LEAST SQUARES ESTIMATES SYMETRICALLY ADJUSTED WITH NO LOSS OF INFORMATION IN THE TRUNCATED TOBIT MODEL

Ezequiel Uriel, Mª Consuelo Colom, Mª Cruz Molés y Mª Luisa Moltó

WP-EC 98-25

Correspondence to E. Uriel: I V I E
C/. Guardia Civil, 22, Esc. 2, 1º. 46020 Valencia.
Tel: 963 930 816 / Fax: 963 930 856 / e-mail: ezequiel.uriel@ivie.es

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Octubre 1998.
ISBN: 84-482-1936-8
Depósito Legal: V-4088-1998

---

* E. Uriel: Universitat de València and Instituto Valenciano de Investigaciones Económicas; MªC. Colom, MªC. Molés y MªL. Moltó: Universitat de València.

LEAST SQUARES ESTIMATES SYMETRICALLY ADJUSTED WITH NO LOSS OF INFORMATION IN THE TRUNCATED TOBIT MODEL

Ezequiel Uriel, Mª Consuelo Colom, Mª Cruz Molés y Mª Luisa Moltó

A B S T R A C T

The objective of this paper is to improve the estimator of Powell (1986) in the truncated Tobit model. The Powell estimator is the least squares alternative to the maximum likelihood estimator for the Tobit model. Only symmetry of the distribution of the error term is assumed, but no distribution function is needed. In order to attain symmetry, our proposal predicts the values of the left hand side of the distribution, using the information contained in the right hand side, instead of eliminating sample information as the Powell estimator does. The paper takes an appropriate family of estimators as the point of departure, deriving the Power estimator as a particular case. The behaviour of the estimates of interest within the family is analysed both in terms of the theoretical properties and the small sample properties.

Keywords: Semiparametric estimation, Tobit model, truncated regression, Powell estimator.

RESUMEN

El objetivo de este trabajo es mejorar el estimador propuesto por Powell (1986) para el modelo Tobit truncado. El estimador de Powell supone una alternativa al estimador máximo verosímil del modelo Tobit y su ventaja es que no necesita admitir ninguna forma funcional conocida para la distribución del término de error. La única hipótesis distribucional que se impone es la simetría. Para conseguir la simetría, nuestra propuesta predice los valores de la cola inferior de la distribución, utilizando la información contenida en la cola derecha, en vez de eliminar información muestral como hace el estimador de Powell.

El trabajo toma un adecuada familia de estimadores como punto de partida, derivando el estimador de Powell como un caso particular. Se analiza el comportamiento de los estimadores de interés dentro de la familia tanto en términos de propiedades teóricas como propiedades en muestras finitas.

Palabras clave: Estimación semiparamétrica, modelo Tobit, regresión truncada, estimador de Powell.

1. INTRODUCTION

Limited dependent variable models are of much use in empirical economic applications. Estimation procedures based on maximum likelihood provide consistent and asymptotically normal estimators when a parametric probability distribution is assumed for the error term, which is known. However, these estimates are inconsistent when there is a wrong specification of the distribution or when the error term is heteroscedastic.

Powell (1986) comments some suggestions appeared in the econometric literature, to find out estimators more robust than the maximum likelihood estimator. He proposes a new estimator that significantly improves the existing ones. The drawback of his proposal, which seems adequate for censored and truncated regression models, is the loss of information involved.

The Powell estimator is the least squares alternative to the maximum likelihood estimator for the Tobit model. Only symmetry of the distribution of the error term is assumed, but no distribution function is needed.

In order to attain symmetry for the distribution of the error term, Powell truncates the right hand side of the distribution of the dependent variable, thus eliminating same sample information.

The objective of this paper is to improve the estimator of Powell (1986) in the truncated Tobit model. In order to attain symmetry, our proposal predicts the values of the left hand side of the distribution, using the information contained in the right hand side, instead of eliminating sample information.

The paper takes an appropriate family of estimators as the point of departure, deriving the Power estimator as a particular case. The behaviour of the estimates of interest within the family is analysed both in terms of the theoretical properties and the small sample properties.

The paper is structured in the following manner; in the second section the new estimator for the Tobit model, denominated modified Powell estimator, which improves the Powell estimator is developed and the behaviour of the estimator is studied. Finally, the properties in small samples are explored in section 3, through simulation.

2. MODIFIED POWELL ESTIMATOR

Consider the specification of the truncated Tobit model

$$y_i = \begin{cases} y_i^* & sii \quad y_i^* > 0 \\ 0 & sii \quad y_i^* \leq 0 \end{cases}$$

where $y_i^* = x_i'\beta_0 + u_i$, x_i is a vector of explanatory variables, β_0 is the vector of unkown parameters and u_i is the error term.

Note that in our model the variable y_i^* and the vector of explanatory variables x_i are observable only when $y_i^* > 0$, no information being available when $y_i^* \leq 0$.

Ordinary least squares provides consistent estimates of the parameter vector β_0 when the error term u_i is symmetrically distributed around zero and the dependent variable y_i^* is fully observed. However, the truncation in the dependent variable introduces asymmetry in the distribution and consequently the least squares procedure will not provide consistent estimators.

Powell suggestion lies in restoring symmetry by truncating the right hand side of the distribution of the dependent variable. Thus, the sample observations such that $y_i^* \geq 2x_i'\beta_0$ are eliminated; given that there is no information when $y_i^* \leq 0$, the interval for the sample observations will be $(0, 2x_i'\beta_0)$, as shown in figure 1, which represents the density function of the model $y_i^* = x_i'\beta_0 + u_i$.

When only the observations of the interval $(0, 2x_i'\beta_0)$ are used, the distribution is symmetric and the least squares estimator for the truncated Tobit model called symmetrically truncated least squares (STLS) is consistent and asymptotically Normal (Powell, 1986). From now on, we will denote this estimator by P.

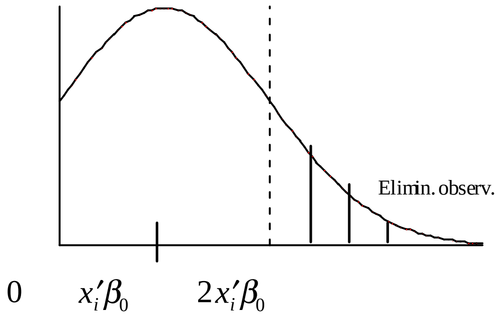

Figure 1 : Probability density function of the truncated response variable

In order to avoid the information loss and improve the Powell estimator, symmetry of the distribution is restored by including the predicted observations of the left hand side of the distribution of the dependent variable.

The procedure consists in generating sample observations such that $y_i^* \leq 0$. Predicted values must be symmetric to those verifying $y_i^* \geq 2x_i'\beta_0$. For an individual which is situated on the right-hand side of the distribution, whose dependent variable value verifies $y_i^* \geq 2x_i'\beta_0$, a prediction of the dependent variable on the left hand side of the distribution is generated, such that $\tilde{y}_i = 2\tilde{x}_i'\beta_0 - y_i$, where the vector of characteristics used for prediction is $\tilde{x}_i = x_i$.

In this way the prediction of the response variable $\tilde{y}_i$ is associated to an individual with predicted vector of characteristics $\tilde{x}_i$, as the original response is unknown, given that negative values are unobserved. This is the way to attain symmetry in the distribution.

Figure 2 shows the newly generated observation $\tilde{y}_i$ in the distribution.

