Estimating Censored Regression Models in R using the censReg Package

Arne Henningsen
University of Copenhagen

Abstract

We demonstrate how censored regression models (including standard Tobit models) can be estimated in R using the add-on package **censReg**. This package provides not only the usual maximum likelihood (ML) procedure for cross-sectional data but also the random-effects maximum likelihood procedure for panel data using Gauss-Hermite quadrature.

Keywords: censored regression, Tobit, econometrics, R.

1. Introduction

In many statistical analyses of individual data, the dependent variable is censored, e.g. the number of hours worked, the number of extramarital affairs, the number of arrests after release from prison, purchases of durable goods, or expenditures on various commodity groups (Greene 2008, p. 869). If the dependent variable is censored (e.g. zero in the above examples) for a significant fraction of the observations, parameter estimates obtained by conventional regression methods (e.g. OLS) are biased. Consistent estimates can be obtained by the method proposed by Tobin (1958). This approach is usually called "Tobit" model and is a special case of the more general censored regression model.

This paper briefly explains the censored regression model, describes function `censReg` of the R package **censReg**, and demonstrates how this function can be used to estimate censored regression models.

There are also some other functions for estimating censored regression models available in R. For instance function `tobit` from the **AER** package (Kleiber and Zeileis 2008, 2009) and function `cenmle` from the **NADA** package are front ends to the `survreg` function from the **survival** package. Function `tobit` from the **VGAM** package estimates the censored regression model by using its own maximum likelihood routine. Function `MCMCtobit` from the **MCMCpack** package uses the Bayesian Markov Chain Monte Carlo (MCMC) method to estimate censored regression models.

2. Censored regression model for cross-sectional data

2.1. Standard Tobit model

In the standard Tobit model (Tobin 1958), we have a dependent variable y that is left-censored at zero:

$$y_i^* = x_i'\beta + \varepsilon_i \tag{1}$$

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } y_i^* > 0 \end{cases} \tag{2}$$

Here the subscript $i = 1, \ldots, N$ indicates the observation, y_i^* is an unobserved ("latent") variable, x_i is a vector of explanatory variables, β is a vector of unknown parameters, and ε_i is an disturbance term.

2.2. Censored regression model

The censored regression model is a generalisation of the standard Tobit model. The dependent variable can be either left-censored, right-censored, or both left-censored and right-censored, where the lower and/or upper limit of the dependent variable can be any number:

$$y_i^* = x_i'\beta + \varepsilon_i \tag{3}$$

$$y_i = \begin{cases} a & \text{if } y_i^* \leq a \\ y_i^* & \text{if } a < y_i^* < b \\ b & \text{if } y_i^* \geq b \end{cases} \tag{4}$$

Here a is the lower limit and b is the upper limit of the dependent variable. If $a = -\infty$ or $b = \infty$, the dependent variable is not left-censored or right-censored, respectively.

2.3. Estimation Method

Censored regression models (including the standard Tobit model) are usually estimated by the Maximum Likelihood (ML) method. Assuming that the disturbance term ε follows a normal distribution with mean 0 and variance σ^2, the log-likelihood function is

$$\begin{aligned} \log L = \sum_{i=1}^{N} \Bigg[& I_i^a \log \Phi\left(\frac{a - x_i'\beta}{\sigma}\right) + I_i^b \log \Phi\left(\frac{x_i'\beta - b}{\sigma}\right) \\ & + \left(1 - I_i^a - I_i^b\right) \left(\log \phi\left(\frac{y_i - x_i'\beta}{\sigma}\right) - \log \sigma\right) \Bigg], \end{aligned} \tag{5}$$

where $\phi(.)$ and $\Phi(.)$ denote the probability density function and the cumulative distribution function, respectively, of the standard normal distribution, and I_i^a and I_i^b are indicator functions with

$$I_i^a = \begin{cases} 1 & \text{if } y_i = a \\ 0 & \text{if } y_i > a \end{cases} \tag{6}$$

$$I_i^b = \begin{cases} 1 & \text{if } y_i = b \\ 0 & \text{if } y_i < b \end{cases} \tag{7}$$

The log-likelihood function of the censored regression model (5) can be maximised with respect to the parameter vector $(\beta', \sigma)'$ using standard non-linear optimisation algorithms.

