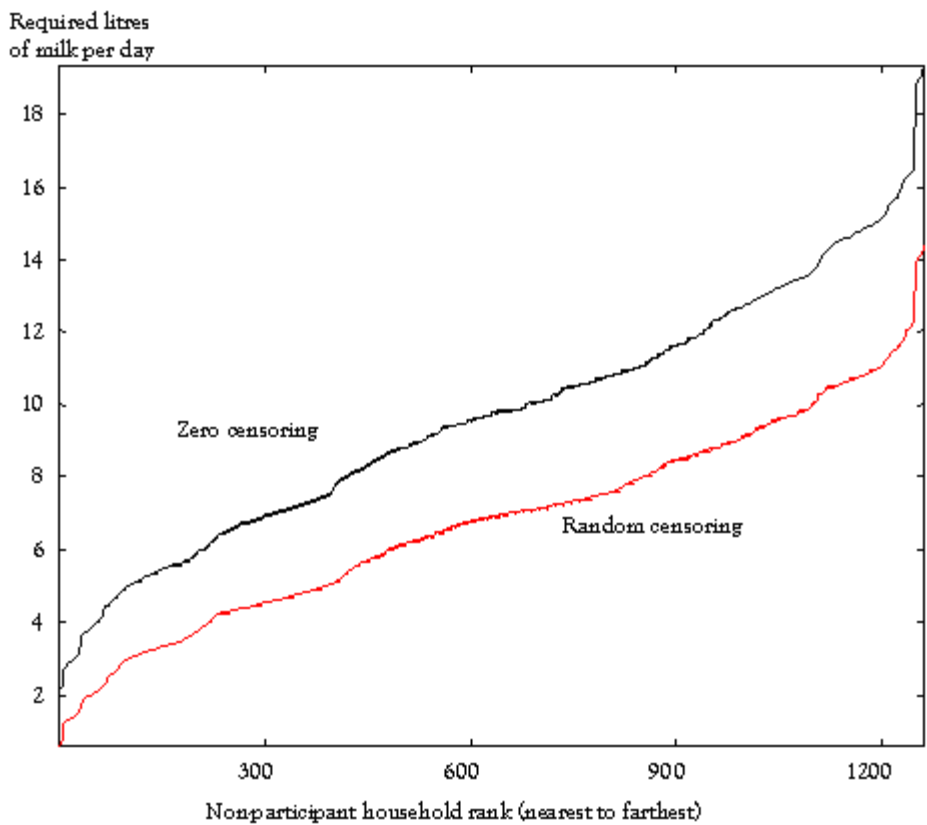
Figure 11. *Distance to market estimates: Milk.*

Figure 12 reports estimates of amounts by which time to walk to the milk group must be *lowered* in order to make marketing viable in the household. Unlike the reports corresponding to the remaining covariates, the time reports are bounded by the actual time required to walk to the milk group. For households who require greater time reductions than their observed travel time, this need can only be met through additions of other resources. Of the 1260 non-participants, there are 915 households for which travel time is not constraining and there are 345 households for which it is. Of those 345 households, there is no household for which the estimated *reduction* in travel time (under both formulations) is *greater* than the observed travel time. This point is important. It suggests that, for a significant proportion of the non-participating sample, a *ceteris paribus* reduction in travel time appears to be a potent policy alternative. The remaining observation of significance in Figure 12 is that the median reports under the two formulations are quite similar across the entire non-participating sample.
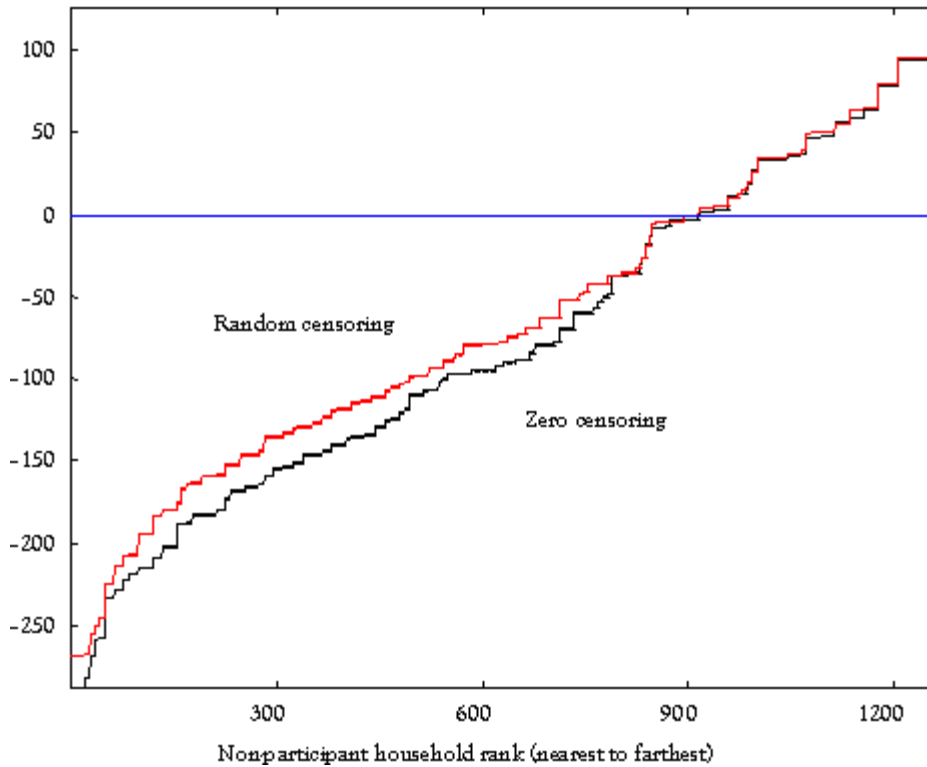
Figure 12. *Distance to market estimates: Time to the milk group.*