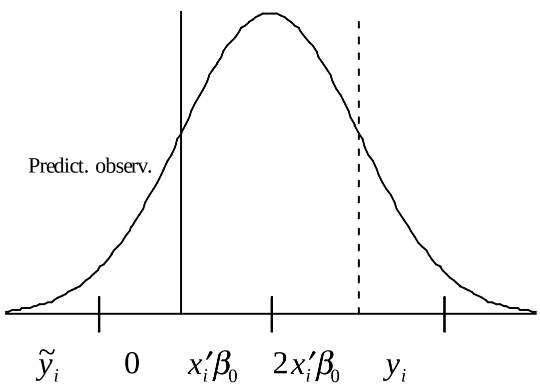

Figure 2 : Probability density function of the non-truncated response variable

Objective function

If we consider both sample observations and predicted observations, the objective function is the following[1] :

$$(1)\quad R_N(\beta) = \sum_{i=1}^{N} 1(0 < y_i < 2x_i'\beta)(y_i - x_i'\beta)^2 + \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta; \tilde{y} > 0)(y_i - x_i'\beta)^2$$
$$+ w_1 \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta; \tilde{y}_i \leq 0)(y_i - x_i'\beta)^2 + w_2 \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta;\ \tilde{y}_i \leq 0)(\tilde{y}_i - \tilde{x}_i'\beta)^2$$

Note that those sample individuals whose response is in the right hand side of the distribution appear twice in the objective function with weights w_1 corresponding to the original data of the right hand tail, and w_2 corresponding to the predicted data in the left-hand tail. In order to avoid that these sample individuals have a higher weight than the other sample individuals, $w_1 + w_2 = 1$.

When $2x_i'\beta \leq y_i \leq 2x_i'\beta_0$ the copy $\tilde{y}_i$ has a positive sign. Consequently, the predicted observations will not be taken into account in the sample, including only the observations of the right-hand tail.

Least-squares estimator symmetrically adjusted without loss of information.

A estimator for β_0 would be drawn from the following set of normal equations:

$$(2)\quad \sum_{i=1}^{N} 1(0 < y_i < 2x_i'\beta)(y_i - x_i'\beta)(-2)x_i + \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta; \tilde{y}_i > 0)(y_i - x_i'\beta)(-2)x_i$$
$$+ w_1 \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta\ \ \tilde{y}_i \leq 0)(y_i - x_i'\beta)(-2)x_i + w_2 \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta,\ \tilde{y}_i \leq 0)(\tilde{y}_i - x_i'\beta)(-2)x_i = 0$$

The following estimator is derived from the normal equations:

[1] The parameter used in the indicator function of the objective function, should be the true vector of parameters β_0 ; but given that it is the objective of our estimation, it is unknown and, thus following Powell (1986), the generic vector β is used.

$$(3)\ \hat{\beta}_{MP} = \Bigg[\sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(x_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(x_i x_i'\right)$$

$$+ \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(x_i x_i'\right)\Bigg]^{-1}$$

$$\Bigg[\sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(y_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(y_i x_i'\right)$$

$$+ w_1 \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(y_i x_i'\right) + w_2 \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(\tilde{y}_i x_i'\right)\Bigg]$$

When $w_1 = w_2 = \frac{1}{2}$, the modified Powell estimator is consistent and asymptotically normal. This is shown in the Appendix.

Given that the true value of the parameter vector β_0 is unknown, the predictions of the lower tail are generated such that $\tilde{y}_i = 2x_i'\beta^* - y_i$; where β^* is an estimate of the parameter vector. Consequently, the so called feasible modified Powell estimator (MPF) has the following expression:

$$(4)\ \hat{\beta}_{MPF} = \Bigg[\sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(x_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(x_i x_i'\right)$$

$$+ \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(x_i x_i'\right)\Bigg]^{-1}$$

$$\Bigg[\sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(y_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(y_i x_i'\right)$$

$$+ \left(w_1 - w_2\right) \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(y_i x_i'\right) + 2w_2 \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(x_i x_i'\right)\beta^*\Bigg]$$

We can rewrite the new estimator (MPF) as a function of the Powell estimator $\hat{\beta}_P$ (Powell, 1996):

$$(5)\ \hat{\beta}_{MPF} = \Bigg[\sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(x_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(x_i x_i'\right)$$

$$+ \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(x_i x_i'\right)\Bigg]^{-1}$$

$$\Bigg[\hat{\beta}_P \sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)\left(x_i x_i'\right) + \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)\left(y_i x_i'\right)$$

$$+ \left(w_1 - w_2\right) \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(y_i x_i'\right) + \beta^*\, 2w_2 \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)\left(x_i x_i'\right)\Bigg]$$

Let us define

$$a = \sum_{i=1}^{N} 1\left(0 < y_i < 2x_i'\hat{\beta}\right)x_i x_i'$$

$$b = \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)x_i x_i'$$

$$c = \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)x_i x_i'$$

substituting those definitions into equation (5),

$$(6)\quad \hat{\beta}_{MPF} = \hat{\beta}_P \frac{a}{a+b+c} + \beta^* \frac{2w_2 c}{a+b+c} + \frac{\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)y_i x_i}{a+b+c}$$

$$+ \frac{\left(w_1 - w_2\right)\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right)y_i x_i}{a+b+c}$$

Next, the estimator $\hat{\beta}_{MPF}$ is analysed in detail, according to different values of w_1 and w_2, and also to different alternatives for the estimator β^*, that generates the copies.

When the weights are symmetric $w_1 = w_2 = \frac{1}{2}$, then given that the weight of the observations in both tails are the same, the symmetric feasible modified Powell estimator is

$$(7)\quad \hat{\beta}_{MPF_S} = \hat{\beta}_P \frac{a}{a+b+c} + \beta^* \frac{c}{a+b+c} + \frac{\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right)y_i x_i}{a+b+c}$$

When Montecarlo simulation experiments are made, we observe that the number of observations generating positive copies $\tilde{y}_i > 0$ diminishes systematically throughout the iterative process.

As a result in the last iteration, both the term b and the term $\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta};\ \tilde{y}_i > 0\right)y_i x_i$, are equal to zero, and consequently, the symmetric feasible modified Powell estimator is the following convex linear combination:

(8) $\hat{\beta}_{MPF_s} = \hat{\beta}_P \frac{a}{a+c} + \beta^* \frac{c}{a+c}$

When a consistent and asymptotically normal estimator β^* is used, the estimator $\hat{\beta}_{MPF_S}$ will maintain all the properties.

In particular, if $\beta^* = \hat{\beta}_p$, the symmetric feasible modified Powell estimator will coincide with the Powell estimator. We will denote it by $\hat{\beta}_{MPF_{SP}}$.

In practice, the Powell estimator is a particular case of the modified Powell estimator, $\hat{\beta}_{MPF_{SP}} = \hat{\beta}_P$, keeping its properties, but also its drawbacks. A loss of efficiency can appear, given that the estimator $\hat{\beta}_P$ only takes into account those observations appearing in the centre of the distribution, rejecting the information contained in the observations of the upper tail.

When $w_1 > w_2$, the asymmetric feasible modified Powell estimator will have the following expression:

(9) $\hat{\beta}_{MPF_A} = \hat{\beta}_P \frac{a}{a+c} + \frac{(w_1 - w_2)\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right) y_i x_i}{a+c} + \beta^* \frac{2w_2 c}{a+c}$

taking into account that both the terms b and $\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i > 0\right) y_i x_i$ tend to zero.

The case $w_1 < w_2$ is not considered here, given that it implies that a higher weight is given to predictions $\tilde{y}_i$ than to observations in the objective function.

When $\beta^* = \hat{\beta}_P$, the asymmetric feasible modified Powell estimator is:

(10) $\hat{\beta}_{MPF_{AP}} = \hat{\beta}_P \frac{a + 2w_2 c}{a+c} + \frac{w_1 - w_2}{a+c} \sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta}; \tilde{y}_i \leq 0\right) y_i x_i$

This estimator is not consistent, as it used to be when $w_1 = w_2$.

The mean square error (MSE) allows joint consideration of consistency and efficiency. We will analyse the behaviour of the last estimator through the MSE.

