Örebro University

Örebro University School of Business

Master Thesis

Sune Karlsson

Panagiotis Mantalos

Regression with a Partially Censored Dependent Variable Under Simulation Study

Abel Shiferaw

19841203

Abstract

A partially censored regression analysis on a simulated Swedish exports data to other Europe countries and non-Europe countries at a certain threshold is considered. A suitable regression model is proposed and the parameters of this model are estimated using maximum likelihood estimation method. The bias and variance of the estimates are studied under different residual correlation coefficient and vector of constant. Findings from the simulation study revealed that an increase on the value of the constant vector or residual correlation coefficient enhance the estimates of the partially censored model in terms of reduced bias and variance.

Key words: Truncation, Censoring, partial censoring, Maximum likelihood Estimation,

Reparameterization, Gradient.

Table of contents

1. Introduction

Dealing with a model based statistical estimation using complete sample on the dependent variable will give quite good estimate of the parameter of the model and as a result yields a fairly nice prediction of the population total. In contrast with regression model over full samples there are also models with limited information on the dependent variable. Such models are classified as truncated model and censored model.

To start with let us consider definition of a truncated regression model, its kind as left truncation and right truncation and some counter examples.

(Hayashi, 2000) Defined truncated model (sample) as observations fulfilling a certain pre specified threshold condition being included in the sample while those violating the threshold are discarded from the sample. In this model the values of the dependent variable is observable only if it lies within the pre specified limit otherwise it is unobservable.

As stated in (Greene, 2008) truncation is a property of making inference about the true population from a sample obtained from subpopulation. It can be categorized as left, right, and double truncation based on the boundary of the threshold point.

Left truncation (truncation from below) occurs when values of the dependent variable are only observed above a certain threshold point (c). In other words no information can be known about the values of the variables below the specified threshold. As explained in (Mitra, 2013) a quite nice example of such a situation is truncation on total production of electronic goods as a specific minimum working hour is taken as a threshold point. In this situation it is difficult to know the exact number of goods the company has produced since some goods has already been missed from being counted as products due to not fulfilling the pre-specified threshold (minimum working hour), moreover no information is reviled on the quantity of the truncated goods.

One would face a right truncation, if values of the dependent variable cannot be observed above a certain pre specified threshold point(c). In other words one gets no information about variables above the upper bound c. As an elaboration to this truncation one can consider a study of incomes of the poor where households above a certain threshold of poverty line are excluded from the sample. A double truncation arise when the dependent variable is observed within a certain lower limit (c_1) and upper limit (c_2).

Next up one can demonstrate censoring along with its distinction with truncation in a manner as follows.

Analogous to truncation, censoring also occurs when the random values of the dependent variable is limited to a certain range of values. It arises from a situation where a certain experiment (measurement) does not have an intention of including every observation of unit in the sample.

Censoring can be from below or from above or else from both below or above depending on the interval of the threshold. As an example for left censoring (Tobin, 1958) considered household expenditure of expensive goods where low income individuals expenditure is censored to zero threshold, while for right censoring (Greene, 2008)considered ticket demand and sale for a certain event to be held in a particular stadium where the number of seat is a threshold point, in this situation the demand is censored from above to a threshold of maximum number of seat in a stadium.

(Heij, et al., 2004) Stated that, in censored regression, random samples are taken from the whole population while in truncated regression, random samples are taken from part of the population due to some restriction. In other words in censored model there is information for the whole sample while this does not appear in truncation. Moreover (Wooldridge, 2002) explained this distinction as , when censoring is under taken the explanatory variable(s) remain in the sample as the dependent variable is restricted to a certain threshold interval while during truncation the covariate is already excluded from being sampled when restriction is made on the dependent variable at a particular threshold .In general Censoring is a property of data selection mechanism while truncation is characteristic of the population that yields the data in hand.

In this paper one will study estimation of a regression model with partially censored dependent variable. By partially censored model means the dependent variable is a combination of full regression model and censored regression model.

In particular, a simulated export data on Swedish firm to other Europe countries and countries outside Europe is generated where exports to other Europe countries are only collected if a particular firm's exports exceed a threshold (currently four million five hundred thousand Swedish kronor). On the other hand exports to non-Europe countries are always recorded. This causes a problem for the statistical analysis of the export behavior of firms where the dependent variable is total exports.

In other words, for firm that exports more than the threshold to other Europe countries total export is available while, for firm that export less to other Europe countries only exports to non-Europe countries are available.

The aim of this thesis is to propose a regression model for total export and come up with suitable estimators then finally evaluates the properties of the estimates under different statistical tests.

The rest part of the thesis will proceed in a manner as follows. Chapter two concerns with searching for an appropriate estimator for a model exposed to either truncation or censoring. Chapter three deals with defining an appropriate model for the partially censored dependent variable and reparmetrization technique. Chapter four emphases the method how the estimation is carried out. Chapter five reveals the data generation procedure and test results achieved under the simulation study. Chapter six focuses on drawing conclusion from test results obtained in the simulation study.

2. Convenient Estimator for Dependent Variable Subject to Truncation or Censoring

2.1. Ordinary Least Square (OLS) Estimator

By Gauss-Markov Theorem and (Hayashi, 2000), when

$$y_t = \boldsymbol{X}'_t\beta + u_t \text{ , where } u_t \sim \text{N}(0, \sigma^2) \text{ and } \text{cov}(u_t, \boldsymbol{X}'_t) = 0 \text{ for all t=1,…,N}$$

then, an ordinary least square estimator(OLS) is the Best Linear Unbiased Estimator (BLUE).

However as stated by (Karlsson, 2006) in a regression model where the dependent random vector is subject to truncation or censoring, the introduction of OLS estimator results in a bias and inconsistent estimate for the parameter of interest since E$(u_t\,|\boldsymbol{X}'_t) \neq 0$.

Truncated case:

$$y_t^* = \boldsymbol{X}'_t\beta + u_t \text{ , where } u_t \sim \text{N}(0, \sigma^2)\text{, t=1,…,N}$$

Where,

$$\begin{aligned} y_t &= \boldsymbol{X}'_t\beta + u_t && \text{if } y_t^* > c \\ y_t &\text{ is unobservable} && \text{if } y_t^* \leq c\,. \end{aligned} \qquad (1)$$

As $y_t^* \sim N(\boldsymbol{X}'_t\beta\ , \sigma^2)$, then, it follows that E $(y_t^*|\boldsymbol{X}'_t) = \boldsymbol{X}'_t\beta$.