Unlike the travel-time estimates, the reports for education, crossbreed cows, local breed cows and visits by extension agents (Figures 13–16, respectively) are unbounded. For each of these covariates, the reports from the two formulations are quite similar. Also, *for each of the covariates in question*, the zero censoring model predicts that 604 households are resource deficient whereas the random censoring formulation reports 653 households are deficient. Small, but significant differences prevail across the reports from the two formulations, with the reports from the random censoring formulation generally greater than the reports from the zero-censoring formulation (there are 7 exceptions in the case of education, 471 in the case of crossbreed cows, 471 in the case of local breed cows and 0 in the case of extension).

Collectively, Figures 13–16 highlight the importance of three aspects of the distance-to-market estimates using the Tobit regression. First, the results suggest that random censoring leads to significantly different estimates of marketable surplus requirements. Second, despite the latter, differences in resource requirements estimates across the two formulations are, generally, quite small. Third, more generally, potentially important biases in estimates of policy quantities may arise from ignoring random censoring in market-participation studies using household production data.
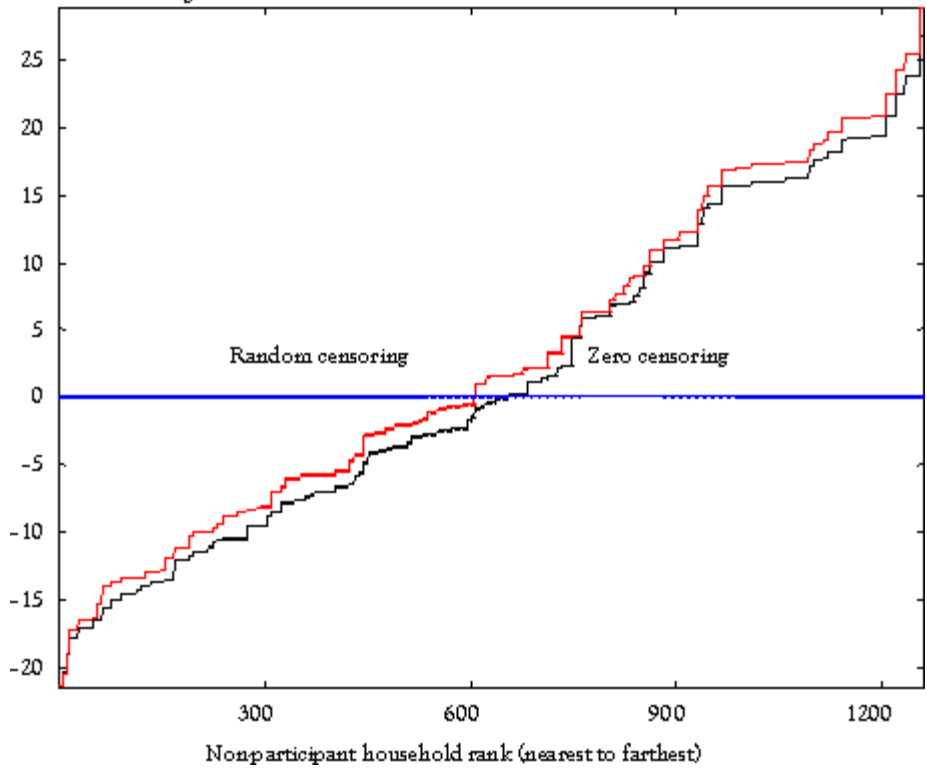
Figure 13. *Distance to market estimates: Education.*
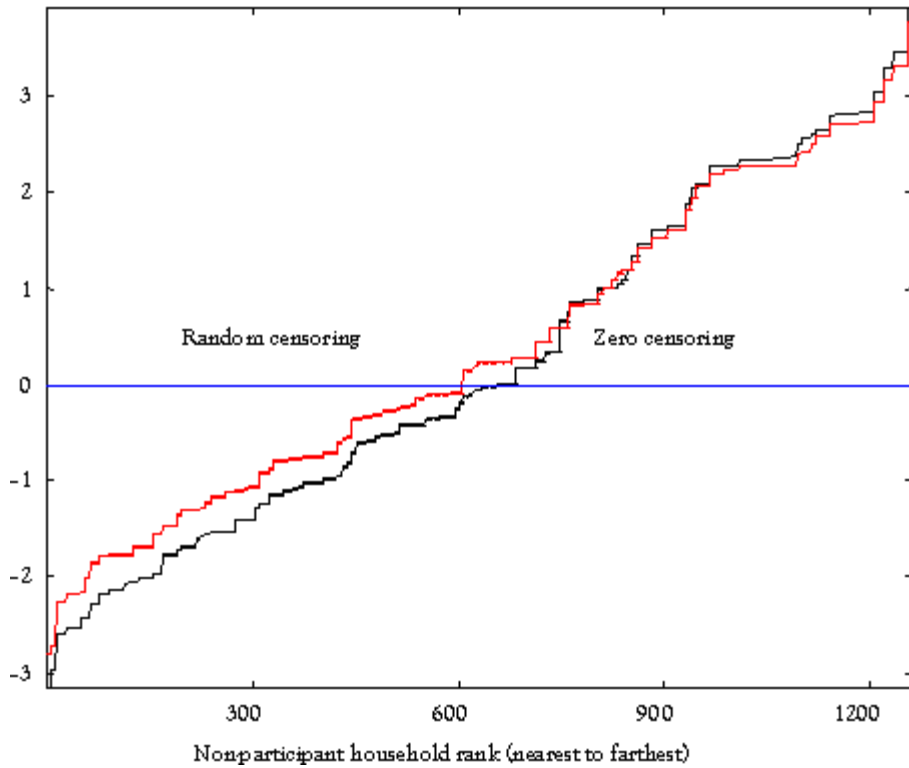
Figure 14. *Distance to market estimates: Crossbred cows.*
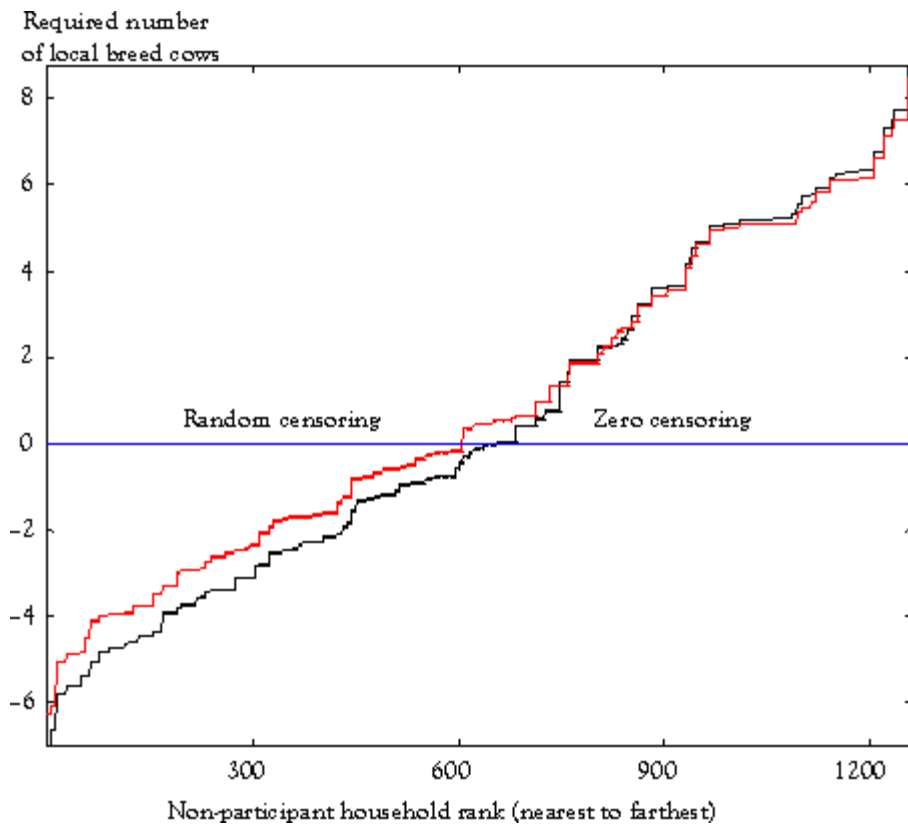
Figure 15. *Distance to market estimates: Local breed cows.*
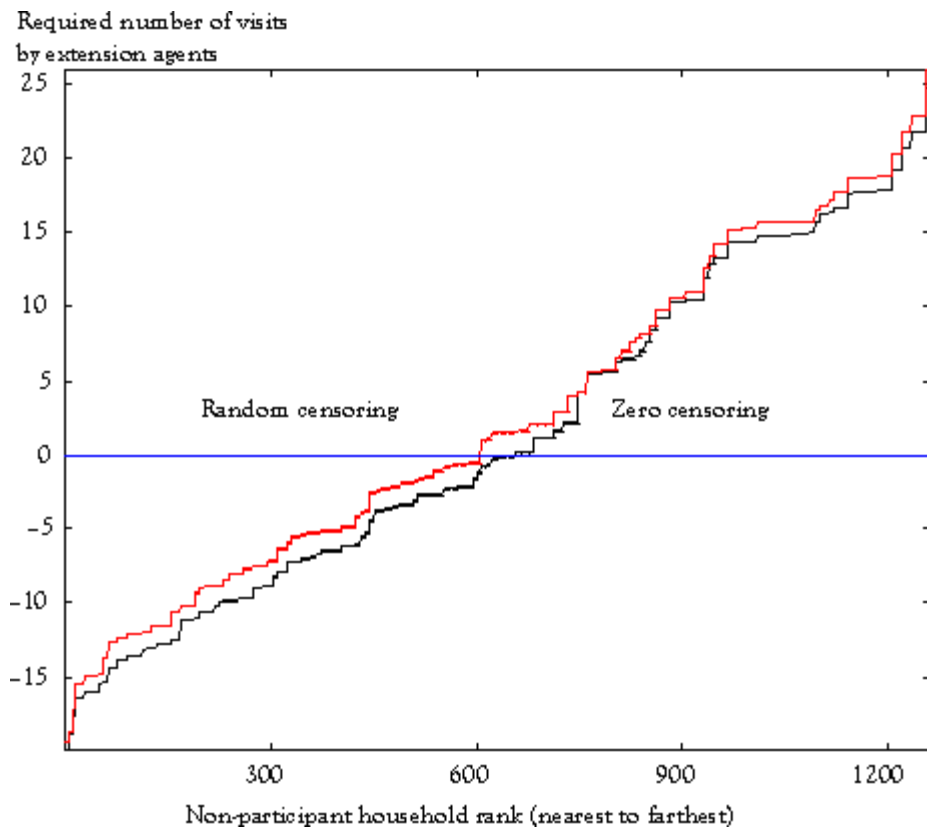
Figure 16. *Distance to market estimates: Visits by extension agents.*