Assuming that the Powell estimator is asymptotically consistent we obtain:

$$V\left[\hat{\beta}_{MPF_{AP}}\right]=\left(\frac{a+2w_2c}{a+c}\right)^2 V\left[\hat{\beta}_P\right]+\left(\frac{w_1-w_2}{a+c}\right)^2 c\, V\left[u / y_i \geq 2x_i'\hat{\beta};\tilde{y}_i \leq 0\right]$$

$$\mathrm{E}\left[\hat{\beta}_{MPF_{AP}}\right]=\beta_0+\left(\frac{w_1-w_2}{a+c}\right)\sum_{i=1}^{N} 1\left(y_i \geq 2x_i'\hat{\beta};\tilde{y}_i \leq 0\right)x_i \mathrm{E}\left[u / y_i \geq 2x_i'\hat{\beta};\tilde{y}_i \leq 0\right]$$

Consequently the MSE for large samples, is given by:

$$\mathrm{E}\left[\left(\hat{\beta}_{MPF_{AP}}-\beta_0\right)^2\right]=\left(\frac{a+2w_2c}{a+c}\right)^2 V\left[\hat{\beta}_P\right]+\left(\frac{w_1-w_2}{a+c}\right)^2 c\, \mathrm{E}\left[u^2 / y_i \geq 2x_i'\hat{\beta};\tilde{y}_i \leq 0\right]$$

Note that in the right hand side of the above expression the first term is lower than the variance of the Powell estimator, but the appearance of the second term do not allow us to draw definite conclusions on the mean square error.

Given that a direct comparison of the modified Powell estimator MSE and the Powell estimator MSE is not possible, we will next analyse the behaviour of the mean square error through simulation. We will consider different values for the weights w_1 and w_2, for the estimator $\hat{\beta}_{MPF_{AP}}$

3. SMALL SAMPLE PROPERTIES

The behaviour of the estimator in small samples is interesting to explore, given that they are often used in practical problems. Montecarlo simulations provide interesting comparisons between our estimator, the modified Powell (MP) estimator, and other two estimators for the truncated Tobit model, the maximum-likelihood (ML) estimator and the Powell (P) estimator.

Different models for data generation are considered here by changing the distributional assumptions about the error term, in particular the specification of the distribution and the homoscedasticity assumption. The objective is to compare the functioning of the three estimators in face of a wrong specification of the distribution, that can produce non robust estimators. We analyse which estimator shows the best behaviour. Efficiency is evaluated through the mean square error and the mean absolute error criteria.

Table 1 shows the results of different models, under homoscedastic error terms and Normal distribution. Design 1 considers a model with slope 1 and zero constant term; that is, $y = x + u$. The error terms follow a typified Normal distribution and for the explanatory variables x a Uniform distribution (-1.7, 1.7) is assumed. The sample size is 200. The only change introduced in Design 2 is the sample size, which is in this case 100.

Design 3 duplicates the dispersion of the error terms of design 1, being now the standard deviation equal to 2. Design 4 maintains the same hypothesis of design 1, except that the model is now specified with a unit constant and slope term $(y = 1 + x + u)$. Design 5 considers the same model as design 4, but with a standard deviation of 2 for the error terms.

Design 6 and 7 include in the model a new regressor taking alternate values -1 and 1, with zero mean and unit variance. The constant term is now zero, being the first slope one and the second slope zero. The only difference between design 6 and 7 is the sample size (200 for design 6 and 300 for design 7).

Finally, design 8 only differs from design 6 in the dispersion of the error terms. Design 9 includes in the specification of the model of design 6 a unit constant term.

Results of design 1 show that the ML estimator is more efficient than the P and the MP estimators. Also, the MP estimator, irrespective of the weights, is more efficient than the P estimator.

We observe that the higher the difference between w_1 and w_2 , the lower the mean square error is, while the bias increases.

When we compare design 1 and 2, where the only difference is the sample size, the MSE (mean square error) and MAE (mean absolute error) increases when the sample size decreases. Similarly, when the standard deviation of the error term is doubled (design 3), both the MSE and the bias increase for all the estimators, in comparison with design 1.

Table 1 : Simulation results with homoscedastic normal error terms

Design 1 -Standard deviation=1, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0187	0.3770	0.1418	-0.2215	0.0372	0.2420	0.2881
Slope	1.0000	1.0113	0.2558	0.0652	0.8399	0.9707	1.1630	0.2009
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0907	0.8166	0.6716	-0.0282	0.3302	0.5438	0.5397
Slope	1.0000	0.9267	0.6224	0.3909	0.5714	0.7721	1.1060	0.4258
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.1062	0.8007	0.6492	-0.0098	0.3404	0.5508	0.5377
Slope	1.0000	0.9202	0.6119	0.3789	0.5668	0.7663	1.0950	0.4230
MP(w_1=0.55)								
Const.	0.0000	0.1728	0.7388	0.5730	0.0662	0.3988	0.5826	0.5363
Slope	1.0000	0.8841	0.5652	0.3313	0.5562	0.7481	1.0460	0.4098
MP(w_1=0.6)								
Const.	0.0000	0.2611	0.6610	0.5029	0.1686	0.4603	0.6283	0.5422
Slope	1.0000	0.8369	0.5081	0.2835	0.5473	0.7153	0.9853	0.3983
MP(w_1=0.7)								
Const.	0.0000	0.4400	0.5048	0.4471	0.3637	0.5990	0.7117	0.5837
Slope	1.0000	0.7363	0.3921	0.2225	0.5038	0.6310	0.8608	0.3939

Design 2 -Standard deviation=1, N=100, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0994	0.6633	0.4476	-0.3306	0.0503	0.3037	0.4504
Slope	1.0000	1.0628	0.4082	0.1697	0.7984	0.9963	1.2760	0.2982
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0592	1.3606	1.8450	-0.0281	0.4388	0.6318	0.6933
Slope	1.0000	0.9510	0.9787	0.9555	0.5160	0.7358	1.1005	0.5558
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0753	1.3337	1.7750	-0.0112	0.4437	0.6358	0.6881
Slope	1.0000	0.9432	0.9603	0.9208	0.5138	0.7305	1.0875	0.5507
MP(w_1=0.55)								
Const.	0.0000	0.1431	1.2274	1.5190	0.0625	0.4854	0.6521	0.6738
Slope	1.0000	0.9058	0.8842	0.7868	0.5025	0.7172	1.0510	0.5266
MP(w_1=0.6)								
Const.	0.0000	0.2303	1.0944	1.2450	0.1527	0.5142	0.6843	0.6621
Slope	1.0000	0.8604	0.7882	0.6376	0.4940	0.6886	0.9997	0.4982
MP(w_1=0.7)								
Const.	0.0000	0.4107	0.8289	0.8522	0.3518	0.6198	0.7623	0.6740
Slope	1.0000	0.7602	0.5996	0.4153	0.4685	0.6277	0.8737	0.4625

Table 1 (cont.)

Design 3 -Standard deviation=2, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.1847	1.1561	1.3640	-0.7059	0.1168	0.6250	0.8158
Slope	1.0000	1.1044	0.5118	0.2715	0.7400	0.9920	1.3940	0.3906
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.4769	1.3168	1.9530	0.3692	0.7825	1.0075	0.9337
Slope	1.0000	0.8011	0.9966	1.0280	0.2676	0.5629	1.0565	0.6907
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.5032	1.2910	1.9120	0.3962	0.8015	1.0360	0.9421
Slope	1.0000	0.7944	0.9785	0.9949	0.2700	0.5641	1.0435	0.6835
MP(w_1 =0.55)								
Const.	0.0000	0.6182	1.1920	1.7960	0.5147	0.8956	1.1120	0.9876
Slope	1.0000	0.7640	0.9010	0.8635	0.2876	0.5434	0.9927	0.6528
MP(w_1 =0.6)								
Const.	0.0000	0.7762	1.0689	1.7390	0.6845	1.0270	1.2285	1.0690
Slope	1.0000	0.7284	0.8052	0.7189	0.2869	0.5118	0.9261	0.6186
MP(w_1 =0.7)								
Const.	0.0000	1.0765	0.8104	1.8120	1.0745	1.2700	1.4060	1.2350
Slope	1.0000	0.6476	0.6125	0.4975	0.3243	0.4633	0.7946	0.5650

Design 4 -Standard deviation=1, N=200, Censoring=25%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	0.9658	0.1634	0.0277	0.8777	0.9947	1.0670	0.1258
Slope	1.0000	1.0337	0.1441	0.0218	0.9459	1.0260	1.1255	0.1148
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	0.9711	0.2956	0.0878	0.8772	1.0410	1.1505	0.1973
Slope	1.0000	1.0264	0.2580	0.0669	0.8703	0.9574	1.1250	0.1825
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	0.9799	0.2904	0.0843	0.8860	1.0490	1.1555	0.1951
Slope	1.0000	1.0198	0.2534	0.0643	0.8680	0.9547	1.1175	0.1804
MP(w_1 =0.55)								
Const.	1.0000	1.0194	0.2689	0.0723	0.9358	1.0825	1.1825	0.1895
Slope	1.0000	0.9883	0.2365	0.0558	0.8479	0.9221	1.0840	0.1757
MP(w_1 =0.6)								
Const.	1.0000	1.0733	0.2421	0.0637	1.0035	1.1215	1.2170	0.1920
Slope	1.0000	0.9462	0.2162	0.0494	0.8208	0.8927	1.0400	0.1753
MP(w_1 =0.7)								
Const.	1.0000	1.1807	0.1830	0.0660	1.1310	1.2180	1.2910	0.2301
Slope	1.0000	0.8581	0.1745	0.0504	0.7521	0.8232	0.9370	0.1942

Table 1 (cont.)

