Hedonic Prices for Timber Auctions with Endogenous Participation

Raphaële PRÉGET (INRA)[§]
Patrick WAELBROECK (ENST)[¥]

October 23rd, 2006

Abstract

Timber stumpage appraisal is an important issue in timber markets. Public Forest Services need to set a relevant reserve price and to obtain a fair market value. The hedonic price function approach is a useful method to infer the appraisal timber value since so many characteristics may influence the stumpage price. We adopt the transaction-evidence timber appraisal using data from timber auctions in Lorraine (Eastern France). We propose a Bayesian Metropolis-Gibbs Monte Carlo Markov Chain (MCMC) algorithm to estimate parameters of a sample selection model in which the hedonic equation includes a binary endogenous explanatory variable. We use a three simultaneous equation model: equation (1) defines the probability that there is no bid, equation (2) determines among submitted lots if there is only one bid (no competition), equation (3) is the hedonic price equation that explains the highest bid. We find that the MCMC algorithm provides stable results across different model specifications, whereas the Heckman sample selection procedure results in unreliable inference on the coefficient associated with the binary endogenous variable as well as the correlation coefficient. When there is no competition (one bid) the submitted price is -40.3% lower than when there is competition.

Key words: Timber auctions, hedonic prices, unsold lots, sample selection, binary endogenous explanatory variable, Metropolis-Gibbs, endogenous participation.

Code JEL: C11, C15, C34, C35, C63, D44, L73.

[§] Laboratoire d'Economie Forestière, UMR ENGREF/INRA, 14 rue Girardet, 54042 Nancy cedex, France; email preget@nancy-engref.inra.fr
[¥] EGSH/ENST, 46 rue Barrault, 75013 Paris, France; email: waelbroe@enst.fr

1 Introduction

Timber stumpage appraisal is an important issue in timber markets. Public Forest Services need to set a relevant reserve price and to obtain a fair market value. But what is a fair market price for a heterogeneous product such as standing timber lots? There is quite a difference between forest production (which spreads over a long period of time) and timber supply (which is a harvesting decision). So without real references to production costs, the seller has to consider other elements to assess the value of her products and to determine her own reservation value. Concretely, the seller has to estimate market demand. Introduced in the 1980s in the USDA Forest service, the “transaction evidence appraisal” method (TEA) estimates timber value from market prices obtained during past timber sales.[1]

In this article, we are interested in the results of French timber auctions. Most of French timber sales are sequential first-price sealed-bid auctions of heterogeneous lots. Heterogeneity in the product is probably the most important feature of standing timber sales. Lots differ from each other with respect to volume, composition, localization, harvesting conditions, etc. (inter lots heterogeneity). But a lot is also composed of different species and qualities (intra lot heterogeneity). Thus, lots are not only different from one another, but there are also made up of heterogeneous wood. These inter- and intra-lot heterogeneities raise various questions about the valuation of the lots that are put on sale, about their optimal composition, as well as about the allocative efficiency of a sequential mechanism.

Heterogeneity makes the hedonic price function approach useful in order to infer appraisal timber value since many characteristics may influence the stumpage price. The hedonic price method is based on the implicit price of each characteristic and determines how the buyers value a lot as a set of characteristics. This issue is important for the seller to better anticipate the demand in order to compose attractive lots and to set relevant reserve prices. We are thus

[1] The buyers estimate the value of a standing timber lot in a different way since they have more information on harvesting costs, on what they will produce with the wood and at what price they will be able to sell their

interested in finding the variables that influence the highest bid in order to construct a hedonic price function.

There are two problems that arise when we analyze timber auction data sets. Both arise from the endogenous participation of the bidders to the auctions. First, there are many lots for which there is no bid and there are good reasons to think that this outcome is not random: bidders may not bid for timber lots that are of bad quality or have difficult harvesting conditions, etc. It is important to note that in French timber auctions, the seller does not announce any reserve price. The seller might withdraw the lot if she thinks that the highest bid is too low, but the reserve price is kept secret. Thus, the lack of bids can not be explained by a reserve price that is set to high, since no minimum amount is required to bid for a lot. Of course, lots with no submission remain unsold. However, we have to take them into account in the estimation procedure so as to prevent a possible sample selection bias. Secondly, the degree of competition varies from one auction to another and we know that the number of bidders has an impact on the bidding strategies. Indeed, for some lots, bidders bid aggressively because there are many other bidders who want to win the auction. On the other hand, there are many auctions in which there is only one bid. We show that, when there is no competition, the highest bid is significantly lower. Again, the number of bids can not be explained by the value of the reserve price, so it is sensible to think that such an outcome is driven by the observed and unobserved characteristics of the lot. In other words, the number of bidders has to be included in the hedonic price equation as an endogenous explanatory variable of the highest bid. In our data set, the number of bidders is a discrete variable. Thus we have to deal with at least one binary endogenous variable in addition to the issue of sample selection.

Endogenous participation raises two econometric issues: sample selection and endogenous binary variable in the hedonic price equation. The main problem is related to the correlation between unobservable variables that determine the participation process and the highest bid. We solve this challenge by specifying a 3-equation model that we estimate using Monte Carlo Markov Chain (MCMC) Bayesian simulation algorithms. The idea is to simulate the

---

products. It is therefore easier for buyers to measure their willingness to pay for the lot, i.e. to determine their reservation value.

(latent) variables that determine the participation outcomes, which greatly simplifies the analysis of the joint posterior distribution of the parameters. This technique is called data augmentation. Using this methodology, we find that highest bids of lots for which there is no competition (in other words, for which there is only one bidder) are much lower (-40.3%) than when there are two or more bidders.

Despite the importance of the issue of sample selection with binary endogenous variable, we are not aware of a study that deals simultaneously with these two issues. On the one hand, the problem of sample selection has been widely analyzed in the econometrics literature starting with the seminal work of Nobel price winner James Heckman, who proposed a method (Heckit) to correct sample selection bias. Van Hasselt (2005) has recently proposed a Bayesian Monte Carlo Markov Chain (MCMC) algorithm to make inference on the correlation coefficient of the sample selection model. The author conducts a Monte Carlo Study that shows that Gibbs sampling algorithm performs well regardless of whether the parameters of the model are fully identified or not.[2] On the other hand, Chakravarty and Li (2003) propose a Bayesian algorithm to test the effect of an endogenous binary variable on the profits of a trader (we are not aware of another similar study). They propose a simple Gibbs sampling algorithm that alternates between conditional posterior probability distribution of the parameters. They find no evidence of significant correlation between traders' private information and their profits.