However, the mean of the observed , y_t , (the truncated mean) is not linear with respect to the covariate , $\boldsymbol{X}'_t$, . This can be shown in the following derivation as illustrated in (Heij, et al., 2004) and (Greene, 2008).

$$\text{E}\,(y_t \,|\boldsymbol{X}'_t, y^*_t > c) = \boldsymbol{X}'_t\beta + \sigma\, \text{E}\,(u_t \,|y^*_t > c)$$

$$= \boldsymbol{X}'_t\beta + \sigma\, \text{E}\,(u_t \,|u_t > c - \boldsymbol{X}'_t\beta)$$

$$= \boldsymbol{X}'_t\beta + \sigma\lambda_t$$

$$\neq \boldsymbol{X}'_t\beta$$

Where $\lambda_t = \frac{\phi(\frac{c - X'_t \beta}{\sigma})}{1 - \Phi(\frac{c - X'_t \beta}{\sigma})}$ known as the inverse Mills ratio, is a truncation correction factor that helps to minimize the bias of the parameter estimate(s) in OLS estimation of truncated and censored regression model and was introduced by John P. Mills.

As stated in (Heij, et al., 2004) and (Heckman, 1976) the omission of the truncation correction factor, λ_t, while estimating , y_t , on, $\boldsymbol{X}'_t$, OLS estimator yields biased and inconsistent estimate of the parameter, β. The magnitude of this bias as seen from the above derivation depends on the parameters β, σ^2, the covariates $\boldsymbol{X}'_t$ and the truncation point c.

Censored case:

$$y^*_t = \boldsymbol{X}'_t\beta + u_t\,, \text{ where } u_t \sim \text{N}\,(0, \sigma^2), \text{t=1,…,N}$$

Where,

$$\begin{aligned} y_t &= \boldsymbol{X}'_t\beta + u_t \quad && \text{if} \quad y^*_t > c \\ y_t &= \text{c} && \text{if} \quad y^*_t \leq c \;. \end{aligned} \tag{2}$$

The conditional mean of the observed y_t is a combination of discreet part and continuous part which is illustrated in (Greene, 2008) in the following manner.

$$\text{E}\,(y_t \,|\boldsymbol{X}'_t) = \text{E}\,(y^*_t|y^*_t \leq c\boldsymbol{X}'_t)\text{*p}\,(y^*_t \leq c) + \text{E}\,(y^*_t|y^*_t > c\boldsymbol{X}'_t)\text{*p}\,(y^*_t > c)$$

$$= \text{c*}\,\Phi(\frac{c - \boldsymbol{X}'_t\beta}{\sigma}) + (\boldsymbol{X}'_t\beta + \sigma\lambda_t)\,(1 - \Phi(\frac{c - \boldsymbol{X}'_t\beta}{\sigma}))$$

$$= \text{c*}\,\Phi(\frac{c - \boldsymbol{X}'_t\beta}{\sigma}) + \boldsymbol{X}'_t\beta(1 - \Phi(\frac{c - \boldsymbol{X}'_t\beta}{\sigma})) + \sigma\phi(\frac{c - \boldsymbol{X}'_t\beta}{\sigma})$$

$$\neq \boldsymbol{X}'_t\beta$$

The above derivation has revealed that ordinary least square estimation is inconvenient method for estimating parameter(s) of a model subjected to truncation or censoring hence searching for an appropriate estimator for a regression model with limited information on the dependent variable is the succeeding duty.

2.2. Maximum Likelihood Estimator (MLE)

In situations where missing information on the dependent variable exists for a certain regression model, estimation of parameter(s) of the model will require a special estimator so as to obtain fair estimate for the unknown population parameter. The commonly known estimator called maximum likelihood estimator (MLE) is applicable in a regression models where the dependent variable is exposed to either truncation or censoring.

As stated by (Aldrich, 1997) maximum likelihood estimator was introduced in 1922 by one of the most influential statistician of the 20th century called Sir Ronald Aylmer Fisher. MLE is well known by its asymptotic properties namely consistency, asymptotic normality and efficiency.

Formally the maximum likelihood estimator is the value of the parameter, β, that maximizes the likelihood as a function of the unknown parameter, $L(\beta|\boldsymbol{X}_t')$. That is, if we consider a random sample, x_t, from a particular pdf, $f(x_t, \beta)$, then, $\hat{\beta}_{mle}$, will yield maximum of $L(\beta|\boldsymbol{X}_t')$.

In most cases it is recommended to maximize the logarithm of the likelihood instead of the likelihood due to the monotonic behavior of the logarithmic function. Moreover the value of estimator that maximizes the log-likelihood will also maximize the likelihood.

Based on the random sample, x_t, and pdf, $f(x_t, \beta)$, one can provide the generalized form of the log-likelihood as follows.

$$lnL(\beta|\boldsymbol{X}_t') = ln(\prod_{t=1}^{n} f(x_t, \beta)) = \sum_{i=1}^{n} \ln f(x_t, \beta)$$

2.3. Log Likelihood for Left Truncated and Left Censored Dependent Variable

Once again consider a linear regression of latent dependent variable, y_t^*, on a matrix of nonrandom explanatory variable, $\boldsymbol{X}_t'$, and a normally distributed and uncorrelated stochastic disturbance, u_t .

That is;

$$y_t^* = \boldsymbol{X}_t'\beta + u_t, \text{ where } u_t \sim \mathrm{N}(0, \sigma^2), \text{ t=1,…,N}$$

According to (Greene, 2005) , for left truncated and left censored dependent variables shown in (1) and (2) respectively, one can rewrite their corresponding log likelihood as follows.

I. Log likelihood for left truncated dependent variable in (1)

$$lnL(y_t) = \sum_{t=1}^{N} \left[ln\frac{1}{\delta}\phi\left(\frac{y_t - \boldsymbol{X}_t'\beta}{\sigma}\right) - ln\left(1 - \Phi\left(\frac{c - \boldsymbol{X}_t'\beta}{\sigma}\right)\right) \right]$$

II. Log likelihood for left censored dependent variable in (2)

$$lnL(y_t) = \sum_{y_t^*>c} ln\frac{1}{\sigma}\phi\left(\frac{y_t - \boldsymbol{X}_t'\beta}{\sigma}\right) + \sum_{y_t^* \le c} ln\,\Phi\left(\frac{c - \boldsymbol{X}_t'\beta}{\sigma}\right)$$

3. Model

For the sack of convenient simulation to be carried out later in section 5, it is of interest to re-model total export, $e_t = e_{0,t} + e_{eu,t}$, based on logarithm of export to non-Europe countries, $lne_{0,t}$, and logarithm of export to other Europe countries, $lne_{eu,t}$, which can be shown in a manner as follows.