8.5 Conclusions

There are two key conclusions. First, small but significant differences arise from incorporating

random censoring into the Tobit regression. Second, Markov chain Monte Carlo (MCMC) methodology has the ability to answer important research questions in a robust, appealing manner.

9   A double hurdle model of market participation

Our last application investigates the consequences for milk sales and distance-to-market estimates of viewing the entry decision as a two-step procedure. This type of framework has recently garnered support for modelling diverse consumption decisions and is motivated by the possibility that the factors affecting the participation decision (e.g. whether to smoke cigarettes) may be different from the factors affecting the decision about quantities (e.g. how many cigarettes to smoke per day).

9.1 Motivation

As we did in the previous chapter, we focus much of our attentions towards the fixed costs issue. 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 chapter we demonstrate a method for estimating the double hurdle model of market participation and supply volume determination in the face of unobservable fixed costs.

9.2 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 an asterisk (*) 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, x_{pi}. Without loss of generality, we set $y_{ri}^* = 0$ and denote the difference between the incurred and reserve utility levels y , and their relationship

to the characteristics by the function $f_i(Y)$. The condition characterising the discrete choice about whether to participate in the market can then be written as:

$$y_{pi} = f_i(x_{pi})$$

(43)

with participation when $y_{pi} > 0$ and non-participation 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 non-participation.

Statistical implementation depends on the information structure of this choice problem, in particular whether the discrete participation decision occurs *before* a corresponding quantity decision is undertaken about the *intensity* of participation, in this case, as to *how much* 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 $ji(\,\cdot\,)$ 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 x_{si} denote these characteristics. Across each of the households i = 1, 2, …, N, we are concerned, once again, with the problem:

$$\max_{y_{si}} \Phi_i(y_{si}|x_{si}) \qquad \text{subject to} \qquad y_{si} \geq 0$$

(44)

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

$$\varphi_i(y_{si}|x_{si}) \leq 0$$

(45)

the non-negativity restriction on choice,

$$y_{si} \geq 0$$

(46)

and the complementary-slackness condition,

$$\varphi_i(y_{si}|x_{si})y_{si} = 0$$

(47)

Equations (43)–(47) form the basis for a double hurdle interpretation of the household's supply decision, on which we now expand.

9.3 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:

$$y_{si} = \delta_i y_{si}^{**} \tag{48}$$

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 43) has the form:

$$y_{pi} = \beta_p x_{pi} + u_{pi} \tag{49}$$

where $\delta_i = 1$ if $y_{pi} > 0$ and $\delta_i = 0$ otherwise, where β_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:

$$y_{si}^* = \beta_s x_{si} + u_{si} \tag{50}$$

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, x_{si}, by the vector β_s, with u_{si} a random error.

Equations (48)–(50), along with their restrictions, combine to yield the double hurdle motivation for participation. This notion is exhibited clearly in equation (48), 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 (49). Second, the latent quantity y_{si}^* must exceed zero in equation (50). Hence, both the participation- and the sales-equations 'constraints' must be satisfied in order for positive sales to arise.