Even when we analyze the issues of sample selection and binary endogenous explanatory variable separately, we know that maximum likelihood procedures might be unreliable, even though there is only one correlation coefficient to estimate. We are not aware of any study that deals with both issues at the same time as it would require three correlation coefficients to estimate. The existing maximum likelihood estimation procedures (such as simulated maximum likelihood) do not perform well with multiple correlation coefficients and sample selection (see Waelbroeck, 2005). This justifies our Bayesian approach.

[2] The Gibbs algorithm is an MCMC algorithm that iteratively draws from the conditional posterior distributions of the parameters and always accepts such draws.

There is a large number of articles on timber auctions in the literature. Instead of an exhaustive review, we focus on the regression analysis approach to stumpage appraisal. Unsold timber lots were not considered in early regression-based models for predicting stumpage prices (e.g. Jackson and McQuillan, 1979, McQuillan and Johnson-True, 1988). Buongiorno and Young (1984) modeled winning bids using OLS conditional on timber auctions that received at least two bids. Prescott and Puttock (1990) and Puttock, Prescott and Meilke (1990) propose a standard hedonic price function to forecast stumpage prices in Southern Ontario timber sales. However, as Huang and Buongiorno (1986) argued, the fact that some timber lots remained unsold is an important market information. Thus, transaction evidence appraisal models should include this market information to prevent biased predictions of market values. To take into account unsold timber, they propose censored regressions (Tobit model) since the reserve price is known and announced before the auctions in U.S. timber sales and they have information on the no-bid timber sales.

Beyond the treatment of unsold lots, the number of bidders also appears as an important variable in the estimation of the winning bid in stumpage appraisal literature. Indeed, the degree of competition in auctions has an impact on the bidding strategies. Participants do not necessarily know the actual number of bidders, but they bid according to the expected or potential competition (Brannman 1996). Many studies on timber auctions such as Jonhson (1979), Hansen (1986), Sendack (1991) empirically supports the auction theory prediction that there is a positive relationship between the number of bidders and the value of the highest bid. Sendack (1991) explicitly examines the impact of the number of bidders on the winning bid by including a transformation of the number of bids submitted as an explanatory variable. Assigning a dummy variable to each number of bidders category ($n = 1$, $n = 2$, …, $n = 11$), Brannman, Klein and Weiss (1987) obtained estimated coefficients that support first price auction theory: bid shading is decreasing with the number of bidders. Boltz, Carter and Jacobson (2002) highlight the importance of intra-lot heterogeneity on auction prices of mixed species lots from timber auctions in North Carolina. Tobit estimation results, corrected for the effects of bidder participation, for market conditions, for production costs and for the quality and species characteristics, show that increased heterogeneity leads to lower sale prices. Their study gives in some way an estimation of the opportunity cost for biodiversity. None of these studies endogenize participation.

However, to use stumpage appraisal models as predictive tools it is necessary to endogenize the actual number of bidders. Examining the impact of the (announced) reserve prices in sealed bid Federal timber auctions, Carter and Newman (1998) endogenize the number of bidders in a simultaneous-two-equations Tobit framework, but the expected number of bidders is determined strictly by the reserve price. [3] Niquidet and van Kooten (2004) seek to predict the fair market value of standing timber in British Columbia using a two-stage truncated regression procedure because they do not have sufficient information on no-bid auctions.

Thus, the main difference between the stumpage appraisal literature and our analysis is the fact that the seller's reserve price is not announced in French timber auctions, and so we can not use their two-stage procedure: we can not explain bidders' participation by the level of the reserve price. In this article, bidder participation directly depends on the characteristics of the timber lot. Moreover, we treat participation as a discrete variable.

Finally, our study contributes to the literature on timber value appraisal as well as to the econometric literature since we propose a methodology to assess the value of heterogeneous goods from sequential auctions with secret reserve price and endogenous participation. In the next section, we describe the institutional framework of French public timber auctions. Section 3 describes the data set. The methodology is detailed in section 4 and section 5 presents the results. Section 6 concludes our research.

2 Institutional framework

Competitive bidding is widely used in timber sales in France. In particular, the French National Public Forest Service (ONF[4]) uses sealed-bid first-price auctions to sell timber from public forest. Timber auctions of the ONF, which represent 40% of the timber sold each year in France, generally concern standing timber. The auction mechanism seems to be the best way to determine an "objective" or a fair market value for such a heterogeneous product.

[3] They treat the number of bidders as a continuous variable.

Timber auctions are sequential auctions of heterogeneous goods since many different lots (usually more than one hundred) are put on sale one after the other; the result of the auction of a lot is given before the next lot is put on sale. The first lot is usually randomly drawn, and then the auctioneer follows the catalogue order. The sale catalogue details all the lots. During a sale, bidders are not interested in every lot. Each bidder has a specific demand about species, volume, and quality. Thus, the number of bidders for a particular lot is fairly small and it is quite usual that there is only one bid or even no bid at all. Besides, bidders are asymmetric: they have different goals (sawyer, merchant, etc.), different business sizes, different needs and different localizations.

2.1 Heterogeneity in timber auctions

At harvesting time, the ONF does not choose the characteristics of the products. It has to sell what came out of the forest, which is heterogeneous by nature. Thus, lots are heterogeneous (different from one another), but they are also made up of heterogeneous wood. In particular in standing timber sales, a lot may contain many species of different diameter and of different quality. Auctioning such a product raises the problem of the optimal lot composition. The successive auctions correspond to different lots, but there might be interrelated. Some lots may be close substitutes. On the contrary, others may present synergies. For example, it may be only profitable for some buyers to harvest two or more lots that are close to each other.

Taking into account the heterogeneity of the lots raises many practical issues. First, although there is a catalogue published before the sale that detail the characteristics of the lots, potential buyers visit the lots they intend to buy. Moreover, since bidders are not guaranteed to obtain the lots they want, they have to prospect 5 to 10 times more lots. This leads to non-negligible prospecting and estimation costs for the bidders. These search costs, which are linked to the heterogeneity of the product, are wasteful from a social perspective. Reducing the cost of preparing a bid in timber auctions may increase the number of bidders. It is then possible that the seller would be better off sharing all the information she has and announcing

---

[4] ONF stands for Office National des Forêts.

a reserve price. In any case, the seller also needs to assess as correctly as possible the value of any given lot so as to define a relevant reserve price.