Let, $y_{0,t} = lne_{0,t}$ and, $y_{eu,t} = lne_{eu,t}$ then,

$$y_{0,t} = \boldsymbol{X}_t'\beta + u_t \tag{3}$$

And the latent, $y_{eu,t}^* = \boldsymbol{X}_t'\gamma + v_t$, where $y_{eu,t}^* = \begin{cases} y_{eu,t} & \text{if } y_{eu,t}^* > \mathrm{C} \\ 0 & \text{if } y_{eu,t}^* \le \mathrm{C} \end{cases}$ as $\mathrm{C} = lnc$ (4)

Where, $\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix} \right)$ and uncorrelated with, $\boldsymbol{X}_t'$.

Hence, based on $y_{0,t} = lne_{0,t}$ and, $y_{eu,t} = lne_{eu,t}$ the new model for, e_t, can be written as follows.

$$e_t = e_{0,t} + e_{eu,t}$$

$$e_t = exp(y_{0,t}) + exp(y_{eu,t})$$

$$e_t = exp(\boldsymbol{X}_t'\beta + u_t) + exp(\boldsymbol{X}_t'\gamma + v_t) \tag{5}$$

3.1. Log likelihood of the Model

The likelihood of logarithm of total export, $L(lne_t)$, can be written as a joint likelihood of the marginal likelihood of logarithm of export to non-Europe countries , $L(lne_{0,t})$, and conditional likelihood of logarithm of export to other Europe countries , $L(lne_{eu,t})$,as below.

Let $y_t = lne_t$,

As, $y_{0,t} = lne_{0,t}$, and, $y^*_{eu,t} = lne^*_{eu,t}$, then, define the likelihood for , y_t , as:

$$L\,(y_t) = L\,(y_{0,t})*\,L\,(y^*_{eu,t}|y_{0,t}) \qquad \text{where } L = \text{likelihood}$$

$$\rightarrow \qquad l\,(y_t) = l\,(y_{0,t}) + l\,(y^*_{eu,t}|y_{0,t}) \qquad \text{where } \; l = lnL \tag{6}$$

The marginal log likelihood of, $y_{0,t}$, in (3) can be written as:

$$l(y_{0,t}) = \sum_{t=1}^{N} ln\,\frac{1}{\sigma_u}\,\phi\left(\frac{y_{0,t} - \boldsymbol{X}'_t\beta}{\sigma_u}\right)$$

$$= \sum_{t=1}^{N} ln\,(\frac{1}{\sigma_u\sqrt{2\pi}}\;exp\;(\frac{-(y_{0,t} - \boldsymbol{X}'_t\beta\;)^2}{2{\sigma_u}^2})) \tag{7}$$

Up next consider $y^*_{eu,t} = \boldsymbol{X}'_t\gamma + v_t$;

The latent, $y^*_{eu,t} = \begin{cases} y_{eu,t} & \text{if } y^*_{eu,t} > \mathrm{C} \\ 0 & \text{if } y^*_{eu,t} \leq \mathrm{C} \end{cases}$, $\mathrm{C} = lnc$

Where $v_t\,|u_t \sim \mathrm{N}\,(\;\mu_v + \frac{\sigma_v}{\sigma_u}\,\rho(u_t - \mu_u), (1 - \rho^2)\,\sigma_v^2)$ and uncorrelated with, $\boldsymbol{X}'_t$.

As $\binom{\mu_v}{\mu_u} = \binom{0}{0}$ then, $v_t\,|u_t \sim \mathrm{N}\,(\frac{\sigma_v}{\sigma_u}\,\rho u_t\,, (1 - \rho^2)\sigma_v^2\,)$, where $\rho = \frac{cov(u_t\,,v_t)}{\sigma_u\sigma_v}$

The conditional log likelihood for $y^*_{eu,t}|y_{0,t}$ in (4) can be defined as follows:

$$l(y^*_{eu,t}|y_{0,t}) = \sum_{t=1\;\;y^*_{eu,t}>\mathrm{C}}^{N} ln\;\frac{1}{\sigma_v\sqrt{(1-\rho^2)}}\,\phi(\frac{y_{eu,t} - \boldsymbol{X}'_t\gamma - \frac{\sigma_v}{\sigma_u}\,\rho u_t}{\sigma_v\sqrt{(1-\rho^2)}}) +$$

$$\sum_{t=1\;\;y^*_{eu,t}\leq\mathrm{C}}^{N} ln\,\Phi\left(\frac{-(\boldsymbol{X}'_t\gamma + \frac{\sigma_v}{\sigma_u}\,\rho u_t\,)}{\sigma_v\sqrt{(1-\rho^2)}}\right)$$

$$= \sum_{t=1\;\;y^*_{eu,t}>\mathrm{C}}^{N} ln(\frac{1}{\sigma_v\sqrt{2\pi(1-\rho^2)}}\;exp\;(\frac{-(y_{eu,t} - \boldsymbol{X}'_t\gamma - \frac{\sigma_v}{\sigma_u}\rho\left(y_{0,t} - \boldsymbol{X}'_t\beta\right))^2}{2\sigma_v^2(1-\rho^2)}) +$$

$$\sum_{t=1\;\;y^*_{eu,t}\leq\mathrm{C}}^{N} ln\,\Phi\left(\frac{-\,(\boldsymbol{X}'_t\gamma + \frac{\sigma_v}{\sigma_u}\,\rho\left(y_{0,t} - \boldsymbol{X}'_t\beta\right))}{\sigma_v\sqrt{(1-\rho^2)}}\right) \tag{8}$$

Plugging in (7) and (8) in to (6) yields the joint log likelihood of, y_t , shown below;

$$l(y_t) = \sum_{t=1}^{N} ln\,(\frac{1}{\sigma_u\sqrt{2\pi}}\, exp\,(\frac{-(y_{0,t} - \boldsymbol{X}_t'\beta)^2}{2{\sigma_u}^2})) +$$

$$\sum_{t=1\ \ y_{eu,t}^*>\mathrm{C}}^{N} ln(\frac{1}{\sigma_v\sqrt{2\pi(1-\rho^2)}}\ exp\,(\frac{-(y_{eu,t} - \boldsymbol{X}_t'\gamma - \frac{\sigma_v}{\sigma_u}\rho\left(y_{0,t} - \boldsymbol{X}_t'\beta\right))^2}{2\sigma_v^2(1-\rho^2)})) +$$

$$\sum_{t=1\ \ y_{eu,t}^*\leq\mathrm{C}}^{N} ln\,\Phi\left(\frac{-(\boldsymbol{X}_t'\gamma + \frac{\sigma_v}{\sigma_u}\rho\left(y_{0,t} - \boldsymbol{X}_t'\beta\right))}{\sigma_v\sqrt{(1-\rho^2)}}\right) \tag{9}$$

3.2. Re-parameterization procedure for $l(y_t)$

Applying the concept of (Olson, 1978) re-parameterization one can rewrite the parameters; β, σ_u, σ_v, and, ρ, in (9) as a function of other parameter so as to make the log likelihood function easy for computation. These re-parameterization techniques can be handled in the following manner.