2.4. Implementation in function `censReg`

Censored regression models can be estimated in R with function `censReg`, which is available in the **censReg** package (Henningsen 2011). The most important steps done by the `censReg` function are:

1. perform basic checks on the arguments provided by the user
2. prepare the data for the estimation, i.e. the vector of the dependent variable $y = (y_1, \ldots, y_N)'$ and the matrix of the regressors $X = (x_1', \ldots, x_N')'$
3. obtain initial values of the parameters β and σ from an OLS estimation using function `lm` (if no initial values are provided by the user)
4. define a function that calculates and returns the log-likelihood value and its gradients[1] given the vector of parameters $(\beta', \sigma)'$
5. call function `maxLik` of the **maxLik** package (Toomet and Henningsen 2010) for the maximisation of the likelihood function
6. add class `"censReg"` to the returned object

2.5. Using function `censReg`

Before function `censReg` can be used, the **censReg** package (Henningsen 2011) must be loaded:

```
R> library( "censReg" )
```

The user interface (e.g. function arguments and printed output) of function `censReg` follows rather closely the user interface of function `tobit` from the **AER** package (Kleiber and Zeileis 2008, 2009). The first argument of both functions is `formula`. It is the only mandatory argument and must provide a symbolic description of the model to be fitted. The optional argument `data` can be used to provide a data set (`data.frame`) that contains the variables used in the estimation. We demonstrate the usage of `censReg` by replicating an example given in Kleiber and Zeileis (2008, p. 142). The data used in this example are available in the data set `Affairs` that is included in the R package **AER** (Kleiber and Zeileis 2008, 2009). This data set can be loaded by the following command:

```
R> data( "Affairs", package = "AER" )
```

In the example of Kleiber and Zeileis (2008, p. 142), the number of a person's extramarital sexual intercourses ("affairs") in the past year is regressed on the person's age, number of years married, religiousness, occupation, and own rating of the marriage. The dependent variable is left-censored at zero and not right-censored. Hence, this is a standard Tobit model. It can be estimated by following command:

[1] The gradients of the log-likelihood function are presented in appendix A.1.

```
R> estResult <- censReg( affairs ~ age + yearsmarried + religiousness +
+    occupation + rating, data = Affairs )
```

Function `censReg` returns an object of class `"censReg"`. The corresponding `summary` method can be used to obtain summarized estimation results:

```
R> summary( estResult )

Call:
censReg(formula = affairs ~ age + yearsmarried + religiousness +
    occupation + rating, data = Affairs)

Observations:
         Total  Left-censored     Uncensored Right-censored
           601            451            150              0

Coefficients:
             Estimate Std. error t value  Pr(> t)
(Intercept)   8.17420    2.74145   2.982  0.00287 **
age          -0.17933    0.07909  -2.267  0.02337 *
yearsmarried  0.55414    0.13452   4.119 3.80e-05 ***
religiousness -1.68622   0.40375  -4.176 2.96e-05 ***
occupation    0.32605    0.25442   1.282  0.20001
rating       -2.28497    0.40783  -5.603 2.11e-08 ***
logSigma      2.10986    0.06710  31.444  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Newton-Raphson maximisation, 7 iterations
Return code 1: gradient close to zero (gradtol)
Log-likelihood: -705.5762 on 7 Df
```

In case of a censored regression with left-censoring not at zero and/or right-censoring, arguments `left` (defaults to zero) and `right` (defaults to infinity) can be used to specify the limits of the dependent variable. A lower (left) limit of minus infinity (`-Inf`) and an upper (right) limit of infinity (`Inf`) indicate that there is no left-censoring and right-censoring, respectively. For instance, minus the number of extramarital sexual intercourses is not left-censored but right-censored at zero. The same model as above but with the negative number of affairs as the dependent variable can be estimated by

```
R> estResultMinus <- censReg( I( - affairs ) ~ age + yearsmarried + religiousness +
+    occupation + rating, left = -Inf, right = 0, data = Affairs )
```