2.2 Secret reserve prices

Contrary to North American timber auctions, the reserve price of the seller is not announced in French public timber auctions. It is kept secret. This singular practice has been studied in the literature, but is difficult to justify theoretically. Elyakime, Laffont, Loisel and Vuong (1994) show in an independent private value auction model that the seller is strictly better for by determining a minimum bid instead of keeping the reserve price secret. According to their model, the seller is always better off announcing his reserve price, but if the reserve price is kept secret, then the optimal reserve price should be set to her reservation value. Nevertheless, the practice of a secret reserve price is sometimes justified either by the fact that announcing a reserve price reduces the participation of the bidders or by a common value component (Vincent, 1995). Risk aversion is also mentioned to justify a secret reserve price (Li and Tan, 2000). A lack of competition for some lots and the willingness of the ONF to maintain a reasonable timber price may also explain this practice. Finally, a secret reserve price may be used to prevent collusion between bidders at the reserve price.

In fact, we believe that the seller prefers not to announce and commit to any reserve price because she does not know her reservation value at the auction time. Indeed, a secret reserve price is reported for each lot in the database, but this price is not generally the seller's real reservation value since many lots are sold under the reserve price. This means that the French public Forest Service decides to sell a lot or not at the last moment and does not commit to any reserve price before the auction. So, the seller uses the bids to adjust her estimation of a lot value. Besides, she can lower the reserve price if she sees that many lots remain unsold. With this privilege, the seller keeps a certain flexibility to manage the sale, but that practice may be costly for the seller from an auction theory point of view. Without firm and credible commitment, the ONF may lose a part of the benefit of an auction. If the bidders anticipate that the seller can modify the rules of the game, then they will modify their bidding strategy, which may lower the efficiency of the bidding mechanism. Nevertheless, the fact that the seller updates her reserve prices translates the difficulty to assess the value of a lot. Announcing a reserve price might have negative consequences if the model used to set reserve prices is mis-specified. Indeed, a reserve price set too high can result in no bids, while a reserve price set too low may result in excessive bid shading since the number of bidders is usually low in timber auctions.

3 Data

The data set we use is part of the data collected by Costa and Préget (2004).

3.1 Fall 2003 Lorraine timber sales

The data set of Costa and Préget (2004) relies on the auction results of the ten Fall 2003 timber sales of Lorraine, a Region of the eastern part of France. A total of 2262 lots were put on sale. Since there are many differences between hardwood and softwood valuations, we select only pure hardwood lots, i.e. lots that are composed of more than 99% of hardwood. 1205 hardwood lots have been put on sale between September 9th and October 28th 2003. Lots may be very heterogeneous and made up of many species. The Herfindahl index is used to measure intra lot heterogeneity.[5] Out of the 1205 hardwood lots put on sale, only 48% of the lots are put on sale for the first time; thus 52% of the lots correspond to previously unsold lots.

At the end of the auctions, lots may be classified according to the auction results. A lot sold during the auction is said to be "auctioned", whereas the others are what we call "unsold lots". The percentage of unsold lots is 42% and shows a relatively difficult wood market environment in the Lorraine area during that period. It is useful to distinguish between lots that got one or more bids but have nevertheless been withdrawn by the seller and those that got no bid at all, referred to as the "no bid" category. Table 1 presents sale results.

[5] The Herfindahl index is the sum of the square volume proportion of each species. Here the number of species is limited to 7, then the Herfindahl index varies from 0.14 to 1. The more homogeneous the lot, the closer the index to one.

Table 1. Timber auction results

Auctioned lots		695 (58%)
Unsold lots	Withdrawn lots	318 (26%)
	No bid	192 (16%)
Total		1205 (100%)

3.2 Database

The database of Costa and Préget (2004) includes more than one hundred variables that represent a large part of the information available in the catalogues. It also includes private information from the ONF (harvesting conditions, quality of the lot, secret reserve price), data about the auction results (the number of bids, the auctioned prices) and computed data such as the Herfindahl index. This database is particularly rich. Moreover, it is exhaustive since it contains all the standing timber lots from public forests put on sale in the region during the fall of 2003. The following two tables give summary statistics of variables used in our econometric study.

Table 2. Descriptive statistics for binary variables

Variable	Mean
No restrictions	0.3718
Cutting	
arranged cutting	0.5270
other cutting	0.0440
selection cutting	0.0108
accidental products	0.0274
regeneration cutting	0.3909
Previously unsold	0.5178
Harvesting conditions	
easy logging & extraction	0.2722
normal logging	0.5876
difficult logging	0.0274
difficult logging & extraction	0.0797
very difficult logging & extraction	0.0315
Mitraille (scrap-iron, grape-shot from the first world war)	0.2257
Stand, crop	
high forest	0.2971
conversion of a stand	0.6241
coppice forest	0.0058
coppice with standards	0.0730
Landing area	
unarranged	0.8041
arranged	0.1593

none	0.0365
Quality	
very good	0.0407
good	0.3485
normal	0.4564
mediocre	0.1261
bad	0.0266

Table 3. Descriptive statistics for continuous variables

Variable	Mean	Std. Dev.	Min	Max
Surface (in hectare)	12.41	10.38	0.20	104.04
Number of trees	238.27	205.63	21	2259
Number of poles	267.07	663.76	0	11366
Herfindahl index	0.6007	0.1949	0.3337	1.0000
Stem volume of the mean-tree	1.0623	0.7314	0.0596	4.7190
Oak volume without crown	94.51	115.98	0	859.98
Beech volume without crown	136.83	164.09	0	1365.80
Other hardwood volume without crown	67.66	97.25	0	838.60
Crown hardwood volume	166.62	153.64	0	1196.47
Coppice volume	0.33	5.39	0	153.83
Relative order of the auction	0.50	0.29	0	1

All continuous variables are define in log except variables in percentage such as the Herfindahl index and the variable used to give the relative order of the auction in the sale and the stem volume of the mean-tree. 36% of the auctioned lots are sold at a price lower than the seller reserve price. These figures show that the seller does not commit to a credible reserve price and takes her decision to sale or not at the last moment. Thus, the reserve price of our data set has no clear signification.

Table 4 reports the number of lots for which there is no bid, one bid, two or more bids.