Equation (49) is simply a linear, statistical interpretation of the participation decision in equation (43) and, when the error is normal, has the important connotation of a probit equation. Equation (50) follows from relaxing the non-negativity constraint in equation (46), ignoring the complementary-slackness condition in equation (47) and acknowledging that, when one does so, a latent, censored (Tobit) regression is implied in which observed sales are left-censored at zero.

9.4 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

$$\ell(y_{si} > 0) = f(y_{si} | \delta_i = 1 \text{ and } y_{si} > 0) \times prob(\delta_i = 1 \text{ and } y_{si}^* > 0) \tag{51}$$

Consequently, the likelihood for observing zero sales is the probability that neither of the two conditions in question prevail, or

$$\ell(y_{si} = 0) = 1 - prob(\delta_i = 1 \text{ and } y_{si}^* > 0) \tag{52}$$

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 (49) and (50) follow a multivariate-normal distribution. In this context, equation (49) depicts a traditional probit regression, equation (50) depicts a traditional Tobit regression, and the multivariate-normal assumption allows correlation between the errors, as in Nelson (1977), Cogan (1981)

and Goetz (1992). By combining results in Chib (1992) and in Albert and Chib (1993), 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 (49) and (50) as:

$$y = x\beta + u \qquad (53)$$

where $y \equiv (y_p', y_s')'$, $yp \equiv (y_{p1}, y_{p2}, ..., y_{pN})'$, $y_s \equiv (y_{s1}, y_{s2}, ..., y_{sN})'$; $x \equiv (x_1, x_2)'$, $x_1 \equiv (x_p, 0_s)'$, $x_2 \equiv (0_p, x_s)'$, $x_p \equiv (x_{p1}, x_{p2}, ..., xp_N)'$, $x_{p1} \equiv (x_{p11}, x_{p12}, ..., x_{p1kp})$, $x_{p2} \equiv (x_{p21}, x_{p22}, ..., x_{p2kp})$, ..., $x_pN \equiv (x_{pN1}, x_{pN2}, ..., x_{pNkp})$, $xs \equiv (x_{s1}, x_{s2}, ..., x_{sN})'$, $xs1 \equiv (x_{s11}, x_{s12}, ..., x_{s1ks})$, $x_{s2} \equiv (x_{s21}, x_{s22}, ..., x_{s2ks})$, ..., $x_{sN} (x_{sN1}, x_{sN2}, ..., x_{sNks})$; $\beta \equiv (\beta p', \beta s')'$, $\beta p \equiv (\beta_{p1}, \beta_{p2}, ..., \beta_{pkp})'$, $\beta s \equiv (\beta_{s1}, \beta_{s2}, ..., \beta_{sks})'$; 0_p and 0_s are null vectors of dimensions N × ks and N × kp, respectively; and the 2N vector $u \equiv (up', us')'$, $up \equiv (u_{p1}, u_{p2}, ..., u_{pN})$, $u_s \equiv (u_{s1}, u_{s2}, ..., u_{sN})'$, is assumed to have a multivariate normal distribution with mean the 2N null vector and covariance $\Sigma Ä I_N$. The parameters of the 2 × 2 covariance matrix Σ 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 (53) is in the form of Zellner's (1996 seemingly unrelated regressions model equations 8.72–8.78). As such, the model plays an important role in another discrete-choice setting that has received considerable attention of late, namely the multinomial-probit model (see, for example, Geweke et al. 1994; Dorfman 1996; Geweke et al. 1997; McCulloch et al. 2000). 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 2N × 1 vector y contains both observed and latent components. The first N components, y_p, are all latent and some proportion of the second component, y_s, will also be unobserved. In particular, define $c \equiv \{ i \mid y_{si} = 0 \}$ 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 (non-positive) quantity of sales is implied. These quantities facilitate estimation—a point that is demonstrated to great effect in the seminal paper, Chib (1992)—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 characterised with reference to the probability masses of the four, respective events: $E1 \equiv$ event ($\delta_i = 1$ and $y_{si}^* > 0$), $E_4 \equiv$ the event ($\delta i = 1$ and $ysi^* £ 0$), $E_3 \equiv$ the event ($\delta i = 0$ and $ysi^* > 0$) and $E_4 \equiv$ the event ($\delta i = 0$ and $ysi^* \le 0$). 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 $ysi^* > 0$ as perfectly correlated. This interpretation, in effect, assigns zero probability to events E_2 and E_3 (prob($di = 1$ and $ysi^* \le 0$) = prob($\delta i = 0$ and $ysi^* > 0$) = 0). Then, according to the restrictions implied by the probit model in equation (49) all N elements of yp are latent with ypi truncated to the positive (negative) part

of the real line according to $\delta i = 1$ ($\delta i = 0$) and, in addition, the censored components of ys 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 ysi* ≤ 0) = 0 but prob($\delta i = 0$ and ysi* > 0) ≠ 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 ysi* > 0) = 0 but prob($\delta i = 1$ and ysi* ≤ 0) ≠ 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 example, Cragg 1971; Fin and Schmidt 1984; Jones 1989).