This estimation returns β parameters that have the opposite sign of the β parameters estimated in the original model, but the (logarithmised) standard deviation of the residuals remains unchanged.

```
R> cbind( coef( estResult ), coef( estResultMinus ) )
```

```
                    [,1]       [,2]
(Intercept)    8.1741974 -8.1741974
age           -0.1793326  0.1793326
yearsmarried   0.5541418 -0.5541418
religiousness -1.6862205  1.6862205
occupation     0.3260532 -0.3260532
rating        -2.2849727  2.2849727
logSigma       2.1098592  2.1098592
```

2.6. Methods for objects of class "censReg"

The **censReg** package provides several methods for objects of class "censReg": the print method prints some basic information on the estimated model, the coef method returns the coefficient vector, the vcov method returns the variance-covariance matrix of the coefficients, the logLik method returns the log-likelihood value, and the summary method prepares summary results and returns an object of class "summary.censReg". Furthermore, there are two methods for objects of class "summary.censReg": the print method prints summarized estimation results and the coef method returns a matrix consisting of the estimated coefficients, their standard errors, z statistics, and P values.

The print, coef, and vcov methods have an argument logSigma. This argument must be logical and determines whether the estimated standard deviation of the residuals (σ) should be printed or returned in logarithmic terms (if argument logSigma is TRUE) or in natural units (if it is FALSE). As the logarithm of the residuals' standard deviation ($\log \sigma$) is used during the estimation procedure, argument logSigma defaults to TRUE. If this argument is FALSE, the (unlogarithmized) standard deviation (σ) is calculated by applying the exponential function and the covariance matrix of all parameters is calculated by the Delta method.

3. Censored regression model for panel data

3.1. Specification

The censored regression model for panel data with individual specific effects has following specification:

$$y_{it}^* = x_{it}'\beta + \varepsilon_{it} = x_{it}'\beta + \mu_i + \nu_{it} \tag{8}$$

$$y_{it} = \begin{cases} a & \text{if } y_{it}^* \leq a \\ y_{it}^* & \text{if } a < y_{it}^* < b \\ b & \text{if } y_{it}^* \geq b \end{cases} \tag{9}$$

Here the subscript $i = 1, \ldots, N$ indicates the individual, subscript $t = 1, \ldots, T_i$ indicates the time period, T_i is the number of time periods observed for the ith individual, μ_i is a time-invariant individual specific effect, and ν_{it} is the remaining disturbance.

3.2. Fixed effects

In contrast to linear panel data models, we cannot get rid of the individual effects by the

within transformation. Theoretically, the fixed-effects panel Tobit model is affected by the *incidental parameters problem* (Neyman and Scott 1948; Lancaster 2000), i.e. the estimated coefficients are inconsistent unless the number of time periods (T_i) approaches infinity for each individual i. However, Greene (2004) showed with a Monte Carlo study that the slope parameters (but not the variance) of a fixed-effects Tobit model can be estimated consistently even if the number of time periods in small.

Assuming that the disturbance term ν follows a normal distribution with mean 0 and variance σ^2, the log-likelihood function is

$$\log L = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[I_{it}^a \log \Phi \left(\frac{a - x_{it}'\beta - \mu_i}{\sigma} \right) + I_{it}^b \log \Phi \left(\frac{x_{it}'\beta + \mu_i - b}{\sigma} \right) + \left(1 - I_{it}^a - I_{it}^b \right) \left(\log \phi \left(\frac{y_{it} - x_{it}'\beta - \mu_i}{\sigma} \right) - \log \sigma \right) \right]. \tag{10}$$

This log-likelihood function can be maximized with respect to the parameter vectors β and $\mu = [\mu_i]$. However, the number of coefficients increases linearly with the number of individuals. As most empirical applications use data sets with large numbers of individuals, special maximization routines, e.g. as described in Greene (2001) and Greene (2008, p. 554-557), are usually required.