Table 4. Number of lots according to the number of bidders

Number of bids	0	1	2 and more	Total
Number of lots	192 (16%)	227 (19%)	786 (65%)	1205 (100%)

4 Methodology

We propose to estimate a hedonic price function based on the highest bids. The highest bid of an auction is not necessary a winning bid (and thus a market price) since the seller might withdraw the lot if she believes that the highest bid is too low. However, we choose to estimate the highest bid and not the sale price because the sale price is not independent from the seller's decision (because of the secrete reserve price) and is less informative about market demand.[6] As explained in the introduction, participation to timber auction raises two problems. First, many lots receive no bid and thus remain unsold at the end of the sale. Secondly, the number of bidders in an auction has an impact on the result of the auction. Particularly it makes a big difference if there is only one bidder (no competition) and if there are two or more bidders that compete for the same lot.[7] Nevertheless, participation depends on the (unobserved) characteristics of the lots and might be endogenous from an econometric point of view.

We propose a Bayesian Monte Carlo Markov Chain (MCMC) sampling algorithm that uses data augmentation and exploits the conditional posterior distributions of the parameters. Indeed, latent variables can be simulated and, conditional on these variables, the model is a simple Seemingly Unrelated Regression (SUR) model that is easy to deal with. We use a Metropolis step to draw from the conditional posterior distribution of the elements of the covariance matrix of the unobservable variables. So we propose a slightly different MCMC algorithm for the sample selection part of the model than Van Hasselt (2005). We write the latent model as a SUR model with an unequal number of observations and thus inference on the coefficients of the observed equation only relies on observations that are not censored.

We contribute to the literature on two points. First, we deal with three correlation coefficients because we have three unobservable variables in our model, while Chakravarty and Li (2003) and Van Hasselt (2005) only have to deal with one correlation coefficient. Secondly, both

[6] In Préget and Waelbroeck (2006a) we consider the auctioned price and unsold lots are taken into account in a sample selection model of Heckman. Nevertheless, the selection equation depends on the seller strategy, what we want to avoid here.
[7] Even when there is only one bidder, the submitted bid can not be too low because it has to reach the secret reserve price of the seller to be a winning bid.

articles reparameterize the elements of the covariance matrix that simplify the sampling procedure and speeds up the rate of convergence of the simulated Markov chain. Their algorithms might not be optimal with likelihood functions of irregular shapes. We have included a Metropolis step from the conditional posterior distribution of the covariance matrix that sometimes accepts draws that decrease the likelihood function.[8]

We analyze endogenous participation to French public timber auctions using a system of three equations. Equation (1) determines the probability that there is no bid in which case the bidders do not participate to the auction, i.e. the expected payoff of participating, $w_{1,i}$, is zero or negative. Thus, we define $y_{1,i} = 1$ if at least one bidder participates to the auction and $y_{1,i} = 0$ otherwise.

$$y_{1,i} = \begin{cases} 1 & \text{if } w_{1,i} > 0 \\ 0 & \text{if } w_{1,i} \leq 0 \end{cases} \tag{1}$$

where $w_{1,i} = x_{1,i}' \beta_1 + e_{1,i}$, β_1 is of dimension k_1 and $x_{1,i}$ is a set of control variables.

Equation (2) determines the probability that we only observe one bid, among the submitted lots, i.e. if there is no competition for that lot. We define $y_{2,i} = 1$ if there is only one bid and $y_{2,i} = 0$ otherwise (i.e. if there are two bids or more).[9]

$$y_{2,i} = \begin{cases} 1 & \text{if } w_{2,i} > 0 \\ 0 & \text{if } w_{2,i} \leq 0 \end{cases} \quad \text{if } y_{1,i} = 1 \tag{2}$$

where $w_{2,i} = x_{2,i}' \beta_2 + e_{2,i}$, β_2 is of dimension k_2 and $x_{2,i}$ is a set of control variables.

[8] See Chen *et al*. (2000).
[9] Generally, we only observe the binary endogenous variable in the selected sample.

Finally, equation (3) is the hedonic price equation that explains the highest bid $w_{3,i}$ as a function of lot characteristics and the endogenous participation variable $y_{2,i}$. Equation (3) is only observed for lots that have received at least one bid ($y_{1,i} = 1$).

$$w_{3,i} = z_{3,i}' \gamma_3 + y_{2,i} a_2 + e_{3,i} = x_{3,i}' \beta_3 + e_{3,i} \quad \text{observed for } y_{1,i} = 1 \qquad (3)$$

where $x_{3,i} = (z_{3,i}', y_{2,i})'$ and $\beta_3 = (\gamma_3', a_2)'$.

We assume that $\varepsilon_i = (e_{1,i}', e_{2,i}', e_{3,i}')'$ is normally distributed with mean (0, 0, 0)′ and covariance Σ for $i = 1, \ldots, n$:

$$\Sigma = \begin{bmatrix} 1 & r_{12} & r_{13}s_3 \\ r_{12} & 1 & r_{23}s_3 \\ r_{13}s_3 & r_{23}s_3 & s_3^2 \end{bmatrix}$$

Parameters r_{12}, r_{13} and r_{23} represent the correlations between the unobservable variables. Hence, r_{13} is the correlation coefficient of the Heckman sample selection procedure, while r_{23} is related to the lack of competition for the lot in the hedonic price equation. Parameter s_3^2 is the variance of $e_{3,i}$. Since probit equations (1) and (2) are not identified, we had to impose two restrictions. We chose to normalize the variances of the selection equation and of the endogenous binary variable to 1. These are standard restrictions in probit models.

We always observe $(x_{1,i}, y_{1,i})$, but we only observe $y_{2,i}$ and $w_{3,i}$ when $y_{1,i} = 1$.[10] Moreover, variables $w_{1,i}$ and $w_{2,i}$ are latent. The vector of explanatory variables can be stacked in order to write the (partially) latent model as a Seemingly Unrelated Regressions (SUR) model with an unequal number of observations. Let n_1 be the number of observations for which $y_{1,i} = 0$ and n_2 the number of observations such that $y_{1,i} = 1$, with $n = n_1+n_2$. We now assume for notational convenience that the data have been sorted according to the values of y_1. We also note the vector of binary dependent variables as $y = (y_1', y_2')'$. Let $\beta = (\beta_1', \beta_2', \beta_3')'$, $w_1 =$

[10] Depending on the data set, $x_{2,i}$ and $x_{3,i}$, may not be observed when $y_{1,i} = 0$. In that case, censored data are not used to make inference in equation (2) and (3).