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 (models two and three) and, hence, reports are made only for the model 1 specification. In addition, further experiments led to the conclusion that the *same* covariates were significant in explaining both the participation and the supply decisions.

In this case, the seemingly-unrelated regressions model in equation (54) reverts to the traditional multivariate regression system (Zellner 1996, p. 224, equation 8.1) and estimation is slightly simplified. In terms of equation (53), the modifications implied are $y \equiv (y_p, y_s)$; $x \equiv x_p \equiv x_s$; x has dimensions $N \times k$; $b \equiv (bp, bs)$; and $u \equiv (up, us)$ is now assumed to have a multivariate normal distribution with mean the $N \times 2$ null vector and covariance $\sum Ä I_N$. Additionally, due to the facts that the vector yp is latent, and a subset of the components of ys is also latent, we use the symbols z_p and z_s to signify the corresponding observed vectors with the latent components included. Hence, $z \equiv (z_p, z_s)$. Finally, in a conventional notation, we note that there are $m = 2$ equations in the system.

With this notation at hand, under a non-informative prior $\pi(\sum, \beta, z_p, z_s) \mu |\sum|^{-(m + 1)/2}$, the full conditional distributions comprising the joint posterior for the unknown parameters and the latent data, $\pi(\sum, \beta, z_p, z_s \mid y, x)$, have the following forms:

$$z_p \mid \Sigma, \beta, z_s \quad \sim \quad \text{truncated-normal}(Ez_p, Vz_p)$$

$$z_s \mid k_s, \Sigma, \beta \quad \sim \quad \text{truncated-normal}(Ez_s, Vz_s)$$

$$\beta \mid k_{\cdot,\cdot}, z_s, \Sigma \quad \sim \quad \text{normal}(E\beta, V\beta) \qquad\qquad (54)$$

$$\Sigma \mid \beta, z_{\cdot,\cdot}, z_s \quad \sim \quad \text{Inverted-Wishart}(W, v)$$

where $Ez_p \equiv x\,\beta_p + \Sigma_{ps}\Sigma_{ss}^{-1}(z_s - x\,\beta_s)$, $Vz_p \equiv \Sigma_{pp} - \Sigma_{ps}\Sigma_{ss}^{-1}\Sigma_{sp}$; $Ez_s \equiv x\,\beta_s + \Sigma_{sp}\Sigma_{pp}^{-1}(z_p - x\,\beta s)$, $Vz_s \equiv \Sigma_{ss} - \Sigma_{sp}\Sigma_{pp}^{-1}\Sigma_{ps}$; $E\beta \equiv (x'x)^{-1}z$, $V\beta \equiv \Sigma\ddot{A}\,(x'x)^{-1}$; $W \equiv (z - x\,\beta)'(z - x\,\beta)$, $v \equiv N{-}k + m + 1$; and the 2×2 matrix Σ has (scalar) components Σ_{pp}, Σ_{ps}, Σ_{sp} and Σ_{ss}. Consequently, simulations from the joint posterior can be undertaken through a Gibbs sampling, data augmentation algorithm that samples sequentially through the distributions in (54), and the outputs $\{\Sigma^{(s)} \ s = 1, 2, \ldots, S\}$, $\{\beta^{(s)} \ s = 1, 2, \ldots, S\}$, $\{z_p(s) \ s = 1, 2, \ldots, S\}$ and $\{z_s(s) \ s = 1, 2, \ldots, S\}$ can be used to derive inferences about policy measures of interest.

However, unlike the previous algorithms, three additional features are necessary for convergence. First, due to identification problems, the draw from the inverted-Wishart in step 2 is normalised 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 subset of the vector z_s, corresponding to the households in the censor set, are drawn from the conditional normal distribution and the draws for both z_p and z_s 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 $S^1 = 2000$ observations followed by a 'collection phase' of $S^2 = 2000$ observations.

In closing this section it seems natural to ask the extent to which the well-known problem of *sample selection bias* (see, for example, Greene (2000), 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 (1992). 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 *selection bias* 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. However, 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 sampling algorithm 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 (1992) are similarly circumvented. Hence the proposed algorithm appears to offer a number of attractive features compared with more traditional methodology.

9.5 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, and the results of the previous chapter, 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, will be 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 conceptualised in the context of the decision-making model (equations 43–47), the statistical description of the hurdle model (equations 48–50) and the estimation model (equations 51–54), 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 the household level, trade is commonly discontinuous in time, with individual households selling in 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 optimisation problem in (44), the constraint $y_{si} \geq 0$ is replaced by the condition $y_{si} \geq q$. 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 (50), 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 Σ and β, the latent z_p and the latent components of z_s.