3.3. Random effects

If the individual specific effects μ_i are independent of the regressors x_{it}, the parameters can be consistently estimated with a random effects model. Assuming that the individual specific effects μ follow a normal distribution with mean 0 and variance σ_μ^2, the remaining disturbance ν follows a normal distribution with mean 0 and variance σ_ν^2, and μ and ν are independent, the likelihood contribution of a single individual i is

$$L_i = \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \left[\Phi \left(\frac{a - x_{it}'\beta - \mu_i}{\sigma_\nu} \right) \right]^{I_{it}^a} \left[\Phi \left(\frac{x_{it}'\beta + \mu_i - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi \left(\frac{y_{it} - x_{it}'\beta - \mu_i}{\sigma_\nu} \right) \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \right\} \phi \left(\frac{\mu_i}{\sigma_\mu} \right) \, d\mu_i \tag{11}$$

and the log-likelihood function is

$$\log L = \sum_{i=1}^{N} \log L_i \tag{12}$$

(see Bruno 2004, p. 2).

Given that we assumed that μ follows a normal distribution, we can calculate the integrals in the log-likelihood function by the Gauss-Hermite quadrature and then maximise the log-likelihood function using standard non-linear optimisation algorithms (see Butler and Moffit 1982).

Alternatively, the log-likelihood function can be maximized using the method of Maximum Simulated Likelihood (MSL), which allows some flexibility in the specification of the disturbance terms (Greene 2008, p. 799).

Random effects estimation using the Gauss-Hermite quadrature

The Gauss-Hermite quadrature is a technique for approximating specific integrals with a weighted sum of function values at some specified points. Applying the Gauss-Hermite quadrature to equation (11), we get

$$L_i = \frac{1}{\sqrt{\pi}} \sum_{h=1}^{H} w_h \left\{ \prod_{t=1}^{T_i} \left[\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{I_{it}^a} \left[\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{(1 - I_{it}^a - I_{it}^b)} \right\}, \tag{13}$$

where H is number of quadrature points, $\psi_1, \ldots, \psi_H$ are the abscissae, and $w_1, \ldots, w_H$ are the corresponding weights (Greene 2008, p. 553)

3.4. Implementation in function `censReg`

Currently, only the random-effects model using the Gauss-Hermite quadrature is implemented in **censReg**.[2] Most steps internally done by `censReg` in case of panel data are similar to the steps described in section 2.4:

1. perform basic checks on the arguments provided by the user
2. prepare the data for the estimation, i.e. the matrix of the dependent variable $y = [y_{it}]$ and the 3-dimensional array the regressors $X = [x_{itk}]$
3. obtain initial values of the parameters β, σ_μ and σ_ν from a linear random-effects estimation using function `plm` of the **plm** package (Croissant and Millo 2008) (if no initial values are provided by the user)
4. obtain the abscissae ψ and the corresponding weights w for the Gauss-Hermite quadrature using function `ghq` of the **glmmML** package (Broström 2009)
5. define a function that calculates (using the Gauss-Hermite quadrature) and returns the log-likelihood value and its gradients[3] given the vector of parameters $(\beta', \sigma_\mu, \sigma_\nu)'$
6. call function `maxLik` of the **maxLik** package (Toomet and Henningsen 2010) for the maximisation of the likelihood function
7. add class `"censReg"` to the returned object

3.5. Using function `censReg`

Function `censReg` automatically estimates a random effects censored regression model if argument `data` is of class `"pdata.frame"`, i.e. created with function `pdata.frame` of package **plm** (Croissant and Millo 2008).

First, we prepare a small artificial panel data set with 15 individuals, 4 time periods and a censored dependent variable:

[2] Of course, also the "pooled" model can be estimated by `censReg`—simply by ignoring the panel structure of the data set.