Let $\alpha = \frac{1}{\sigma_u}$, $\tau = \frac{\beta}{\sigma_u}$, $\gamma = \gamma$, $\omega = \frac{1}{\sigma_v\sqrt{(1-\rho^2)}}$ and $\theta = \frac{\rho}{\sqrt{(1-\rho^2)}}$

Where, $-\infty < \beta < \infty$, $-\infty < \gamma < \infty$, $-1 < \rho < 1$, $\sigma_u > 0$, $\sigma_v > 0$,

$$l(y_t) = \sum_{t=1}^{N} ln\,(\frac{\alpha}{\sqrt{2\pi}}\, exp\,(\frac{-(\alpha y_{0,t} - \boldsymbol{X}_t'\tau)^2}{2})) +$$

$$\sum_{t=1\ \ y_{eu,t}^*>\mathrm{C}}^{N} ln(\frac{\omega}{\sqrt{2\pi}}\, exp\,(\frac{-(\omega y_{eu,t} - \boldsymbol{X}_t'\gamma\omega - \theta\left(\alpha y_{0,t} - \boldsymbol{X}_t'\tau\right))^2}{2})) +$$

$$\sum_{t=1\ \ y_{eu,t}^*\leq\mathrm{C}}^{N} ln\,\Phi(-(\boldsymbol{X}_t'\gamma\omega + \theta(\alpha y_{0,t} - \boldsymbol{X}_t'\tau)))$$

$$l(y_t) = \sum_{t=1}^{N}(-\frac{1}{2}ln\,2\pi + ln\alpha - \frac{1}{2}\left(\alpha y_{0,t} - \boldsymbol{X}_t'\tau\right)^2) +$$

$$\sum_{t=1\ \ y_{eu,t}^*>\mathrm{C}}^{N}(-\frac{1}{2}ln\,2\pi + ln\omega - \frac{1}{2}(\omega y_{eu,t} - \boldsymbol{X}_t'\gamma\omega - \theta\alpha y_{0,t} + \boldsymbol{X}_t'\theta\tau)^2) +$$

$$\sum_{t=1\ \ y_{eu,t}^*\leq\mathrm{C}}^{N} ln\,\Phi(-\boldsymbol{X}_t'\gamma\omega - \theta\alpha y_{0,t} + \boldsymbol{X}_t'\theta\tau) \tag{10}$$

4. Method of Estimation

When simulating random sample for the dependent variables, $e_{0,t}$,(export to non-Europe countries) and, $e_{eu,t}$,(export to other Europe countries) it was invalid for negative valued observations to exist, as a result, the logarithm of export to non-Europe countries, $lne_{0,t}$, and logarithm of export to other Europe countries, $lne_{eu,t}$, was considered and a new model (5) was derived. In order to estimate this new model the following steps were followed.

Maximum likelihood estimator (MLE) for estimating parameters, $\beta, \gamma, \sigma_u, \sigma_v, \rho$, of model (5) was introduced. While deriving the likelihood of logarithm of total export , $L(\ y_t\ )$, the joint likelihood of, $L(\ y_{0t}\ )$, and, $L(\ y_{eut}\ )$, was considered. Moreover, reparametrization technique to modify the log likelihood for, $l\ (y_t\ )$, was followed.

A gradient search method known as Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was employed for optimizing (maximizing) the gradient of the log likelihood for, $l\ (y_t\ )$, in (10) . The initial guess(s) for the purpose of maximization was obtained from an OLS estimate of model (3) and (4) assuming no censoring in model (4). The gradient for (10) as well as R code for MLE of the parameters and R code for OLS estimates for initial values are presented in appendix (I) and (II) respectively.

5. Data and Test Result

A sample of five thousand observations was generated. The generated data consists of matrix of non-stochastic covariate, $\boldsymbol{X}_t'$, where, $\boldsymbol{X}_t'$, comprises vector of constants , vector of discrete and continuous uniform distribution where the discrete samples are chosen to be between ,1 ,and ,3 ,while the continuous one are between, 0.05 ,and, 0.5 .

Moreover, the data contains vector of random residual, u_t, generated from a normal distribution with zero mean and standard deviation, σ_u , and conditional vector of random residual, $v_t|u_t$, generated from a normal distribution with mean , $\frac{\sigma_v}{\sigma_u}\ \rho u_t$, and standard deviation, $\sqrt{(1\ -\ \rho^2)\sigma_v^2}$, where , ρ, is the residual correlation coefficient.

The assumed true parameter values were, $\beta = \begin{pmatrix}1\\1\\1\end{pmatrix},\ \gamma = \begin{pmatrix}1\\1\\1\end{pmatrix},\ \rho = 0.5,\ \ \sigma_u = 2$ and, $\sigma_v = 3.5$.

Furthermore, the parameters are restricted in the interval as follows:

$-\infty < \beta < \infty, \quad -\infty < \gamma < \infty, \quad -1 < \rho < 1, \quad \sigma_u > 0, \quad \sigma_v > 0.$

The test results shown below studies the properties of the parameter estimate(s) especially , $\hat{\gamma}$, and, $\widehat{\sigma_v}$, (estimates for censored regression model) in relation to bias and variance for five hundred replicates for different constant, and residual correlation coefficient, ρ,.

To start with, consider bias and variance of parameter estimates attained for a constant vector of ones on, $\boldsymbol{X}_t'$, and a residual correlation coefficient, ρ,= 0.5 as below.

Table 1. Bias and variance of parameter estimates for five hundred replicates for, ρ = 0.5, and constant = 1

ρ = 0.5 constant = 1	$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$	$\hat{\gamma} = \begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}$	$\hat{\rho}$	$\widehat{\sigma_u}$	$\widehat{\sigma_v}$
Bias	$\begin{pmatrix} -0.0006 \\ 0.0013 \\ 0.0030 \end{pmatrix}$	$\begin{pmatrix} -0.9463 \\ 0.1852 \\ 0.1515 \end{pmatrix}$	-0.0068	1.7927	0.6654
Variance	$\begin{pmatrix} 0.0109 \\ 0.0015 \\ 0.0553 \end{pmatrix}$	$\begin{pmatrix} 0.0645 \\ 0.0091 \\ 0.3788 \end{pmatrix}$	0.0002	0.0004	0.0031

Next up, an increase on the constant vector of , $\boldsymbol{X}_t'$, was considered and outcome obtained were as follows.

Table 2. Bias and variance of parameter estimates for five hundred replicates, for, ρ = 0.5, and constant = 2.