$(w_{1,1}, \ldots, w_{1,n})'$, $w_2 = (w_{2,1}, \ldots, w_{2,n_2})'$, $w_3 = (w_{3,1}, \ldots, w_{3,n_2})'$ and define $w = (w_1', w_2', w_3')'$. We define ε_1, ε_2, ε_3 and ε in a similar fashion.

For notational convenience, we decompose the vectors of unobservable variables according to the selection process: $\varepsilon = (\varepsilon_{11}', \varepsilon_{12}', \varepsilon_2', \varepsilon_3')'$, where the second index equals to 1 if $y_{1,i} = 0$ and equals to 2 if $y_{1,i} = 1$. Thus the covariance of the unobservable variables is simply

$$\Omega = E\varepsilon\varepsilon' = \begin{bmatrix} I_{n_1} & 0 \\ 0 & \Sigma \otimes I_{n_2} \end{bmatrix}$$

where I_j denotes the identity matrix of dimension $j \times j$. Thus Ω^{-1} is readily obtained. We also decompose and stack the vector of the partially latent dependent variables as $w = (w_{11}', w_{12}', w_2', w_3')'$ and define similarly

$$X = \begin{bmatrix} x_{11} & 0 & 0 \\ x_{12} & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix} \quad (n_1+3n_2)\times(k_1+k_2+k_3)$$

The (partially) latent model can be written in matrix format:

$$w = X\beta + \varepsilon \tag{4}$$

Hence conditionally on w and Ω, the estimates of β are simply obtained by a Generalized Least Squares (GLS) regression of (4).[11] Moreover, matrices $X'\Omega^{-1}X$ and $X'\Omega^{-1}w$ required for the GLS estimates of the parameters of the model are easily computed.

The system of equations defined by (1) - (3) could be estimated by simulated maximum likelihood where multiple integrals are estimated by the GHK algorithm, a method to compute multiple normal integrals using recursive simulation draws. However, Waelbroeck (2005) has illustrated the poor performance of simulated maximum likelihood procedure compared to Bayesian MCMC methods that use the SUR structure of the latent model (4) to simulate draws from the joint posterior distribution of the parameters when there is data attrition.

The Metropolis-Gibbs sampling algorithm proceeds in 3 steps. The first step is a standard data augmentation step. We use a uniform prior for β, r_{12}, r_{13}, r_{23} and a non-informative prior for s_3: $p(\beta, r_{12}, r_{13}, r_{23}, s_3) \propto 1/s_3$.[12] To simplify notations we have dropped the dependence of Ω on Σ and the dependence of Σ on $(r_{12}, r_{13}, r_{23}, s_3)'$ when there is no ambiguity.

Step 1. $\mathbf{w}_1$, $\mathbf{w}_2$ | b, S, **$\mathbf{w}_3$, y, X**

In the first step, we only need to draw from w_1 and w_2 since w_3 is observed. When $y_{1,i} = 0$, we know that $w_{1,i} < 0$, hence for those observations ($i = 1, \ldots, n_1$), we draw $w_{1,i}$ from the standard truncated normal distribution with mean $\mathrm{x}_{1,i}'\beta_1$ and variance 1 truncated on $(-\infty, 0)$. We use the optimal algorithm of Robert (1995) to draw from the truncated normal distribution.[13] For the other observations ($i = n_1+1, \ldots, n$), we know that conditionally on β, Σ, y, X, $(w_{1,i}, w_{2,i}, w_{3,i})'$ has a joint normal distribution with mean $(\mathrm{x}_{1,i}'\beta_1,\ \mathrm{x}_{2,i}'\beta_2,\ \mathrm{x}_{3,i}'\beta_3)'$ and covariance Σ. Thus,

$$w_{1,i} \mid w_{2,i}, \beta, \Sigma, \mathrm{y}, \mathrm{w}_3, \mathrm{X} \sim TN(m_{1|23}, S_{1|23}; B_1)$$

where $TN(a, b; c)$ denotes the normal distribution with mean a, variance b truncated in subspace c and $B_1 = \{z_1 \in \mathrm{R}: z_1 > 0\}$. The conditional moments $m_{1|23}$ and $S_{1|23}$ are given by

---

[11] Since each stage contains different number of observations and generally different sets of explanatory variables, we can not estimate the SUR model with ordinary least squares regression applied to each latent equation separately.

[12] The choice of the prior distribution does not matter much when there are a large number of observations, which is usually the case for auction data. Moreover, using the uniform prior distribution provides a direct mean of comparison with the maximum likelihood procedures.

[13] Using the inverse c.d.f. method yielded unreliable results.

the standard formulas of the conditional distribution from a multivariate normal distribution. Similarly,

$$w_{2,i} \mid w_{1,i}, \beta, \Sigma, \mathrm{y}, \mathrm{w}_3, \mathrm{X} \sim TN(m_{2|13}, S_{2|13}; B_2)$$

where $B_2 = \{z \in \mathrm{R}: z \leq 0\}$ if $y_{2,i} = 0$ and $B_2 = \{z \in \mathrm{R}: z > 0\}$ if $y_{2,i} = 1$.

Step 2. b | S**, y, w, X**

As discussed in the presentation of the (partially) latent model, the conditional distribution of β is readily seen to be

$$\beta \mid \Sigma, \mathrm{y}, \mathrm{w}, \mathrm{X} \sim N((\mathrm{X}'\Omega^{-1}\mathrm{X})^{-1}\ \mathrm{X}'\Omega^{-1}\mathrm{w}, (\mathrm{X}'\Omega^{-1}\mathrm{X})^{-1}).$$

Step 3. S | b**, y, w, X**

The conditional posterior distribution of Σ is not standard,

$$\Sigma \mid \beta, \mathrm{y}, \mathrm{w}, \mathrm{X} \propto |\Sigma|^{-n2/2} \exp(-\varepsilon'\Omega^{-1}\varepsilon/2) \,/\, s_{3,}$$

but can be simulated using Metropolis step. Define $\sigma = (r_{12}, r_{13}, r_{23}, s_3)'$. We use a normal jumping distribution $\mathrm{N}(\sigma, \theta'\, \mathrm{I}_{4\times 4})$.[14]

[14] We set the elements of θ in the Metropolis-Hastings algorithm to obtain an acceptance rate between 0.2 and 0.25. In general a large step size decreases the speed of convergence of the algorithm but enables to get out of problematic areas of the likelihood function more quickly, while a small value would make the algorithm converge faster at the cost of getting stuck in undesirable areas. The range of values that we have used is standard for the number of parameters used in the application and was found to be a good compromise between the two effects mentioned above. Note that this range of acceptance rates has been shown to be optimal for MCMC algorithms that use a normal jumping distribution. For other sampling schemes, the optimal acceptance rate has to be computed and could be different from the above values. Draws that resulted in values of the correlation coefficients below -1 or above 1, as well as draws not resulting in a positive covariance matrix were rejected. Note also that we used a log transformation of the various probabilities in order to avoid numerical underflows.