Design 5 - Standard deviation =2, N=200, Censoring =25%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	0.9850	0.4995	0.2485	0.6953	1.0590	1.3350	0.3939
Slope	1.0000	0.9843	0.2878	0.0827	0.8154	0.9609	1.1535	0.2248
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	1.1335	0.9359	0.8894	1.0650	1.3635	1.5745	0.5800
Slope	1.0000	0.8528	0.7617	0.5990	0.4817	0.6686	0.9623	0.5228
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	1.0000	1.1549	0.9169	0.8605	1.0860	1.3755	1.5850	0.5826
Slope	1.0000	0.8479	0.7495	0.5821	0.4784	0.6788	0.9550	0.5191
MP(w_1 =0.55)								
Const.	1.0000	1.2589	0.8477	0.7820	1.2075	1.4500	1.6635	0.6160
Slope	1.0000	0.8184	0.6929	0.5107	0.4890	0.6599	0.9169	0.5036
MP(w_1 =0.6)								
Const.	1.0000	1.3935	0.7571	0.7252	1.3540	1.5580	1.7400	0.6698
Slope	1.0000	0.7775	0.6228	0.4355	0.4675	0.6318	0.8669	0.4929
MP(w_1 =0.7)								
Const.	1.0000	1.6452	0.5665	0.7356	1.6320	1.7450	1.9110	0.7867
Slope	1.0000	0.6821	0.4783	0.3287	0.4402	0.5691	0.7259	0.4867

Design 6 - Standard deviation =1, N=200, Censoring =50%, Two slope coefficients

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0063	0.3855	0.1479	-0.2119	0.0503	0.2836	0.2838
Slope1	1.0000	0.9978	0.2561	0.0652	0.8276	0.9610	1.1480	0.2043
Slope2	0.0000	-0.0190	0.1437	0.0209	-0.1140	-0.0098	0.0935	0.1181
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.1056	0.7965	0.6424	-0.0358	0.3256	0.5234	0.5052
Slope1	1.0000	0.9102	0.5701	0.3314	0.5588	0.7747	1.0615	0.4024
Slope2	0.0000	-0.0025	0.1950	0.0378	-0.1273	0.0115	0.1081	0.1501
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.1211	0.7809	0.6215	-0.0167	0.3361	0.5308	0.5036
Slope1	1.0000	0.9030	0.5598	0.3212	0.5560	0.7715	1.0595	0.4002
Slope2	0.0000	-0.0037	0.1927	0.0369	-0.1259	0.0100	0.1064	0.1485
MP(w_1 =0.55)								
Const.	0.0000	0.1868	0.7203	0.5512	0.0549	0.3831	0.5677	0.5044
Slope1	1.0000	0.8674	0.5171	0.2836	0.5426	0.7414	1.0200	0.3906
Slope2	0.0000	-0.0051	0.1788	0.0318	-0.1239	0.0060	0.0962	0.1378
MP(w_1 =0.6)								
Const.	0.0000	0.2757	0.6451	0.4900	0.1691	0.4457	0.6176	0.5186
Slope1	1.0000	0.8226	0.4641	0.2458	0.5359	0.7133	0.9616	0.3823
Slope2	0.0000	-0.0056	0.1624	0.0263	-0.1176	0.0010	0.0959	0.1271
MP(w_1 =0.7)								
Const.	0.0000	0.4533	0.4917	0.4460	0.3723	0.5903	0.7070	0.5723
Slope1	1.0000	0.7224	0.3613	0.2070	0.4985	0.6340	0.8359	0.3846
Slope2	0.0000	-0.0056	0.1286	0.0165	-0.0969	-0.0076	0.0796	0.1021

Table 1 (cont.)

Design 7 - Standard deviation =1, N=300, Censoring =50%, Two slope coefficients

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0391	0.3244	0.1063	-0.2424	0.0053	0.2099	0.2529
Slope1	1.0000	1.0267	0.2150	0.0467	0.8586	0.9898	1.1680	0.1708
Slope2	0.0000	-0.0046	0.1220	0.0148	-0.0741	-0.0012	0.0682	0.0926
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0037	0.7978	0.6334	-0.2377	0.2146	0.4277	0.5028
Slope1	1.0000	0.9748	0.5662	0.3196	0.6356	0.8443	1.1840	0.3865
Slope2	0.0000	0.0244	0.2115	0.0451	-0.0888	0.0108	0.1206	0.1474
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0138	0.7824	0.6092	-0.2167	0.2298	0.4377	0.4979
Slope1	1.0000	0.9671	0.5562	0.3089	0.6311	0.8372	1.1750	0.3826
Slope2	0.0000	0.0240	0.2081	0.0437	-0.0897	0.0156	0.1187	0.1452
MP(w_1=0.55)								
Const.	0.0000	0.0899	0.7223	0.5272	-0.1250	0.2901	0.4822	0.4855
Slope1	1.0000	0.9274	0.5124	0.2665	0.6126	0.8083	1.1150	0.3656
Slope2	0.0000	0.0202	0.1930	0.0375	-0.0744	0.0130	0.1090	0.1353
MP(w_1=0.6)								
Const.	0.0000	0.1885	0.6471	0.4521	-0.0006	0.3667	0.5357	0.4839
Slope1	1.0000	0.8778	0.4581	0.2238	0.5951	0.7666	1.0425	0.3503
Slope2	0.0000	0.0183	0.1733	0.0302	-0.0747	0.0150	0.0983	0.1216
MP(w_1=0.7)								
Const.	0.0000	0.3912	0.4912	0.3931	0.2611	0.5219	0.6626	0.5262
Slope1	1.0000	0.7689	0.3567	0.1800	0.5412	0.6822	0.8990	0.3489
Slope2	0.0000	0.0129	0.1311	0.0173	-0.0590	0.0063	0.0706	0.0921

Design 8 - Standard deviation =2, N=200, Censoring =50%, Two slope coefficients

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0764	0.9097	0.8293	-0.4333	0.1107	0.5235	0.6486
Slope1	1.0000	1.0477	0.4433	0.1978	0.7552	0.9881	1.3075	0.3359
Slope2	0.0000	-0.0101	0.3243	0.1048	-0.2080	0.0023	0.1685	0.2527
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.5665	0.9175	1.1580	0.4880	0.8204	1.0470	0.9055
Slope1	1.0000	0.7134	0.7539	0.6476	0.2549	0.5290	0.8194	0.6416
Slope2	0.0000	-0.0103	0.3935	0.1542	-0.1947	-0.0049	0.2183	0.2822
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.5899	0.8996	1.1530	0.5194	0.8361	1.0590	0.9145
Slope1	1.0000	0.7093	0.7408	0.6306	0.2583	0.5296	0.8118	0.6370
Slope2	0.0000	-0.0106	0.3888	0.1505	-0.1927	-0.0051	0.2261	0.2791
MP(w_1=0.55)								
Const.	0.0000	0.6919	0.8324	1.1680	0.6203	0.9244	1.1240	0.9617
Slope1	1.0000	0.6883	0.6838	0.5624	0.2879	0.5247	0.7740	0.6155
Slope2	0.0000	-0.0099	0.3612	0.1299	-0.1723	-0.0076	0.1840	0.2586
MP(w_1=0.6)								
Const.	0.0000	0.8302	0.7480	1.2460	0.7684	1.0375	1.2110	1.0370
Slope1	1.0000	0.6596	0.6136	0.4905	0.2999	0.5040	0.7539	0.5937
Slope2	0.0000	-0.0081	0.3268	0.1063	-0.1459	-0.0014	0.1802	0.2350
MP(w_1=0.7)								
Const.	0.0000	1.1115	0.5733	1.5620	1.0960	1.2590	1.3970	1.2070
Slope1	1.0000	0.5880	0.4762	0.3954	0.2977	0.4535	0.6593	0.5674
Slope2	0.0000	-0.0090	0.2429	0.0588	-0.1402	0.0007	0.1331	0.1791

Table 1 (cont.)