ρ = 0.5 constant = 2	$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$	$\hat{\gamma} = \begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}$	$\hat{\rho}$	$\widehat{\sigma_u}$	$\widehat{\sigma_v}$
Bias	$\begin{pmatrix} -0.0008 \\ 0.0006 \\ -0.0024 \end{pmatrix}$	$\begin{pmatrix} -0.3288 \\ 0.1349 \\ 0.0975 \end{pmatrix}$	-0.0049	-0.0001	0.4967
Variance	$\begin{pmatrix} 0.0022 \\ 0.0011 \\ 0.0450 \end{pmatrix}$	$\begin{pmatrix} 0.0109 \\ 0.0058 \\ 0.1753 \end{pmatrix}$	0.0001	0.0003	0.0027

Comparing table 1 and table 2 above reviles that an increase in constant (intercept) gives rise to a decrease in bias and variance for most of the estimates especially for the parameter estimates of the censored regression model ($\hat{\gamma}$ and $\widehat{\sigma_v}$) .

Moreover, an increasing on the residual correlation coefficient, ρ, has provided the result shown below.

Table 3. Bias and variance of parameter estimates for five hundred replicates, for, ρ = 0.8, and constant = 1.

ρ = 0.8 constant = 1	$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$	$\hat{\gamma} = \begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}$	$\hat{\rho}$	$\widehat{\sigma_u}$	$\widehat{\sigma_v}$
Bias	$\begin{pmatrix} 0.0113 \\ -0.0013 \\ -0.0288 \end{pmatrix}$	$\begin{pmatrix} -0.8082 \\ 0.1529 \\ 0.0456 \end{pmatrix}$	-0.0095	-0.0008	0.5820
Variance	$\begin{pmatrix} 0.0074 \\ 0.0011 \\ 0.0307 \end{pmatrix}$	$\begin{pmatrix} 0.0336 \\ 0.0053 \\ 0.1037 \end{pmatrix}$	0.00003	0.0003	0.0021

Once again the association between table 1 and table 3 claims that, an increase on the residual correlation coefficient, ρ, can result in reduction on bias and variance of the parameter estimates, mainly for the estimates of the censored regression model, $\hat{\gamma}$, and, $\widehat{\sigma_v}$.

In general, for comparing the enhancement in terms reduced bias and standard error, one can focus on the mean square error (MSE) of the parameter estimates from the results displayed in table 1, table 2 and table 3 as below.

Table 4. Mean square error for parameter estimates from table 1, table 2 and table 3.

	$\mathrm{MSE}\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$	$\mathrm{MSE}\begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix}$	MSE($\hat{\rho}$)	MSE($\widehat{\sigma_u}$)	MSE($\widehat{\sigma_v}$)
ρ = 0.5 constant = 1	$\begin{pmatrix} 0.0109 \\ 0.0015 \\ 0.0553 \end{pmatrix}$	$\begin{pmatrix} 0.9599 \\ 0.0433 \\ 0.4017 \end{pmatrix}$	0.0002	3.2141	0.4458
ρ = 0.5 constant = 2	$\begin{pmatrix} 0.0022 \\ 0.0011 \\ 0.0450 \end{pmatrix}$	$\begin{pmatrix} 0.1190 \\ 0.0239 \\ 0.1848 \end{pmatrix}$	0.0001	0.0003	0.2494
ρ = 0.8 constant = 1	$\begin{pmatrix} 0.0075 \\ 0.0011 \\ 0.0315 \end{pmatrix}$	$\begin{pmatrix} 0.6867 \\ 0.0286 \\ 0.1057 \end{pmatrix}$	0.0001	0.0003	0.3408

6. Conclusion

From, empirical findings of the simulation study of the partially censored regression model using maximum likelihood estimation one can observe that, an increase on the constant vector of, $\boldsymbol{X}_t'$, in model (5) has resulted a reduction on the percentage of censoring and this accordingly decreases the bias and standard error of the estimates of the model. The study also reviles that an increase in the residuals correlation coefficient, ρ, enhances the estimates of model (5).

Furthermore, by concentrating on the mean square error revealed in table 4 one can conclude that increasing the constant vector of, $\boldsymbol{X}_t'$, provides better result in terms of less bias and minimized variance for the parameter estimates of the censored regression model than increasing the residuals correlation coefficient, ρ,.

References

Aldrich, J., 1997. R. A. Fisher and the Making of Maximum Likelihood 1912-1922. *Statistical Science,* 12(3), pp. 162-176.

Greene, W., 2008. Limited Dependent Variables - Truncation, Censoring and Sample Selection. In: *Econometric Analysis.* Upper Saddle River : Prentice Hall, pp. 833-902.

Greene, W., 2008. Maximum Likelihood Estimation. In: *Econometrics Analysis.* Upper Saddle River: Prentice Hall, pp. 509-602.

Greene, W. H., 2005. *Censored Data and Truncated Distributions,* New York: NYU Working Paper No. EC-05-08..

Hayashi, F., 2000. *Econometrics.* Princeton: Princeton University Press.

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(4), pp. 475 - 492.

Heij, C. et al., 2004. *Econometric Methods with Applications in Business and Economics.* New York: Oxford University Press.

Karlsson, M., 2006. Estimators of regression parameters for truncated and censored data. *Metrika,* 63(3), p. 329–341.

Mitra, D., 2013. *Likelihood Inference for Left Truncated and Right Censored Lifetime Data,* Ontario: DigitalCommon@McMaster.

Olson, R. J., 1978. Note on the uniqueness of the maximum likelihood estimator for the. *Econometrica,* 46(5), pp. 1211-1215.

Tobin, J., 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica,* 26(1), pp. 24-36.

Wooldridge, J., 2002. *Econometric Analysis of Cross Section and Panel Data.* 2010 ed. Cambridge: Massachusetts Institute of Technology.

Appendix (I)

The gradient of the likelihood

$$l(y_t\ ) = \sum_{t=1}^{N}(-\frac{1}{2} ln\ 2\pi\ + ln\alpha\ - \frac{1}{2}\left(\alpha y_{0,t}\ - X_t'\tau\right)^2) +$$
$$\sum_{t=1\ \ y_{eu,t}^*>T}^{N} (-\frac{1}{2} ln\ 2\pi\ + ln\omega - \frac{1}{2}(\omega y_{eu,t}\ -\ X_t'\gamma\omega - \theta\alpha y_{0,t}\ + X_t'\theta\tau)^2\ )\ +$$
$$\sum_{t=1\ \ y_{eu,t}^*\le T}^{N}\ ln\ \Phi(-X_t'\gamma\omega - \theta\alpha y_{0,t}\ + X_t'\theta\tau)$$

Where $\alpha = \mathbf{z}_1^2$, $\tau = \mathbf{z}_{(2,k_1+1)}$, $\omega = \mathbf{z}_{(k_1+2)}^2$, $\gamma = \mathbf{z}_{(k_1+3,k_1+k_1+2)}$
$\theta = \mathbf{z}_{(k_1+k_1+3)}$. And k_1 stands for number of columns of X_t'.