5 Results

We first estimate the two probit equations (1) and (2) separately and run a Heckit procedure using sample selection equation (1) and hedonic bid equation (3) as benchmarks. Secondly, we compare these estimation results with the Bayesian estimation of parameters of equations (1), (2) and (3) using the MCMC algorithm. Only significant variables have been kept in each equation. The signs of the estimated coefficients are coherent and intuitive, except for the variable 'no restriction' for which the coefficient is surprisingly positive in equation (2) and negative in equation (3).

Table 5 gives the probability a lot will receive at least one bid.

Table 5 - Equation (1) Probit regression results of y_1

y_1		Coef.	Std. Err.
selection cutting & other cutting	**	-0.5148	0.2152
accidental products	***	-1.2061	0.2797
previously unsold	***	-2.7943	0.4102
mitraille	**	-0.2970	0.1365
number of trees	***	0.4839	0.0960
arranged landing area	***	0.5692	0.1654
good & very good quality	***	0.5054	0.1372
other hardwood volume without crown	***	-0.1239	0.0471
first sale	***	-1.2185	0.1838
_cons	**	1.3645	0.5623

Log-lik = -309.72

Table 6 gives the probability a lot will receive only one bid, *i.e.* the probability there is no competition for the lot, conditional on the fact that the lot received at least one bid.

Table 6 - Equation (2) Probit regression results of y_2

y_2		Coef.	Std. Err.
no restrictions	**	0.2513	0.1028
selection cutting & other cutting	**	0.4640	0.2185
previously unsold	***	0.6530	0.1000
easy logging & extraction	***	-0.5696	0.1213
Herfindahl index square	**	-0.6039	0.2550
mitraille	***	0.4508	0.1225
number of trees	***	-0.3038	0.1069
number of poles	**	-0.0764	0.0385
relative order of the auction	**	-0.3988	0.1766
arranged landing area	***	-0.4624	0.1463
no landing area	*	0.4259	0.2349
good & very good quality	**	-0.2224	0.1116
mediocre & bad quality	***	0.3800	0.1420
surface	***	0.3663	0.0979
oak volume without crown	***	-0.1823	0.0524
beech volume without crown	***	-0.1036	0.0409
stem volume of the mean-tree	**	-0.2853	0.1273
_cons	***	1.7231	0.4340

Log-lik = -439.34

Table 7 gives the estimation results of observed equation obtained by the Heckman methodology using the method of maximum likelihood. The hedonic equation is the estimated value of the log of the highest bid (equation (3)). The selection equation indicates the factors that influence whether a lot will receive at least one bid or not and estimated coefficients of this equation were already reported in Table 5. We also ran an OLS regression of equation (3) but results were very similar and are not reported. This is expected since the coefficient associated with the inverse Mills ratio is not significantly different from zero in Table 7. The model explains 83.3% of the variance of the highest bids, which is a good fit.

Table 7 - Equation (3) Heckman regression results of w_3

w_3 = log highest bid		Coef.	Std. Err.
surface * density	***	0.0057	0.0013
no restrictions	***	-0.0896	0.0305
other cutting	**	-0.1495	0.0736
accidental products	***	-0.5435	0.1150
regeneration cutting	***	0.0934	0.0316
previously unsold	***	-0.1712	0.0334
density	***	0.0117	0.0045
density*density/1000	***	-0.0718	0.0220
very difficult logging & extraction	***	-0.2651	0.0843
Herfindahl index	***	0.6037	0.1040
mitraille	***	-0.0913	0.0321
number of trees	***	0.1642	0.0446
number of poles * number of poles	***	-0.0061	0.0016
relative order of the auction	***	0.2164	0.0437
conversion of a stand	***	0.1919	0.0344
coppice forest & coppice with standards	***	0.1966	0.0541
no landing area	**	-0.1540	0.0683
very good quality	***	0.1908	0.0611
good quality	***	0.1314	0.0287
mediocre & bad quality	***	-0.1114	0.0406
surface	***	0.4456	0.0641
oak volume without crown	***	0.0571	0.0163
beech volume without crown	***	0.1631	0.0167
other hardwood volume without crown	***	0.0643	0.0100
crown hardwood volume	***	0.0729	0.0142
stem volume of the mean-tree	***	0.2807	0.0407
last sale	***	0.1748	0.0355
y_2	***	-0.4358	0.0315
_cons	***	4.5391	0.1891
r_{13}		-0.0431	0.1511
s_3		0.3796	0.0084
l		-0.0163	0.0574

Table 8 gives the Bayesian estimation of the 3-equation model. For each equation, we used exactly the same variables than before. In Préget and Waelbroeck (2006b) we estimate different model specifications. We found that estimations of the correlation coefficients r_{13} using the Heckit procedure are sensitive to the choice of the explanatory variables. Hence, it is hard to infer if there is a sample selection bias or not. Moreover, the Heckit procedure leads to unreliable inference on the parameter associated with the endogenous variable $y_{2,i}$. The Bayesian algorithm yields a remarkably stable coefficient for the binary endogenous variable and was able to deal with irregularities in the likelihood function.

Convergence of the MCMC algorithm was reach quickly. We removed the first 100000 iterations and kept the next 1000000 iterations for inference. Figures 1 to 3 display the marginal posterior distribution of the correlation coefficients. They all have a single mode. Comparing the estimation results in Table 8 to the results of the Heckman procedure given in Table 7 reveals interesting facts. Coefficients are quite similar in both methodology, but the Heckit procedure underestimates the negative effect of the endogenous variable $y_{2,i}$ to -.4358 compared to -.5159 for the Bayesian procedure. Furthermore, the associated standard error is underestimated by the Heckit procedure. In fact, the value of -.5159 does not lie in the 95% confidence interval of the estimated a_2 (associated with y_2) of Table 7.