Design 9 - Standard deviation =1, N=200, Censoring =25%, Two slope coefficients								
ML	True	Mean	S.D.	M.S.E.	1º Q	Median	3º Q	M.A.E.
Const.	1.0000	0.9815	0.1556	0.0244	0.8778	1.0012	1.0940	0.1257
Slope1	1.0000	1.0078	0.1496	0.0223	0.9190	1.0145	1.0990	0.1167
Slope2	0.0000	-0.0045	0.0986	0.0097	-0.0680	-0.0039	0.0745	0.0801
P	True	Mean	S.D.	M.S.E.	1º Q	Median	3º Q	M.A.E.
Const.	1.0000	0.9790	0.2675	0.0717	0.8689	1.0380	1.1450	0.1918
Slope1	1.0000	1.0091	0.2494	0.0619	0.8338	0.9849	1.1475	0.1923
Slope2	0.0000	-0.0125	0.1216	0.0149	-0.0991	-0.0132	0.0808	0.0981
MP(w_1=0.51)	True	Mean	S.D.	M.S.E.	1º Q	Median	3º Q	M.A.E.
Const.	1.0000	0.9874	0.2626	0.0688	0.8799	1.0440	1.1490	0.1897
Slope1	1.0000	1.0023	0.2454	0.0599	0.8313	0.9803	1.1400	0.1900
Slope2	0.0000	-0.0123	0.1203	0.0146	-0.0984	-0.0135	0.0795	0.0970
MP(w_1=0.55)								
Const.	1.0000	1.0272	0.2436	0.0598	0.9328	1.0795	1.1795	0.1864
Slope1	1.0000	0.9713	0.2297	0.0533	0.8130	0.9545	1.1060	0.1827
Slope2	0.0000	-0.0116	0.1139	0.0130	-0.0913	-0.0128	0.0758	0.0922
MP(w_1=0.6)								
Const.	1.0000	1.0806	0.2177	0.0537	0.9919	1.1290	1.2190	0.1891
Slope1	1.0000	0.9302	0.2106	0.0490	0.7842	0.9184	1.0545	0.1790
Slope2	0.0000	-0.0097	0.1061	0.0113	-0.0799	-0.0118	0.0743	0.0860
MP(w_1=0.7)								
Const.	1.0000	1.1858	0.1651	0.0616	1.1250	1.2145	1.2870	0.2207
Slope1	1.0000	0.8455	0.1673	0.0517	0.7347	0.8239	0.9373	0.1934
Slope2	0.0000	-0.0060	0.0919	0.0084	-0.0692	-0.0085	0.0679	0.0748

A general reduction of the bias for the mean in the P and MP estimator is shown when the censoring is changed from 50% to 25% in design 4, this reduction being most notable in the case of the MP estimator. In addition, the relative efficiency of both the P and the MP estimator show an improvement in relation to the ML estimator, in comparison with design 1.

When the dispersion of the error terms increase in the model $y = 1 + x + u$, the bias and the MSE of the P and MP estimator increases too, as shown in design 4 and 5. This same result was also obtained for the model $y = x + u$, in the comparison of design 1 and 3.

If a new regressor is included in the model $y = x + u$ (design 6, 7 and 8), the ML estimator is the most efficient, even if the differences in efficiency with respect to other estimators for designs 6 and 8 are smaller than under design 1. If we increase the sample size in the last specification from 200 (design 6) to 300 (design 7), the MSE of all the estimators decreases. When a new regressor is included in the specification $y = 1 + x + u$ in design 9, both the P and the MP estimator gain efficiency in relation to the ML estimator. In general, when a new regressor is included in the model, the MP estimator is more efficient than the P estimator.

In order to explore the effects of heteroscedasticity and non normality of the error terms, the results of new designs are presented in table 2. Design 10 considers the Laplace distribution for the error terms. The Cauchy standard distribution is assumed in design 11.

Design 12 and 13 present a 10% per cent normal mixtures with a relative scale of 4 and 9 for the model $y = x + u$, maintaining the unit total variance.

Heteroscedastic normal error terms are considered in design 14 and 15. Increasing heteroscedaticity is assumed in design 14 and decreasing heteroscedasticity in design 15, but in any case unit total variance is assumed.

Results in table 2 show a good behaviour of the MP estimator when the distribution hypothesis of the error terms of the truncated Tobit model are modified.

The MP estimator under the Laplace distribution shows the best behaviour and the ML estimator the worst behaviour, in terms of MSE. The most efficient estimator under the Cauchy distribution is the P estimator. The ML estimator shows very large errors.

Table 2 : Simulation results with heteroscedastic normal error terms.

Design 10 –Laplace, N=200, Censoring=50%								
ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.4920	0.9347	1.1110	-0.7755	-0.3223	0.0715	0.6410
Slope	1.0000	1.1646	0.4741	0.2508	0.8486	1.0890	1.3825	0.3550
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0169	0.5319	0.2818	-0.1187	0.0883	0.2540	0.2921
Slope	1.0000	1.0128	0.4228	0.1780	0.7975	0.9323	1.1145	0.2483
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E*
Const.	0.0000	0.0003	0.5216	0.2707	-0.1009	0.1044	0.2666	0.2914
Slope	1.0000	1.0043	0.4142	0.1708	0.7953	0.9227	1.1045	0.2462
MP(w_1=0.55)								
Const.	0.0000	0.0762	0.4822	0.2371	-0.0210	0.1786	0.3237	0.3003
Slope	1.0000	0.9573	0.3841	0.1486	0.7731	0.8834	1.0495	0.2457
MP(w_1=0.6)								
Const.	0.0000	0.1748	0.4329	0.2170	0.0985	0.2694	0.3949	0.3322
Slope	1.0000	0.8964	0.3476	0.1309	0.7368	0.8307	0.9804	0.2540
MP(w_1=0.7)								
Const.	0.0000	0.3747	0.3319	0.2500	0.2976	0.4383	0.5299	0.4421
Slope	1.0000	0.7649	0.2779	0.1321	0.6283	0.7156	0.8361	0.3028

Table 2 (cont.)