That is;

$$l(y_t\ ) = \sum_{t=1}^{N}(-\frac{1}{2} ln\ 2\pi\ + ln\mathbf{z}_1^2\ - \frac{1}{2}\left(\mathbf{z}_1^2 y_{0,t}\ - X_t'\mathbf{z}_{(2,k_1+1)}\right)^2) +$$
$$\sum_{t=1\ \ y_{eu,t}^*>T}^{N} (-\frac{1}{2} ln\ 2\pi\ + ln\mathbf{z}_{(k_1+2)}^2 - \frac{1}{2}(\mathbf{z}_{(k_1+2)}^2 y_{eu,t}\ -\ X_t'\mathbf{z}_{(k_1+3,k_1+k_1+2)}\mathbf{z}_4^2 -$$
$$\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_1^2 y_{0,t}\ + X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)})^2\ )\ +$$
$$\sum_{t=1\ \ y_{eu,t}^*\le T}^{N}\ ln\ \Phi(-X_t'\mathbf{z}_{(k_1+3,k_1+k_1+2)}\mathbf{z}_{(k_1+2)}^2 - \mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_1^2 y_{0,t}\ + X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)})$$

With respect to $\mathbf{z}_1^2$, $\mathbf{z}_{(2,k_1+1)}$, $\mathbf{z}_{(k_1+2)}^2$, $\mathbf{z}_{(k_1+3,k_1+k_1+2)}$, and $\mathbf{z}_{(k_1+k_1+3)}$ can be shown as follows.

$$\frac{\partial\ l(e_t)}{\partial \mathbf{z}_1^2} = \sum_{t=1}^{N}\left(\frac{2}{\mathbf{z}_1} - 2\mathbf{z}_1 y_{0,t}\ \left(\mathbf{z}_1^2 y_{0,t}\ - X_t'\mathbf{z}_{(2,k_1+1)}\right)\right) +$$
$$\sum_{t=1\ \ y_{eu,t}^*>T}^{N}\ 2\mathbf{z}_1\mathbf{z}_{(k_1+k_1+3)}y_{0,t}\ \left(\mathbf{z}_{(k_1+2)}^2 y_{eu,t}\ -\ X_t'\ \mathbf{z}_{(k_1+3,k_1+k_1+2)}\mathbf{z}_{(k_1+2)}^2 - \mathbf{z}_7\mathbf{z}_1^2 y_{0,t}\ +\right.$$
$$\left. X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)}\right) +$$
$$\sum_{t=1\ \ y_{eu,t}^*\le T}^{N}\ - 2\lambda_t'\ \mathbf{z}_1\mathbf{z}_{(k_1+k_1+3)}y_{0,t}$$

Where $\lambda_t' = \frac{\phi\left(-X_t'\mathbf{z}_{(k_1+3,k_1+k_1+2)}\mathbf{z}_{(k_1+2)}^2 - \mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_1^2 e_{0,t} + X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)}\right)}{\Phi(-X_t'\mathbf{z}_{(k_1+3,k_1+k_1+2)}\mathbf{z}_{(k_1+2)}^2 - \mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_1^2 e_{0,t} + X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)})}$ is the inverse Mills ratio.

$$\frac{\partial\ l(e_t)}{\partial \mathbf{z}_{(2,k_1+1)}} = \sum_{t=1}^{N}\left(\mathbf{z}_1^2 y_{0,t}\ - X_t'\mathbf{z}_{(2,k_1+1)}\right)X_t' +$$
$$\sum_{t=1\ \ y_{eu,t}^*>T}^{N}\ -(\mathbf{z}_{(k_1+2)}^2 y_{eu,t}\ -\ X_t'\ \mathbf{z}_{(k_1+3,k_1+k_1+2)}\ \mathbf{z}_{(k_1+2)}^2 - \mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_1^2 y_{0,t}\ +$$
$$X_t'\mathbf{z}_{(k_1+k_1+3)}\mathbf{z}_{(2,k_1+1)})\ X_t'\mathbf{z}_{(k_1+k_1+3)} +$$
$$\sum_{t=1\ \ y_{eu,t}^*\le T}^{N}\ \lambda_t'\ X_t'\mathbf{z}_{(k_1+k_1+3)}$$

$$\frac{\partial\ l(e_t)}{\partial\ \mathbf{z}_{(k_1+2)}^2} = \sum_{t=1\ \ y_{eu,t}^*>T}^{N}\ \left(\ \frac{2}{\mathbf{z}_{(k_1+2)}} - 2\mathbf{z}_{(k_1+2)}(y_{eu,t}\ -\ X_t'\ \mathbf{z}_{(k_1+3,k_1+k_1+2)})\ (\mathbf{z}_4^2 y_{eu,t}\ -\right.$$

$$X_t' \, \mathbf{Z}_{(k_1+3,k_1+k_1+2)}\mathbf{Z}^2_{(k_1+2)} - \mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_1^2 y_{0,t} \;\; + X_t'\mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_{(2,k_1+1)})) +$$
$$\textstyle\sum_{t=1 \;\; y^*_{eu,t} \le T}^{N} \; -2\lambda_t' \, \mathbf{Z}_{(k_1+2)}\mathbf{Z}_{(k_1+3,k_1+k_1+2)} \, X_t'$$

$$\frac{\partial \; l(e_t)}{\partial \, \mathrm{z}_{(k_1+3,k_1+k_1+2)}} \; = \; \textstyle\sum_{t=1 \;\; y^*_{eu,t} > T}^{N} \; (\mathbf{Z}^2_{(k_1+2)} y_{eu,t} \; - \; X_t' \, \mathbf{Z}_{(k_1+3,k_1+k_1+2)}\mathbf{Z}^2_{(k_1+2)} - \mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_1^2 y_{0,t} \; + X_t'\mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_{(2,k_1+1)}) \, X_t'\mathbf{Z}^2_{(k_1+2)} + \textstyle\sum_{t=1 \;\; y^*_{eu,t} \le T}^{N} -\lambda_t' \; X_t'\mathbf{Z}^2_{(k_1+2)}$$

$$\frac{\partial \; l(e_t)}{\partial \mathrm{z}_{(k_1+k_1+3)}} \; = \; \textstyle\sum_{t=1 \;\; y^*_{eu,t} > T}^{N} \; (\mathbf{Z}^2_{(k_1+2)} y_{eu,t} \; - \; X_t' \, \mathbf{Z}_{(k_1+3,k_1+k_1+2)}\mathbf{Z}^2_{(k_1+2)} - \mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_1^2 y_{0,t} \; + X_t'\mathbf{Z}_{(k_1+k_1+3)}\mathbf{Z}_{(2,k_1+1)})(2\mathbf{Z}_1 y_{0,t} \; - X_t'\mathbf{Z}_{(2,k_1+1)}) + \textstyle\sum_{t=1 \;\; y^*_{eu,t} \le T}^{N} -\lambda_t' \, (2\mathbf{Z}_1 y_{0,t} \; - X_t'\mathbf{Z}_{(2,k_1+1)})$$