All that remains is to derive the fully conditional distribution for the unknown censoring value, θ, and append this distribution to the sampling algorithm implied by (54). From the results established in the previous chapter we note that this distribution is:

$$\theta|\Sigma,\beta,z,,z_q \sim \mathrm{uniform}(\max\{z_{si}, i\in c\}, \min\{z_{si}, i\notin c\}) \qquad (55)$$

implying a few modifications to the algorithm in (54). The first modification is to select 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 subsequent draws in the algorithm are now conditional on the chosen value $\theta^{(s)}$. Third, at the end of the algorithm we insert the additional step: Draw $\theta^{(s+1)}$ from the uniform distribution with bounds $[\max\{z_{si}(s+1), i\ c\}, \min\{z_{si}, i\ c\}]$, where $\max\{z_{si}(s+1), i\ c\}$ implies conditioning on the maximum component of $z_s(s+1)$ and where $\min\{z_{si}, i\ c\}$ denotes the minimum sales quantity observed in the data.

9.6 Results

Results of the Gibbs sampling, data augmentation algorithm applied to the 204 observations are presented in Table 9. 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 Table 9. The mnemonics in the first column refer, respectively, to θ ('Censor value'); return time (in minutes) to transport fluid milk to the milk co-operative ('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 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.

Table 9. *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
Positive predictions	3.00	4.00	2.00	8.00
Negative predictions	176.00	175.00	177.00	171.00
Participants				
R^2	0.92	0.33	0.84	0.39
Positive predictions	11.00	11.00	25.00	13.00
Negative predictions	14.00	14.00	0	12.00

Note: Figures in parentheses are 95% confidence intervals (highest posterior density values).

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 visits by extension agents) and the animal stock (local and crossbreed animals) but 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 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 litres 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 litres 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 *greater* (in absolute value) than the corresponding estimate in the traditional model and in the supply equation each of the estimates is *smaller* (in absolute value) than the corresponding estimate in the traditional, zero censoring model. Furthermore, in both the participation and supply equations, the site-specific dummy coefficients are greater 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 Table 9. 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.

9.7 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 visits by extension agents) 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 litres of milk per household per week.

10 Concluding comments

Gibbs sampling and data augmentation have revolutionised Bayesian inference, particularly in extensions of the normal-linear model. The separate works collected here present applications of these ideas to data on an emerging milk market in the Ethiopian highlands. Fundamental to this collection is the notion of statistical robustness.

The concept that statistical measures remain robust to a diverse set of alternative model formulations is important for policy. This collection has summarised the results of a search for three measures of particular relevance to market participation policy, namely the levels of three essential inputs in the milk production and selling exercise. In this regard, this collection has highlighted the need for policymakers, administrators, extension specialists etc. to focus attentions on the additions of about 2–3 crossbred cows, 7–8 local breed cows or 9–10 visits by extension agents as paramount in effecting participation among representative non-participants. That these quantities have been discovered under so many alternative specifications suggests an important conclusion from the exercise. These estimates are robust.

Although this search for robustness was the main objective of this exercise—an objective shared also in the genesis of data collection (Nicholson C.F., personal communication)—we discovered, along the way, another important fact. This fact is that routine applications of Markov chain Monte Carlo methods—Gibbs sampling and data augmentation in particular— provide important measures for market-development policy. Indeed, in view of their worth, it is surprising that these methods have received so little attention by agricultural economists and development economists to date. Agricultural economists, we believe, particularly those with empirical interests, should aim to focus more attention on the method; a series of fundamental contributions (Chib 1992, 1995, 1996); their application (Dorfman 1996); and, we hope, their profitable exploitation in policy formation.

References

Akwabi-Ameyaw K. 1997. Producer cooperative resettlement projects in Zimbabwe: Lessons from a failed agricultural development strategy. *World Development* 25:437–456.

Albert J.H. and Chib S. 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* 88(June):669–679.

Cameron L. 1999. The importance of learning in the adoption of high-yielding variety seeds. *American Journal of Agricultural Economics* 81:83–94.

Casella G. and George E. 1993. Explaining the Gibbs sampler. *American Statistician* 46:167–174.

Chib S. 1992. Bayes inference in the Tobit censored regression model. *Journal of Econometrics* 52:79–99.

Chib S. 1993. Bayes estimation of regressions with autoregressive errors: A Gibbs sampling approach. *Journal of Econometrics* 58(2):275–294.

Chib S. 1995. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* 90:1313–1321.

Chib S. 1996. Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics* 75(1):79–97.

Chib S. and Greenberg E. 1995. Understanding the Metropolis-Hastings algorithm. *American Statistician* 49(2):327–335.

Cogan J.F. 1981. Fixed costs and labor supply. *Econometrica* 49:945–963.

Cragg J. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 39:829–844.

Debrah S. and Berhanu Anteneh. 1991. *Dairy marketing in Ethiopia: Markets of first sale and producer's marketing patterns.* ILCA Research Report 19. ILCA (International Livestock Centre for Africa), Addis Ababa, Ethiopia. 21 pp.

Delgado C. 1995. Agricultural diversification and export promotion in sub-Saharan Africa. *Food Policy* 20:225–243.

Delgado C. 1999. Sources of growth in smallholder agriculture in sub-Saharan Africa: The role of vertical integration of smallholders with processors and marketers of high-value items. AGREKON (*Journal of the Agricultural Economics Association of South Africa*) 38(May):165–189.