Design 11 –Std. Cauchy, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-12.559	46.108	2273.0	-26.270	3.4840	7.0585	28.890
Slope	1.0000	-4.0933	18.416	363.40	-5.5685	-2.3770	1.3295	8.8190
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0115	0.8743	0.7606	-0.1797	0.2963	0.4948	0.6001
Slope	1.0000	0.9203	0.8927	0.7992	0.5433	0.7891	1.1220	0.5383
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.2912	1.0409	1.1630	0.0852	0.4549	0.6731	0.7792
Slope	1.0000	0.7897	1.0300	1.1000	0.4374	0.7261	1.1325	0.6607
MP(w_1 =0.55)								
Const.	0.0000	1.3301	2.2638	6.8680	0.5485	1.0275	1.4680	1.6130
Slope	1.0000	0.2761	2.2260	5.4540	0.0332	0.4575	1.0750	1.2060
MP(w_1 =0.6)								
Const.	0.0000	2.5044	3.9571	21.850	1.0360	1.6500	2.3350	2.6210
Slope	1.0000	-0.3138	3.9951	17.610	-0.3457	0.1747	0.8452	1.8690
MP(w_1 =0.7)								
Const.	0.0000	4.5890	7.4025	75.580	1.8215	2.5545	3.9185	4.5920
Slope	1.0000	-1.3357	7.5892	62.760	-0.9870	-0.2139	0.4313	3.1670

Design 12 –Normal Mixture, Relative Scale=4, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0319	0.1190	0.0151	-0.1139	-0.0299	0.0528	0.0972
Slope	1.0000	1.0282	0.1117	0.0132	0.9547	1.0325	1.1030	0.0904
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0040	0.0899	0.0081	-0.0502	0.0082	0.0626	0.0698
Slope	1.0000	1.0039	0.0884	0.0078	0.9464	0.9945	1.0465	0.0670
MP(w_1 =0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0011	0.0887	0.0078	-0.0456	0.0126	0.0657	0.0693
Slope	1.0000	0.9996	0.0875	0.0076	0.9406	0.9899	1.0430	0.0666
MP(w_1 =0.55)								
Const.	0.0000	0.0243	0.0841	0.0076	-0.0202	0.0313	0.0879	0.0698
Slope	1.0000	0.9806	0.0844	0.0075	0.9232	0.9757	1.0250	0.0676
MP(w_1 =0.6)								
Const.	0.0000	0.0545	0.0790	0.0092	0.0118	0.0590	0.1142	0.0785
Slope	1.0000	0.9558	0.0811	0.0085	0.8991	0.9511	0.9993	0.0750
MP(w_1 =0.7)								
Const.	0.0000	0.1123	0.0713	0.0177	0.0623	0.1131	0.1650	0.1169
Slope	1.0000	0.9082	0.0765	0.0143	0.8516	0.9137	0.9530	0.1010

Table 2 (cont.)

Design 13 – Normal Mixture, Relative Scale=9, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0125	0.1334	0.0179	-0.0814	0.0041	0.0707	0.0986
Slope	1.0000	1.0102	0.1286	0.0166	0.9253	0.9884	1.0790	0.0980
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0004	0.0439	0.0019	-0.0306	0.0067	0.0332	0.0360
Slope	1.0000	0.9997	0.0520	0.0027	0.9623	0.9972	1.0360	0.0407
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0046	0.0435	0.0019	-0.0274	0.0124	0.0371	0.0363
Slope	1.0000	0.9961	0.0517	0.0027	0.9592	0.9926	1.0335	0.0410
MP(w_1=0.55)								
Const.	0.0000	0.0224	0.0430	0.0023	-0.0079	0.0292	0.0511	0.0413
Slope	1.0000	0.9811	0.0516	0.0030	0.9456	0.9752	1.0160	0.0457
MP(w_1=0.6)								
Const.	0.0000	0.0447	0.0444	0.0040	0.0139	0.0483	0.0742	0.0537
Slope	1.0000	0.9623	0.0530	0.0042	0.9277	0.9573	0.9980	0.0552
MP(w_1=0.7)								
Const.	0.0000	0.0869	0.0521	0.0102	0.0537	0.0914	0.1172	0.0892
Slope	1.0000	0.9263	0.0596	0.0090	0.8842	0.9177	0.9668	0.0809

Design 14 –Heteroscedastic Error, $3E(u_1)^2 = E(u_{200})^2$, N=200, Censoring=50%

ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.1563	0.5268	0.3006	-0.3887	-0.0399	0.2000	0.3977
Slope	1.0000	1.0528	0.3254	0.1081	0.8051	1.0055	1.2640	0.2614
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0867	0.4098	0.1747	-0.1069	0.1450	0.3795	0.3275
Slope	1.0000	0.9446	0.3548	0.1283	0.7056	0.8968	1.1155	0.2771
MP(w_1=0.51)	True	Mean	S.D.	M.S.E.	1º Q	Median	3º Q	M.A.E.
Const.	0.0000	0.1023	0.4022	0.1714	-0.0890	0.1604	0.3891	0.3273
Slope	1.0000	0.9379	0.3487	0.1248	0.7044	0.8879	1.1105	0.2750
MP(w_1=0.55)								
Const.	0.0000	0.1717	0.3731	0.1680	-0.0177	0.2233	0.4438	0.3377
Slope	1.0000	0.8987	0.3234	0.1144	0.6839	0.8540	1.0535	0.2702
MP(w_1=0.6)								
Const.	0.0000	0.2611	0.3354	0.1801	0.0883	0.3185	0.5013	0.3628
Slope	1.0000	0.8460	0.2941	0.1098	0.6436	0.8037	0.9896	0.2737
MP(w_1=0.7)								
Const.	0.0000	0.4436	0.2580	0.2630	0.3109	0.4842	0.6323	0.4681
Slope	1.0000	0.7321	0.2350	0.1267	0.5635	0.6995	0.8439	0.3107

Table 2 (cont.)

Design 15 –Heteroscedastic Error, $E(u_1)^2 = 3E(u_{200})^2$, N=200, Censoring=50%								
ML	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.2495	0.5131	0.3242	-0.5136	-0.1538	0.1060	0.4030
Slope	1.0000	1.0994	0.3375	0.1232	0.8821	1.0515	1.2850	0.2567
P	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	-0.0045	0.4891	0.2308	-0.1465	0.0991	0.3064	0.3469
Slope	1.0000	1.0055	0.4129	0.1697	0.7033	0.9403	1.1960	0.3126
MP(w_1=0.51)	*True*	*Mean*	*S.D.*	*M.S.E.*	*1º Q*	*Median*	*3º Q*	*M.A.E.*
Const.	0.0000	0.0127	0.4797	0.2291	-0.1277	0.1141	0.3192	0.3446
Slope	1.0000	0.9983	0.4060	0.1640	0.6990	0.9341	1.1885	0.3086
MP(w_1=0.55)								
Const.	0.0000	0.0871	0.4431	0.2029	-0.0329	0.1803	0.3711	0.3460
Slope	1.0000	0.9541	0.3776	0.1440	0.6760	0.8963	1.1240	0.2953
MP(w_1=0.6)								
Const.	0.0000	0.1846	0.3983	0.1919	0.0643	0.2651	0.4411	0.3644
Slope	1.0000	0.8963	0.3438	0.1284	0.6463	0.8551	1.0555	0.2890
MP(w_1=0.7)								
Const.	0.0000	0.3852	0.3066	0.2419	0.2705	0.4420	0.5901	0.4478
Slope	1.0000	0.7723	0.2766	0.1280	0.5682	0.7499	0.9025	0.3105

The MP estimator with weights $w_1 = 0.51$ and $w_1 = 0.55$, respectively, is more efficient than the P estimator in design 12 of a mixture of normal distributions with scale 4. The estimators ML and MP with weight $w_1 = 0.7$ are the least efficient.

In design 13 the most efficient estimators are the P and MP estimators with weight $w_1 = 0.51$, while the worst behaviour is shown by the ML estimator.

Finally, the ML estimator is the most efficient for the slope and the least efficient for the constant term in design 14 and 15 corresponding to models with heteroscedastic errors. On the other hand, in design 14 with weight $w_1 = 0.51$ and in design 15 with weight $w_1 = 0.6$, the P estimator presents the worst behaviour for the slope and the MP estimator presents the least MSE for the constant term.

4. CONCLUDING REMARKS

The modified Powell estimator presents a good behaviour throughout the simulation process.

The ML estimator is the most efficient estimator when the standard hypothesis of the truncated Tobit model apply. The MP estimator is, generally, more efficient than the P estimator for any combination of weights w_1 and w_2. When the higher the difference between the weights, the higher the efficiency of the MP estimator, but the bias on the mean is also higher.

The results obtained in the simulations in table 2 show that the behaviour of both the MP and P estimators is better than the behaviour of the ML estimator. However, when we compare the efficiency of the P and MP estimators, no definite conclusion is drawn, as it depends on the design.