Appendix (II)

R code

```
### Rcode for data generation, parameter estimation,and
    bias/variance calculation####

set.seed(16)
N  = 5000
x1 = sample(1:3, N, replace=TRUE)
x2 = runif(N, 0.05, 0.5)
X  = matrix(cbind(x1,x2),nrow=N,ncol=2)
X  = cbind ( matrix(constant,N,1),X ) # Intercept

### calculation of ln_e_0t ####

#let y_0t =ln_e_0t
sigma_u_t = 2
mu_u_t = 0
u_t = rnorm(N, mu_u_t, sigma_u_t)
beta = 1
beta = as.vector(c(1,1,1))
y_0t = X%*%beta + u_t

### calculation of ln_e_eut ####

#let y_eut = ln_e_eut
sigma_v_t = 3.5
```

```
rho = rho
a = sqrt(1-rho^2)
mu_v_t_given_u_t = (sigma_v_t/sigma_u_t)*rho*u_t
sigma_v_t_given_u_t = a*sigma_v_t
v_t_given_u_t=rnorm(N, mu_v_t_given_u_t, sigma_v_t_given_u_t)

gamma = 1
gamma = as.vector(c(1,1,1))
z = X%*%gamma + v_t_given_u_t
threshhold = log(4.5)    #log(4.5)is in millions of Swedish
                          Kronor

w = 0
y_eut = ifelse(z < threshhold , w, z)
censored = ( y_eut == 0 )
ncensored = sum(censored)
X[censored,]
k1 = ncol(X)
#let y_t = ln_e_t

### reparametrized parameters ####

alpha = 1/sigma_u_t
tau = beta/sigma_u_t
omega = 1/sigma_v_t*a
gamma = gamma
theta = rho/a

### the likelihood function ####

im.exp = function(zeta,y_0t,y_eut,N,X)
 {

   k1 = ncol(X)
   alpha = zeta[1]^2
   tau = zeta[2:(k1+1)]
   omega = zeta[k1+2]^2
   gamma = zeta[(k1+3):(k1+k1+2)]
   theta = zeta[k1+k1+3]
   censored = ( y_eut == 0 )
   ncensored = sum(censored)
  loglik_y_t=N*(-1/2)*log(2*pi)+N*log(alpha)-
             1/2*sum(((alpha*y_0t)-(X %*%tau))^2) +
             (N-ncensored)*(-1/2)*log(2*pi)+
             (N-ncensored)*log(omega)-
```

```
                1/2*sum((omega*(y_eut[!censored]-(X[!censored,]
                %*%gamma))-theta*((alpha*y_0t[!censored])-
                (X[!censored,]%*%tau)))^2)+sum(log(pnorm(
                -1*omega*(X[censored,]%*%gamma)-
               (theta*alpha*y_0t[censored])+theta*(X[censored,]
                %*% tau))))

     return(-loglik_y_t)
   }

  im.exp_grad<- function(zeta,y_0t,y_eut,N,X)
  {
   k1 = ncol(X)
   alpha = zeta[1]^2
   tau = zeta[2:(k1+1)]
   omega = zeta[k1+2]^2
   gamma = zeta[(k1+3):(k1+k1+2)]
   theta = zeta[k1+k1+3]
   censored = ( y_eut == 0 )
   ncensored = sum(censored)
   lamda_t=(dnorm(-1*zeta[k1+2]^2*(X[censored,]
          %*%zeta[(k1+3):(k1+k1+2)])-
          (zeta[k1+k1+3]*zeta[1]^2*y_0t[censored]) +
          zeta[k1+k1+3]*(X[censored,]%*%
          zeta[2:(k1+1)])))/(pnorm(-1*zeta[k1+2]^2*(X[censored,]
          %*%zeta[(k1+3):(k1+k1+2)])-
          (zeta[k1+k1+3]*zeta[1]^2*y_0t[censored])+
          zeta[k1+k1+3]*(X[censored,] %*% zeta[2:(k1+1)])))

   FSH = (zeta[1]^2*y_0t - (X%*%zeta[2:(k1+1)]))

   JMM = (zeta[k1+2]^2*(y_eut[!censored]-(X[!censored,]%*%
          zeta[(k1+3):(k1+k1+2)]))-
          zeta[k1+k1+3]*((zeta[1]^2*y_0t[!censored])-
         (X[!censored,] %*% zeta[2:(k1+1)])))

  Grad=c(N*(2/zeta[1])-sum(2*zeta[1]*y_0t*FSH)+
       sum(2*zeta[1]*zeta[k1+k1+3]*y_0t[!censored]*JMM) +
       sum(-2*lamda_t*zeta[1]*zeta[k1+k1+3]*y_0t[censored]),

       t(X)%*% FSH + -1*zeta[k1+k1+3]*(t(X[!censored,] )%*%JMM)
       + zeta[k1+k1+3]*(t(X[censored,])%*%lamda_t),

      (N-ncensored)*(2/zeta[k1+2])-
       sum(2*zeta[k1+2]*JMM*(y_eut[!censored]-
```

```
    (X[!censored,]%*%zeta[(k1+3):(k1+k1+2)])))-
     2*zeta[k1+2]*crossprod(lamda_t,(X[censored,]
     %*%zeta[(k1+3):(k1+k1+2)])),

    zeta[k1+2]^2*(t(X[!censored,])%*%JMM)-
    zeta[k1+2]^2*(t(X[censored,])%*%lamda_t),

    sum(JMM*(2*zeta[1]*y_0t[!censored]-
    (X[!censored,]%*%zeta[2:(k1+1)]))) +
    sum(-1*lamda_t*(2*zeta[1]*y_0t[censored]-
    (X[censored,]%*%zeta[2:(k1+1)]))))

   return(-grad)
  }
   opt_grad=optim(c(,,,,,,,,),y_0t=y_0t,y_eut=y_eut,X=X,N=N,
                  im.exp, im.exp_grad, method = "BFGS",
                  control=list(maxit=400))
   opt_grad

  alpha_hat = opt_grad$par[1]
  tau_hat =   opt_grad$par[2:(k1+1)]
  omega_hat = opt_grad$par[k1+2]
  gamma_hat = opt_grad$par[(k1+3):(k1+k1+2)]
  theta_hat = opt_grad$par[k1+k1+3]