Controlling for endogeneity and for the characteristics of the lots, we find that, on average, lots with only one bid receive a highest bid that is 40.3% below the highest bid of lots with two bids on more.[15]

Table 8 - Bayesian estimation of the 3-equation model

Variable		Coef.	Std. Dev.
Equation (1)			
_cons	***	1.5219	0.5856
selection cutting & other cutting	**	-0.5154	0.2161
accidental products	***	-1.2209	0.2818
previously unsold	***	-2.9648	0.4416
mitraille	**	-0.2979	0.1364
number of trees	***	0.4877	0.0961
arranged landing area	***	0.5791	0.1656
good & very good quality	***	0.5086	0.1375
other hardwood volume without crown	***	-0.1245	0.0475
first sale	***	-1.2301	0.1839
Equation (2)			
_cons	***	1.7890	0.4389
no restrictions	**	0.2520	0.1029
selection cutting & other cutting	**	0.4507	0.2197
previously unsold	***	0.6493	0.1057
easy logging & extraction	***	-0.5797	0.1208
Herfindahl index square	**	-0.6104	0.2559
mitraille	***	0.4494	0.1226
number of trees	***	-0.3117	0.1070
number of poles	**	-0.0770	0.0385
relative order of the auction	**	-0.4168	0.1774
arranged landing area	***	-0.4767	0.1464
no landing area	*	0.4176	0.2361
good & very good quality	**	-0.2190	0.1119
mediocre & bad quality	***	0.3796	0.1428
surface	***	0.3827	0.0991

[15] The percentage was obtained using the formula exp(-.5159)-1.

oak volume without crown	***	-0.1858	0.0523
beech volume without crown	***	-0.1111	0.0416
stem volume of the mean-tree	**	-0.2947	0.1281
Equation (3)			
_cons	***	4.6094	0.2067
surface * density	***	0.0058	0.0013
no restrictions	***	-0.0862	0.0313
other cutting	*	-0.1397	0.0759
accidental products	***	-0.5498	0.1165
regeneration cutting	***	0.0929	0.0321
previously unsold	***	-0.1625	0.0341
density	***	0.0116	0.0045
density*density/1000	***	-0.0714	0.0224
very difficult logging & extraction	***	-0.2628	0.0856
Herfindahl index	***	0.5871	0.1073
mitraille	***	-0.0845	0.0336
number of trees	***	0.1577	0.0459
number of poles * number of poles	***	-0.0063	0.0017
relative order of the auction	***	0.2083	0.0455
conversion of a stand	***	0.1911	0.0350
coppice forest & coppice with standards	***	0.1952	0.0550
no landing area	**	-0.1398	0.0714
very good quality	***	0.1857	0.0626
good quality	***	0.1266	0.0297
mediocre & bad quality	**	-0.1021	0.0426
surface	***	0.4551	0.0659
oak volume without crown	***	0.0557	0.0166
beech volume without crown	***	0.1589	0.0176
other hardwood volume without crown	***	0.0643	0.0101
crown hardwood volume	***	0.0701	0.0147
stem volume of the mean-tree	***	0.2749	0.0420
last sale	***	0.1802	0.0365
y_2	***	-0.5159	0.0913
r_{12}		0.0073	0.0478
r_{13}		-0.0133	0.1326
r_{23}		0.1231	0.1300
s_3	***	0.3888	0.0095

Figure 1 - Marginal posterior distribution of r_{12}

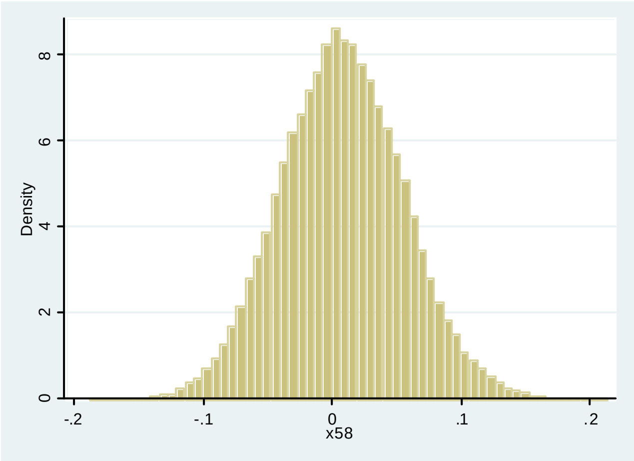


Figure 2 - Marginal posterior distribution of r_{13}

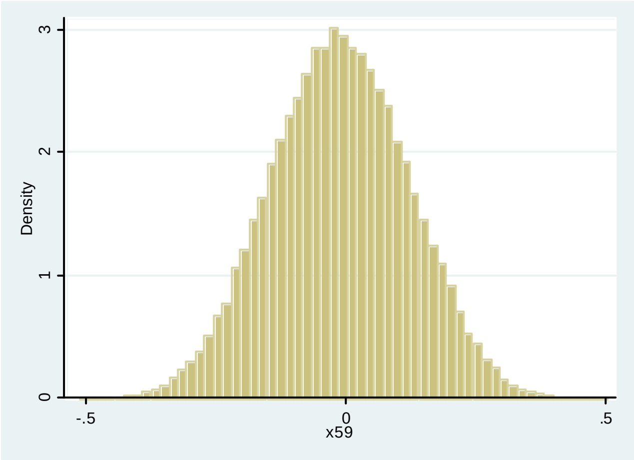
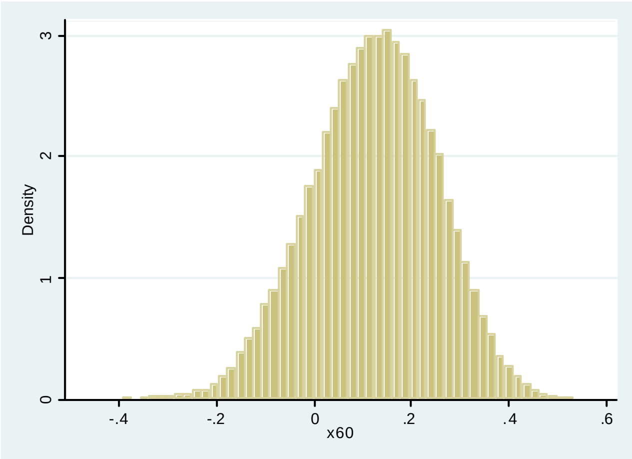