APPENDIX

The asymptotic properties of consistency and normality of the modified Powell estimator when $w_1 = w_2 = \frac{1}{2}$ are studied in this appendix.

Required hypothesis for consistency and asymptotic normality

H.1) The true vector of parameters β_0 is an interior point of the compact parametric space B.

H.2) Regressors x_i are independent random vectors, where $E\|x^{4+\eta}\| < K_0$ for any K_0 and η positive, and ν_N, which is the minimum characteristic root of the matrix $T_N = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(x_i'\beta_0 \geq \varepsilon_0\right)x_i x_i'\right]$, verifying $\nu_N > \nu_0$ with $N > N_0$ for some positive ε_0, ν_0 and N_0.

H.3) Error terms u_i are independent and identically distributed. Conditioned on x_i they are continuous and symmetrically distributed around zero, with bounded density, continuous and positive at zero, uniformly in *i*. In other words, consider $F\left(u / x_i, i\right) = F_i(u)$ is the cumulative distribution function of u_i given x_i, then $dF_i(u) = f_i(u)\, du$, where $f_i(u) = f_i(-u)$, $f_i(u) < L_0$ and $f_i(u) > \xi_0$ with $|u| < \xi_0$ for some positive L_0 y ξ_0.

Hypothesis H.1 is usual in estimation asymptotic distribution theory. Even if the only requirement for consistency is a compact parametric space, asymptotic normality requires that the true value of the parametric vector β_0 is an interior point of the space B.

Hypothesis H.2 introduces some restrictions on the regressors that are less often encountered. The minimum characteristic root restriction is necessary for the identification of the true value of the parameter vector.

Hypothesis H.3 assumes continuity of the distribution of the error terms, for convenience, despite that continuity is only required around zero. The density function bounds of the error terms allow control over heteroscedasticity.

Under the above conditions the modified Powell estimator behaviour in large samples is accordingly:

THEOREM A.1 : *Under the above hypothesis H.1-H.3, the modified Powell estimator* $\hat{\beta}_{MPT}$ *defined in equation (3) is strongly consistent and asymptotically normal, that is,* $\hat{\beta}_{MPT} \xrightarrow{c.s.} \beta_0 \quad, N \to \infty$ *and also* $C_N \sqrt{N}\left(\hat{\beta}_{MPT} - \beta_0\right) \xrightarrow{d} N(0\,,1)$ *where* C_N *is the following matrix:*

$$C_N = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(-x_i'\beta_0 < u_i < x_i'\beta_0\right)x_i x_i' + \frac{1}{2}1\left(u_i > x_i'\beta_0\right)x_i x_i' + \frac{1}{2}1\left(u_i > x_i'\beta_0\right)\tilde{x}_i \tilde{x}_i'\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(-x_i'\beta_0 < u_i < +\infty\right)x_i x_i'\right]$$

<u>Proof of the Consistency</u>

In order to show consistency for the modified Powell estimator, we follow the steps of White (1980a).

<u>Lemma A.1</u> (White lemma 2.2)

Consider $Q_N(w,\theta)$ *a measurable function in the measurable space* Ω. *Consider also a continuous function in a compact set* Θ *for each* w *in* Ω. *Then there exists a measurable function* $\hat{\theta}_N(w)$ *such that* $Q_N\left(w,\hat{\theta}_N(w)\right) = \inf_{\theta\in\Theta} Q_N(w,\theta) \qquad \forall w \in \Omega$.

Moreover, if $\left|Q_N(w,\theta) - \overline{Q}_N(\theta)\right| \to 0$ *uniformly* $\forall \theta \in \Theta$ *and if* $\overline{Q}_N(\theta)$ *has a minimum in* θ_0 *which is uniquely identifiable, then* $\hat{\theta}_N \xrightarrow{c.s.} \theta_0$.

The first part of this lemma guarantees the existence of a succession of estimators. The second part shows that the estimators are consistent. It is then necessary to find out adequate functions Q_N y $\overline{Q}_N$.

Let $R_N^*(\beta) = R_N(\beta) - \sum_{i=1}^{N} 1(y_i \geq 2x_i'\beta \,;\, \tilde{y}_i \leq 0)\left(\frac{1}{2}y_i - \frac{1}{2}\tilde{y}_i\right)^2$. This new expression of the objective function avoids the appearance of inconsistent roots. The first order equations are not modified, given that the additional factor does not depend on β.

Let us define the function $Q_N(\beta) = \frac{1}{N}\left[R_N^*(\beta) - R_N^*(\beta_0)\right]$, whose minimisation is equivalent to the minimisation of the objective function $R_N(\beta)$, the first part of the lemma guaranteeing the existence of the least squares estimator.

Let us consider also the function $\overline{Q}_N(\beta) = \mathrm{E}\left[Q_N(\beta)\right]$ which allow us to show the second part of the lemma A.1. We will then follow the following lemma :

<u>Lemma A.2</u> (White lemma 2.3)

Consider (Z_i) independent random variables, taking values in any set Ψ with σ-algebra A.

Let $q_i : \Psi \times \Theta \to R$ where $\Theta \subset R^p$ is compact. Let us suppose :

a) *for any $\theta \in \Theta$, $q_i(z,\theta)$ is A-measurable.*

b) *$q_i(z,\theta)$ is uniformly continuous in Θ in i.*

c) *There exists a measurable function $m_i : \Psi \to R$ such that $\left|q_i(z,\theta)\right| < m_i(z) \quad \forall \theta \in \Theta$ and $\forall i \quad \mathrm{E}\left[m_i(Z_i)\right]^{1+\delta} \leq M < \infty, \quad \delta > 0$.*

Then :

i) $\mathrm{E}\left[q_i(Z,\theta)\right]$ *is uniformly continuous in Θ in i.*

ii) $\sup_{\theta \in \Theta}\left|\frac{1}{N}\sum_{i=1}^{N} q_i(Z_i,\theta) - \overline{Q}_N(\theta)\right| \xrightarrow{c.s.} 0$, *where* $\overline{Q}_N(\theta) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[q_i(Z_i,\theta)\right]$

The regularity hypothesis established above assured that hypothesis a) and b) of lemma A.2 are kept. It is shown now that hypothesis c) is also true in this case.

$$Q_N(\beta) = \frac{1}{N}\sum_{i=1}^{N} 1\left(\frac{1}{2}y_i \le min\{x_i'\beta,\, x_i'\beta_0\}\right)\left[(y_i - x_i'\beta)^2 - (y_i - x_i'\beta_0)^2\right]$$

$$+1(2x_i'\beta \le y_i \le 2x_i'\beta_0)\left[(y_i - x_i'\beta)^2 - (y_i - x_i'\beta_0)^2\right]$$

$$+1(2x_i'\beta_0 \le y_i \le 2x_i'\beta)\left[(y_i - x_i'\beta)^2 - \frac{1}{2}(y_i - x_i'\beta_0)^2 - \frac{1}{2}(\tilde{y}_i - x_i'\beta_0)^2 + \left(\frac{1}{2}y_i - \frac{1}{2}\tilde{y}_i\right)^2\right]$$

$$+\left(\frac{1}{2}y_i \ge max\{x_i'\beta,\, x_i'\beta_0\}\right)\left[\frac{1}{2}(y_i - x_i'\beta)^2 + \frac{1}{2}(\tilde{y}_i - x_i'\beta)^2 - \frac{1}{2}(y_i - x_i'\beta_0)^2 - \frac{1}{2}(\tilde{y}_i - x_i'\beta_0)^2\right]$$

When we look at each term of this function we find out that $Q_N(\beta)$ is bounded by $\|x_i^2\|\left(3\|\beta_0^2\| + 8\|\beta^2\| + 3\|\beta - \beta_0\|^2\right)$.

Thus, the conditions of lemma A.2 are complied, and we can admit that $\lim_{N\to\infty}\left|Q_N(\beta) - \mathrm{E}(Q_N(\beta))\right| = 0$ uniformly in $\beta \in B$.