[3] The gradients of the log-likelihood function are presented in appendix A.2.

```
R> set.seed( 123 )
R> pData <- data.frame(
+    id = rep( paste( "F", 1:15, sep = "_" ), each = 4 ),
+    time = rep( 1981:1984, 15 ) )
R> pData$mu <- rep( rnorm( 15 ), each = 4 )
R> pData$x1 <- rnorm( 60 )
R> pData$x2 <- runif( 60 )
R> pData$ys <- -1 + pData$mu + 2 * pData$x1 + 3 * pData$x2 + rnorm( 60 )
R> pData$y <- ifelse( pData$ys > 0, pData$ys, 0 )
R> library( plm )
R> pData <- pdata.frame( pData, c( "id", "time" ) )
```

Now, we estimate the random-effects censored regression model using the BHHH method (Berndt, Hall, Hall, and Hausman 1974) and show summary results:

```
R> system.time( panelResult <- censReg( y ~ x1 + x2, data = pData, method = "BHHH" ) )

   user  system elapsed
  0.027   0.000   0.028

R> summary( panelResult )

Call:
censReg(formula = y ~ x1 + x2, data = pData, method = "BHHH")

Observations:
         Total  Left-censored     Uncensored Right-censored
            60             20             40              0

Coefficients:
            Estimate Std. error t value  Pr(> t)
(Intercept) -0.36541    0.55503  -0.658  0.51032
x1           1.67990    0.29376   5.719 1.07e-08 ***
x2           2.24030    0.72945   3.071  0.00213 **
logSigmaMu  -0.12952    0.29504  -0.439  0.66065
logSigmaNu  -0.01243    0.14013  -0.089  0.92931
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

BHHH maximisation, 15 iterations
Return code 8: successive function values within relative tolerance limit (reltol)
Log-likelihood: -73.19882 on 5 Df
```

Argument `nGHQ` can be used to specify the number of points for the Gauss-Hermite quadrature. It defaults to 8. Increasing the number of points increases the accuracy of the computation of the log-likelihood value but also increases the computation time. In the following, we re-estimate the model above with different numbers of points and compare the execution times (measured in seconds of user CPU time) and the estimated parameters.

```
R> nGHQ <- 2^(2:6)
R> times <- numeric( length( nGHQ ) )
R> results <- list()
R> for( i in 1:length (nGHQ ) ) {
+    times[i] <- system.time( results[[i]] <- censReg( y ~ x1 + x2, data = pData,
+    method = "BHHH", nGHQ = nGHQ[i] ) )[1]
+ }
R> names(results)<-nGHQ
R> round( rbind(sapply( results, coef ),times),4)
```

```
                  4       8      16      32      64
(Intercept) -0.2858 -0.3654 -0.3656 -0.3652 -0.3652
x1           1.6679  1.6799  1.6838  1.6836  1.6836
x2           2.1621  2.2403  2.2634  2.2633  2.2633
logSigmaMu  -0.2716 -0.1295 -0.1129 -0.1140 -0.1140
logSigmaNu   0.0215 -0.0124 -0.0139 -0.0135 -0.0135
times        0.0160  0.0230  0.0340  0.0540  0.0970
```

4. Marginal Effects

The marginal effects of an explanatory variable on the expected value of the dependent variable is (Greene 2012, p. 889):

$$ME_j = \frac{\partial E[y|x]}{\partial x_j} = \beta_j \left[\Phi\left(\frac{b - x'\beta}{\sigma}\right) - \Phi\left(\frac{a - x'\beta}{\sigma}\right) \right] \tag{14}$$