Figure 3 - Marginal posterior distribution of r_{23}



Two results are worth pointing out. First, the degree of intra-lot heterogeneity significantly increases the probability there will be competition for the lot, which increases the highest bid. Thus, concentrated lots with an Herfindahl index close to 1 (in other words lots that are not heterogeneous) have a sale premium: an increase of 1% of the concentration index increases the highest bid by 0.35% at the mean value of the Herfindahl index of 0.6. Second, the coefficient associated with the 'relative position of a lot' in the sale is significantly positive. This indicates that lots sold at the end of a sale receive higher bids, once we control for quality differences. This last result implies that the decline in prices often observed in sequential auctions is not present in our sample of timber auctions and that prices, on the contrary, increase for hardwood lots. This could be due to cautious behavior of the bidders in the beginning of the auctions and more aggressive bids at the end of the auctions. This interpretation is confirmed by two additional results. First, the probability that a lot does not receive any bid is higher in the first sale of the campaign: the variable 'first sale' has a significant negative impact in equation (1). Bidders wait and see at the beginning of the timber sale campaign. Second, the variable 'last sale' has a significant positive impact in the hedonic bid equation (3). This result reinforces the 'relative position of a lot' variable on a larger scale. Indeed, the highest bid increases during a sale (which is composed of many timber lots put on sale the same day), moreover the bids tends to be higher in the tenth sale (the one that took place the last day of the timber sale campaign).

6 Conclusion

Using detailed data set on timber auctions in Lorraine, we have highlighted the importance of endogenous participation on auction results, while the issue of unsold lots is not crucial in our data set. These results can not be generalized to other data sets, but we have proposed a methodology to deal with both issues at the same time. The econometric method can be extended in two directions. First, we can easily deal with truncated or censored depend variables in the hedonic price equation, when the reserve price is announced. Secondly, it is straightforward to model the endogenous participation variable as an ordinal variable that can take on a finite number of values using an ordinal probit framework. Our results can help public forest services to determine a relevant reserve price. In order to avoid auctions with 0 or 1 bid, the methodology could also be used to propose more attractive lots and to better understand demand factors: our hedonic price function for stumpage value gives interesting information about the implicit price of each lot characteristic for the optimal lot composition.

References

Boltz F., Carter D., Jacobson M, 2002, "Shadow pricing diversity in U.S. national forests", *Journal of Forest Economics*, *8*, p. 185-197.

Brannman L.E., 1996, "Potential competition and possible collusion in Forest Service Timber auctions", *Economic Inquiry*, *34*, p. 730-745.

Brannman L.E., Klein D., Weiss L.W., 1987, "The price effects of increased competition in auction markets", *Review of Economics and Statistics*, *69*, p. 24-32.

Buongiorno J., Young T., 1984, "Statistical appraisal of timber with an application to the Chequamegon National Forest", *Northern Journal of Applied Forestry*, *1*(4), p. 72-76.

Carter D.R., Newman D.H., 1998, "The impact of Reserve Prices in Sealed Bid Federal Timber Sale Auctions", *Forest Science*, *44*(4), p. 485-495.

Chakravarty S., Li K, 2003, "A Bayesian Analysis of Dual Trader Informativeness in Future Markets," *Journal of Empirical Finance*, *10*(3), p. 355-371.

Chen M-H, Shao Q-M, Ibrahim J. G., 2000, "Monte Carlo Methods in Bayesian Computation," Spinger.

Costa S., Préget R., 2004, "Étude de l'adéquation de l'offre en bois de l'Office National des Forêts à la demande de ses acheteurs", rapport final pour l'ONF, décembre 2004, 112 p.

Elyakime B., Laffont J.J., Loisel P., Vuong Q., 1994, "First-Price Sealed-Bid Auctions With Secret Reservation Prices", *Annales d'Economie et de Statistique*, *34*, p.115-141.

Hansen R., 1986, "Sealed-Bid versus Open Auctions: The Evidence", *Economic Inquiry*, *24*(1), p. 125-42.

Huang F.M., Buongiorno J., 1986, "Market Value of Timber When Some Offerings Are Not Sold: Implications for Appraisal and Demand Analysis", *Forest Science*, *32*(4), p. 845-854.

Jackson D.H., McQuillan A.G., 1979, "A technique for estimating timber value based on tree size, management variables, and market conditions", *Forest Science*, *25*(4), p. 620-626.

Johnson R.N., 1979, "Oral Auction versus Sealed Bids: An Empirical Investigation", *Natural Resources Journal*, *19*(2), p. 315-335.

Li H., Tan G., 2000, "Hidden Reserve Prices with Risk Averse Bidders", working paper, University of British Columbia.

McQuillan A.G., **Johnson-True C.,** 1988, "Quantifying Marketplace Characteristics for Use in Timber Stumpage Appraisal", *WJAF*, *3*(3), p. 66-69

Niquidet K., van Kooten C.G., 2004, "Are log markets competitive? Empirical evidence and implications for Canada-U.S. trade in softwood lumber", working paper 2004-04, Resource and Environmental economics and Policy Analysis (REPA) Research Group, Department of Economics, University of Victoria, British Columbia, 34p.

Préget R., Waelbroeck P., 2006a, "Timber appraisal from French public auctions: How to set the reserve price when there are unsold lots?", working paper Laboratoire d'Economie Forestière n°2006-05.

Préget R., Waelbroeck P., 2006b, "Sample Selection with Binary Endogenous Variable: A Bayesian Analysis of Participation to Timber Auctions", working paper Telecom Paris, ESS-06-08.

Prescott D.M., Puttock G.D., 1990, "Hedonic Price Functions for Multi-product Timber Sales in Southern Ontario", *Canadian Journal of Agricultural Economics*, *38*, p. 333-344.

Puttock G.D., Prescott D.M., Meilke K.D., 1990, "Stumpage Prices in Southwestern Ontario: A Hedonic Function Approach", *Forest Science*, *36*(4), p. 1119-1132.

Robert, C.P., 1995, "Simulation of truncated normal variables," *Statistics and Computing*, 5, p. 121-125.

Sendak P.E., 1991, "Timber sale value as a function of sale characteristics and number of bidders", Res. Pap. NE-657. Radnor, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. 7 p.

Van Hasselt M., 2005, "Bayesian Sampling Algorithms for the Sample Selection and

the Two-Part Models", working paper.

Vincent D., 1995, "Bidding Off the Wall: Why Reserve Prices May Be Kept Secret" *Journal of Economic Theory*, *65*, p. 575-584.

Waelbroeck P., 2005, "Computational Issues in the Sequential Probit Model: A Monte Carlo Study," *Computational Economics*, 26 (2), 141-161.