  #print(c(alpha_hat,tau_hat, omega_hat, gamma_hat, theta_hat))

  ### real estimates ####

  sigma_u_t_hat = 1/(alpha_hat)^2
  beta_ha t= tau_hat* sigma_u_t_hat
  gamma_hat = gamma_hat
  rho_hat = sqrt(theta_hat^2/(1+theta_hat^2))
  sigma_v_t_hat = 1/((omega_hat)^2*sqrt(1-rho_hat^2))

  print(c(beta_hat,gamma_hat,rho_hat,sigma_u_t_hat,
          sigma_v_t_hat))

### parameter estimates for 500 replicates ####

  result = matrix(0,500,9)
  x1 = sample(1:3, N, replace=TRUE)
  x2 = runif(N, 0.05, 0.5)
  X = matrix(cbind(x1,x2),nrow=N,ncol=2)
  X = cbind ( matrix(constant ,N,1),X )
```

```
for(i in 1:500){

  sigma_u_t = 2
  mu_u_t = 0
  u_t = rnorm(N, mu_u_t, sigma_u_t)
  beta = 1
  beta = as.vector(c(1,1,1))
  y_0t = X%*%beta + u_t
  sigma_v_t = 3.5
  rho = rho
  a = sqrt(1-rho^2)
  mu_v_t_given_u_t = (sigma_v_t/sigma_u_t)*rho*u_t
  sigma_v_t_given_u_t = a*sigma_v_t
  v_t_given_u_t=rnorm(N, mu_v_t_given_u_t, sigma_v_t_given_u_t)
  gamma = 1
  gamma = as.vector(c(1,1,1))
  z = X%*%gamma + v_t_given_u_t
  threshhold = log(4.5)   #log(4.5)is in millions of Swedish
                              Kronor
  w = 0
  y_eut = ifelse(z < threshhold , w, z)
  censored = ( y_eut == 0 )
  ncensored = sum(censored)
  X[censored,]

  alpha = 1/sigma_u_t
  tau = beta/sigma_u_t
  omega = 1/(sigma_v_t*a)
  gamma = gamma
  theta = rho/a

  opt_grad=optim(c(,,,,,,,,),y_0t=y_0t,y_eut=y_eut,X=X,N=N,
                 im.exp, im.exp_grad, method = "BFGS",
                 control=list(maxit=400))
  opt_grad

 result[i,1:9]<-opt_grad$par
}

 k1 = ncol(X)
 alpha_hat_i = result[,1]
 tau_hat_i = result[,2:(k1+1)]
 omega_hat_i = result[,k1+2]
 gamma_hat_i = result[,(k1+3):(k1+k1+2)]
 theta_hat_i = result[,k1+k1+3]
```

```
 ###  real estimtes in for loop  ####

 res = matrix(0,500,9)
 for(i in 1:500){

  res[,k1+k1+2] = 1/(alpha_hat_i)^2
  res[,1:k1] = tau_hat_i* res[,k1+k1+2]
  res[,(k1+1):(k1+k1)] = gamma_hat_i
  res[,k1+k1+1] = sqrt(theta_hat_i^2/(1+ theta_hat_i^2))
  res[,k1+k1+3] = 1/((omega_hat_i)^2*sqrt(1-res[,k1+k1+1]^2))
}
res

 beta_hat_i = res[,1:k1]
 gamma_hat_i = res[,(k1+1):(k1+k1)]
 rho_hat_i = res[,k1+k1+1]
 sigma_u_t_hat_i = res[,k1+k1+2]
 sigma_v_t_hat_i = res[,k1+k1+3]

### Testing bias of estimates for 500 replictaes ####

colMeans(beta_hat_i)-beta
colMeans(gamma_hat_i)- gamma
mean(rho_hat_i)- rho
mean(sigma_u_t_hat_i)- sigma_u_t
mean(sigma_v_t_hat_i)- sigma_v_t

### Testing variance of estimates for 500 replicates ####

apply(beta_hat_i, 2, var)
apply(gamma_hat_i, 2, var)
var(rho_hat_i)
var(sigma_u_t_hat_i)
var(sigma_v_t_hat_i)

### OLS estimates used as initial values ####

set.seed(16)
N = 5000
x1 = sample(1:3, N, replace=TRUE)
x2 = runif(N, 0.05, 0.5)
X = matrix(cbind(x1,x2),nrow=N,ncol=2)
X = cbind ( matrix(constant,N,1),X ) # Intercept

### calculation of ln_e_0t ####
```

```
#let y_0t = ln_e_0t
sigma_u_t = 2
mu_u_t = 0
u_t = rnorm(N, mu_u_t, sigma_u_t)
beta = 1
beta = as.vector(c(1,1,1))
y_0t = X%*%beta + u_t

### calculation of ln_e_eut ####

#let y_eut = ln_e_eut
sigma_v_t = 3.5
rho = rho
a = sqrt(1-rho^2)
mu_v_t_given_u_t = (sigma_v_t/sigma_u_t)*rho*u_t
sigma_v_t_given_u_t = a*sigma_v_t
v_t_given_u_t = rnorm(N, mu_v_t_given_u_t, sigma_v_t_given_u_t)
gamma = 1
gamma = as.vector(c(1,1,1))
z = X%*%gamma + v_t_given_u_t

# calculate the estimates BETA HAT
A = t(X)%*%X
Ainv = solve (A)
betahat = Ainv %*% ( t(X)%*%y_0t )
resl = y_0t - X%*%betahat
sse = crossprod( resl )
s_u = sqrt( sse/(N-3) )
betahat = as.vector(betahat)
s_u = as.vector(s_u)

# calculate the estimates GAMMA HAT
A = t(X)%*%X
Ainv = solve (A)
gammahat = Ainv %*% ( t(X)%*%z )
resl = z - X%*%gammahat
sse = crossprod( resl )
s_v = sqrt( sse/(N-3) )
gammahat = as.vector(gammahat)
s_v = as.vector(s_v)

resl1 = residuals(lm(y_0t ~ X))
resl2 = residuals(lm(z ~ X))
cor(resl1,resl2)
rhohat = cor(resl1,resl2)
```

```
rhohat
a = sqrt(1-rhohat^2)
a
alpha0 = 1/s_u
tau0   = betahat/s_u
omega0 = 1/s_v*a
gamma0 = gammahat
theta0 = rhohat/a

### initial values ####

print(c(alpha0, tau0, omega0, gamma0, theta0))
```