In order to compute the approximate variance covariance matrix of these marginal effects using the Delta method, we need to obtain the Jacobian matrix of these marginal effects with respect to all estimated parameters (including σ):

$$\frac{\partial ME_j}{\partial \beta_k} = \Delta_{jk} \left[\Phi\left(\frac{b - x'\beta}{\sigma}\right) - \Phi\left(\frac{a - x'\beta}{\sigma}\right) \right] - \frac{\beta_j \, x_k}{\sigma} \left[\phi\left(\frac{b - x'\beta}{\sigma}\right) - \phi\left(\frac{a - x'\beta}{\sigma}\right) \right] \tag{15}$$

and

$$\frac{\partial ME_j}{\partial \sigma} = -\beta_j \left[\phi\left(\frac{b - x'\beta}{\sigma}\right) \frac{b - x'\beta}{\sigma^2} - \phi\left(\frac{a - x'\beta}{\sigma}\right) \frac{a - x'\beta}{\sigma^2} \right], \tag{16}$$

where Δ_{jk} is "Kronecker's Delta" with $\Delta_{jk} = 1$ for $j = k$ and $\Delta_{jk} = 0$ for $j \neq k$. If the upper limit of the censored dependent variable (b) is infinity or the lower limit of the censored dependent variable (a) is minus infinity, the terms in the square brackets in equation (16) that include b or a, respectively, have to be removed.

Appendix

A. Gradients of the log-likelihood function

A.1. Cross-sectional data

$$\frac{\partial \log L}{\partial \beta_j} = \sum_{i=1}^{N} \Bigg[- I_i^a \frac{\phi\left(\frac{a-x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{a-x_i'\beta}{\sigma}\right)} \frac{x_{ij}}{\sigma} + I_i^b \frac{\phi\left(\frac{x_i'\beta-b}{\sigma}\right)}{\Phi\left(\frac{x_i'\beta-b}{\sigma}\right)} \frac{x_{ij}}{\sigma} + \left(1 - I_i^a - I_i^b\right) \frac{y_i - x_i'\beta}{\sigma} \frac{x_{ij}}{\sigma} \Bigg] \tag{17}$$

$$\frac{\partial \log L}{\partial \log \sigma} = \sum_{i=1}^{N} \Bigg[- I_i^a \frac{\phi\left(\frac{a-x_i'\beta}{\sigma}\right)}{\Phi\left(\frac{a-x_i'\beta}{\sigma}\right)} \frac{a - x_i'\beta}{\sigma} - I_i^b \frac{\phi\left(\frac{x_i'\beta-b}{\sigma}\right)}{\Phi\left(\frac{x_i'\beta-b}{\sigma}\right)} \frac{x_i'\beta - b}{\sigma} + \left(1 - I_i^a - I_i^b\right) \left(\left(\frac{y_i - x_i'\beta}{\sigma} \right)^2 - 1 \right) \Bigg] \tag{18}$$

A.2. Panel data with random effects

$$\begin{aligned} \frac{\partial \log L_i}{\partial \beta_j} = & \frac{1}{\sqrt{\pi}\, L_i} \sum_{h=1}^{H} w_h \Bigg\{ \left(\prod_{t=1}^{T_i} \left[\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{I_{it}^a} \right. \\ & \left[\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \Bigg) \\ & \left(\sum_{t=1}^{T_i} \left[- \frac{\phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{x_{jit}}{\sigma_\nu} \right]^{I_{it}^a} \right. \\ & \left. \left[\frac{\phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)}{\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)} \frac{x_{jit}}{\sigma_\nu} \right]^{I_{it}^b} \left[- \frac{\phi'\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{x_{jit}}{\sigma_\nu^2} \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \right) \Bigg\} \end{aligned} \tag{19}$$

$$\frac{\partial \log L_i}{\partial \log \sigma_\mu} = \frac{\sigma_\mu}{\sqrt{\pi}\, L_i} \sum_{h=1}^{H} w_h \left\{ \left(\prod_{t=1}^{T_i} \left[\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{I_{it}^a} \right. \right. \tag{20}$$
$$\left[\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \Bigg)$$
$$\left(\sum_{t=1}^{T_i} \left[- \frac{\phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{\sqrt{2}\psi_h}{\sigma_\nu} \right]^{I_{it}^a} \right.$$
$$\left. \left[\frac{\phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)}{\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)} \frac{\sqrt{2}\psi_h}{\sigma_\nu} \right]^{I_{it}^b} \left[- \frac{\phi'\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{\sqrt{2}\psi_h}{\sigma_\nu^2} \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \right) \Bigg\}$$