Dempster A.P., Laird N. and Rubin D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 37B:1–38.

Dorfman J.H. 1996. Modeling multiple adoption decisions in a joint framework. *American Journal of Agricultural Economics* 78(3):547–557.

Dréze J.H. and Richard J.F. 1983. Bayesian analysis of simultaneous equation systems. In: Griliches Z. and Intriligator M.D. (eds), *Handbook of econometrics.* Volume I. Amsterdam,

North Holland, The Netherlands.

Fieller E.C. 1932. The distribution of the index in a normal bivariate population. *Biometrika* 24(3):428–440.

Fin T. and Schmidt P. 1984. A test of the Tobit specification against an alternative suggested by Cragg. *Review of Economics and Statistics* 66:174–177.

Geary R.C. 1930. The frequency distribution of the quotient of two normal variates. *Journal of the Royal Statistical Society* 93(3):442–446.

Gelfand A. and Smith A. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85:972–985.

Geweke J., Keane M. and Runkle D.1994. Alternative computational approaches to inference in the multinomial probit model. *Review of Economics and Statistics* 76:609–632.

Geweke J., Keane M. and Runkle D. 1997. Statistical inference in the multinomial, multiperiod probit model. *Journal of Econometrics* 80:125–166.

Goetz S. 1992. A selectivity model of household food marketing behavior in sub-Saharan Africa. *American Journal of Agricultural Economics* 74(May):444–452.

Greene W.H. 2000. *Econometric Analysis.* 4th edition. Prentice Hall, New Jersey, USA.

Grosh B. 1994. Contract farming in Africa: An application of the new institutional economics. J*ournal of African Economies* 3:231–261.

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* 5:475–492.

Heckman J.J. 1979. Sample selection bias as a specification error. *Econometrica* 47:53–161.

Hinkley D.V. 1969. On the ratio of two correlated normal random variables. *Biometrika* 56(4):635–639.

Jaffee S. 1994. Perishable profits: Private sector dairy processing and marketing in Kenya. In: Jaffee S. and Morton J. (eds), *Marketing Africa's high-valued foods: Comparative experiences of an emergent private sector.* Kendall-Hunt, Dubuque, Iowa, USA. pp. 199–253.

de Janvry A., Sadoulet E. and Thorbecke E. 1993. Introduction to a special section, state, market, and civil organizations: New theories, new practices, and their implications for rural development. *World Development* 21:565–575.

Jones A. 1989. A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics* 4(January–March):23–39.

Kiwuwa G.H., Trail J.C.M., Kurtu M.Y., Worku G., Anserson F.M. and Durkin J. 1983. *Crossbred dairy cattle productivity in Arsi region, Ethiopia.* ILCA Research Report 11. ILCA (International Livestock Centre for Africa), Addis Ababa, Ethiopia.

Marsaglia G. 1965. Ratios of normal variables and ratios of sums of uniform variables. *Journal of the American Statistical Association* 60(2):193–204.

McCulloch R., Rossi P. and Polson N. 2000. Bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics* 99:173–193.

Merrill A.S. 1928. Frequency distribution of an index when both the components follow the normal law. *Biometrika* 20A(1):53–63.

Mood A.M., Graybill F.A. and Boes D.C. 1974. *Introduction to the theory of statistics.* 3rd edition. McGraw-Hill, New York, USA.

Nelson F. 1977. Censored regression models with unobserved, stochastic censoring thresholds. *Journal of Econometrics* 6:309–327.

Nicholson C.F. 1997. The impact of milk groups in the Shewa and Arsi regions of Ethiopia: Project description, survey methodology, and collection procedures. ILRI Livestock Policy Analysis Project. ILRI (International Livestock Research Institute), Addis Ababa, Ethiopia. (Mimeograph).

Nicholson C.F., Gebru G., Ehui S.K., Shapiro B.I. and Delgado C. 2000. Producer milk groups in Ethiopia's highlands: A framework for assessing impacts and a review of group performance. In: *The role of village dairy co-operatives in dairy development: Prospects for improving dairy in Ethiopia. Proceedings of a workshop, held from 22–24 April 1998*. Ministry of Agriculture, Addis Ababa, Ethiopia. pp. 82–103.

Singh I., Squire L. and Strauss J. (eds). 1986. *Agricultural household models: Extensions, applications and policy.* The Johns Hopkins University Press, Baltimore, USA. 335 pp.

Staal S.J. 1995. *Peri-urban dairying and public policy in Ethiopia and Kenya: A comparative institutional and economic analysis*. PhD dissertation. University of Florida, Gainesville, USA.

Staal S.J., Delgado C. and Nicholson C. 1997. Smallholder dairying under transactions costs in East Africa. *World Development* 25:779–794.

Stiglitz J.E. 1989. Markets, market failures and development. *American Economic Review* 79(May):197–203.

Zellner A. 1996. *An introduction to Bayesian inference in econometrics*. Wiley Classics Library Edition, Wiley, New York, USA.