$$\frac{\partial \log L_i}{\partial \log \sigma_\nu} = \frac{\sigma_\nu}{\sqrt{\pi}\, L_i} \sum_{h=1}^{H} w_h \left\{ \left(\prod_{t=1}^{T_i} \left[\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{I_{it}^a} \right. \right. \tag{21}$$
$$\left[\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right) \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \Bigg)$$
$$\left(\sum_{t=1}^{T_i} \left[- \frac{\phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\Phi\left(\frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{a - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu^2} \right]^{I_{it}^a} \right.$$
$$\left[- \frac{\phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)}{\Phi\left(\frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu} \right)} \frac{x_{it}'\beta + \sqrt{2}\sigma_\mu \psi_h - b}{\sigma_\nu^2} \right]^{I_{it}^b}$$
$$\left. \left[- \frac{1}{\sigma_\nu} - \frac{\phi'\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)}{\phi\left(\frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu} \right)} \frac{y_{it} - x_{it}'\beta - \sqrt{2}\sigma_\mu \psi_h}{\sigma_\nu^2} \right]^{\left(1 - I_{it}^a - I_{it}^b\right)} \right) \Bigg\}$$

References

Berndt EK, Hall BH, Hall RE, Hausman JA (1974). "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement*, **3**(4), 653–665.

Broström G (2009). *glmmML: Generalized Linear Models with Clustering*. R package version 0.81, URL https://CRAN.R-project.org/package=glmmML.

Bruno G (2004). "Limited Dependent Panel Data Models: A Comparative Analysis of Classical and Bayesian Inference among Econometric Packages." *Computing in Economics and Finance 41*, Society for Computational Economics. URL https://ideas.repec.org/p/sce/scecf4/41.html.

Butler J, Moffit R (1982). "A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model." *Econometrica*, **50**, 761–764.

Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The plm Package." *Journal of Statistical Software*, **27**(2), 1–43.

Greene W (2004). "Fixed Effects and Bias Due to the Incidental Parameters Problem in the Tobit Model." *Econometric Reviews*, **23**(2), 125–147.

Greene WH (2001). "Fixed and Random Effects in Nonlinear Models." *Working Paper EC-01-01*, Department of Economics, Stern School of Business, New York University.

Greene WH (2008). *Econometric Analysis*. 6th edition. Prentice Hall.

Greene WH (2012). *Econometric Analysis*. 7th edition. Pearson.

Henningsen A (2011). *censReg: Censored Regression (Tobit) Models*. R package version 0.5, https://CRAN.R-project.org/package=censReg.

Kleiber C, Zeileis A (2008). *Applied Econometrics with R*. Springer, New York.

Kleiber C, Zeileis A (2009). *AER: Applied Econometrics with R*. R package version 1.1, https://CRAN.R-project.org/package=AER.

Lancaster T (2000). "The incidential parameter problem since 1948." *Journal of Econometrics*, **95**, 391–413.

Neyman J, Scott E (1948). "Consistent Estimates Based on Partially Consistent Observations." *Econometrica*, **16**, 1–32.

Tobin J (1958). "Estimation of relationship for limited dependent variables." *Econometrica*, **26**, 24–36.

Toomet O, Henningsen A (2010). *maxLik: Tools for Maximum Likelihood Estimation*. R package version 0.7, https://CRAN.R-project.org/package=maxLik.

Affiliation:

Arne Henningsen
Department of Food and Resource Economics
University of Copenhagen
1958 Frederiksberg, Denmark
E-mail: arne.henningsen@gmail.com
URL: https://www.arne-henningsen.name/