There is only left to show that $\mathrm{E}(Q_N(\beta))$ has an identifiable unique minimum in β_0, but given that by definition $\mathrm{E}(Q_N(\beta_0)) = 0$, it is sufficient to show that $\mathrm{E}(Q_N(\beta))$ is uniformly strictly greater than zero in β for $\|\beta - \beta_0\| > \varepsilon$ and all N sufficiently large. In this way the conditions of lemma A.1, ensuring consistency of $\hat{\beta}_N$ are obtained.

Given the definition of the function $Q_N(\beta)$ we have :

$$\mathrm{E}(Q_N(\beta)/x_1, x_2, \ldots, x_N) = \frac{1}{N}\sum_{i=1}^{N}\Bigg\{1(x_i'\beta > x_i'\beta_0 > 0)\Bigg[\int_{-x_i'\beta_0}^{x_i'\beta_0}(x_i'\delta)^2 f(u)du$$

$$+\int_{x_i'\beta_0}^{2x_i'\beta - x_i'\beta_0}\left((x_i'\delta) + u\right)^2 f(u)du + \int_{2x_i'\beta - x_i'\beta_0}^{+\infty}(x_i'\delta)^2 f(u)du\Bigg]$$

$$+1(x_i'\beta_0 > x_i'\beta > 0)\left[\int_{-x_i'\beta_0}^{x_i'\beta_0}(x_i'\delta)^2 f(u)du + \int_{x_i'\beta_0}^{+\infty}(x_i'\delta)^2 f(u)du\right]$$

$$+1(x_i'\beta_0 > 0 > x_i'\beta)\left[\int_{-x_i'\beta_0}^{x_i'\beta_0}(x_i'\delta)^2 f(u)du + \int_{x_i'\beta_0}^{+\infty}(x_i'\delta)^2 f(u)du\right]$$

$$+1\left(x_i'\beta > 0 > x_i'\beta_0\right)\left[\int_{-x_i'\beta_0}^{2x_i'\beta - x_i'\beta_0} \left(\left(x_i'\delta\right)+u\right)^2 f(u)du + \int_{2x_i'\beta - x_i'\beta_0}^{+\infty} \left(x_i'\delta\right)^2 f(u)du\right]$$

$$+1\left(x_i'\beta;\, x_i'\beta_0 < 0\right)\left[\int_{-x_i'\beta_0}^{+\infty} \left(x_i'\delta\right)^2 f(u)du\right]\Bigg\}$$

where $\delta = \beta_0 - \beta$.

All terms of the expectation are positive ; thus we have :

$$\mathrm{E}\left(Q_N(\beta)/\, x_1, x_2, \ldots, x_N\right) \geq \frac{1}{N}\sum_{i=1}^{N}\Bigg\{1\left(x_i'\beta_0 > 0 > x_i'\beta\right)\int_{-x_i'\beta_0}^{x_i'\beta_0} \left(x_i'\delta\right)^2 f(u)du$$

$$+1\left(x_i'\beta_0 > x_i'\beta > 0\right)\int_{-x_i'\beta_0}^{x_i'\beta_0} \left(x_i'\delta\right)^2 f(u)du$$

$$+1\left(x_i'\beta > x_i'\beta_0 > 0\right)\int_{-x_i'\beta_0}^{x_i'\beta_0} \left(x_i'\delta\right)^2 f(u)du\Bigg\}.$$

Considering symmetry of the distribution of the error terms and using positive quantities ε_0 y ξ_0 appearing in hypothesis H.2 and H.3, we can assume without loss of generality that $\varepsilon_0 = min\left\{\varepsilon_0, \xi_0\right\}$, then it is verified that for any $\iota \in \left(0, \varepsilon_0\right]$:

$$\mathrm{E}\left(Q_N(\beta)/\, x_1, x_2, \ldots, x_N\right) \geq \frac{1}{N}\sum_{i=1}^{N}\Bigg\{1\left(x_i'\beta_0 \geq \varepsilon_0;\, x_i'\beta < 0\right)1\left(\left|x_i'\delta\right| > \iota\right)2\int_0^{\varepsilon_0} \iota^2 \varepsilon_0\, du$$

$$+1\left(x_i'\beta_0 \geq \varepsilon_0;\, x_i'\beta_0 > x_i'\beta > 0\right)1\left(\left|x_i'\delta\right| > \iota\right)2\int_0^{\varepsilon_0} \iota^2 \varepsilon_0\, du$$

$$+1\left(x_i'\beta_0 \geq \varepsilon_0;\, x_i'\beta > x_i'\beta_0 > 0\right)1\left(\left|x_i'\delta\right| > \iota\right)2\int_0^{\varepsilon_0} \iota^2 \varepsilon_0 du\Bigg\} \geq 2\iota^2\varepsilon_0^2\, \frac{1}{N}\sum_{i=1}^{N} 1\left(x_i'\beta_0 \geq \varepsilon_0\right)1\left(\left|x_i'\delta\right| > \iota\right)$$

Taking expectations over x and applying the inequalities of Holder and Jensen we have that, being k_0 the positive constant of hypothesis H.2 (Powell, 1986) :

$$\mathrm{E}\left(Q_N(\beta)\right) \geq 2\iota^2 \varepsilon_0^2 \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(x_i'\beta_0 \geq \varepsilon_0\right)1\left(\left|x_i'\delta\right| > \iota\right)\right]^3 \left[\mathrm{E}\left\{\|x\|^3\right\} \big/ k_0\right]^2$$

$$\geq 2\iota^2 \varepsilon_0^2 k_0^{-2} \frac{1}{N}\sum_{i=1}^{N} \left\{\mathrm{E}\left[1\left(x_i'\beta_0 \geq \varepsilon_0 ; \left|x_i'\delta\right| > \iota\right)\|x\|^2\right]\right\}^3$$

$$\geq 2\iota^2 \varepsilon_0^2 k_0^{-2} \left\{\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(x_i'\beta_0 \geq \varepsilon_0\right)1\left(\left|x_i'\delta\right| > \iota\right)\left(x_i'\delta\right)^2 \|\delta\|^{-2}\right]\right\}^3$$

$$\geq 2\iota^2 \varepsilon_0^2 k_0^{-2} \left\{\left(\upsilon_N \|\delta\|^2 - \iota^2\right)\|\delta\|^{-2}\right\}^3$$

where υ_N is the minimum characteristic root of matrix T_N of hypothesis H.2, then for large N, $\upsilon_N > \upsilon_0$ and for $\|\delta\| = \|\beta - \beta_0\| > \varepsilon$, an ι is chosen, verifying that $\iota^2 < \upsilon_0 \varepsilon^2$ and thus, it is shown that $\mathrm{E}\left(Q_N(\beta)\right)$ is uniformly bounded over zero in β, for all $N > N_0$.

The application of lemma A.1 shows consistency of the estimators.

Proof of the Asymptotic Normality

We show next the second part of theorem A.1, concerning the asymptotic behaviour of the modified Powell estimator.

The modified Powell estimator $\hat{\beta}_{MPT}$, under regularity hypothesis H.1-H.3, verifies that $C_N \sqrt{N}\left(\hat{\beta}_{MPT} - \beta_0\right) \xrightarrow{d} N(0; I)$, where C_N is a matrix defined as follows :

$$C_N = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left[1\left(-x_i'\beta_0 < u_i < +\infty\right)x_i x_i'\right]$$

The proof is based in a simple extension of the theorem of Huber (1967) for maximum likelihood estimators. The necessary conditions for application of this theorem to our estimators are found in Powell (1984).

First, it is necessary to find out a function $\Psi\left(z_i, \hat{\theta}_N\right)$ satisfying that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Psi\left(z_i, \hat{\theta}_N\right) = o_p(1)$. We will then choose the following function :

$$\Psi\left(u_i, x_i, \beta\right) = 1\left(0 < y_i < 2x_i'\beta\right)\left(y_i - x_i'\beta\right)x_i + \frac{1}{2}1\left(y_i \geq 2x_i'\beta\, ;\, \tilde{y}_i \leq 0\right)\left(y_i - x_i'\beta\right)x_i$$

$$+\frac{1}{2}1\left(y_i \geq 2x_i'\beta\, ;\, \tilde{y}_i \leq 0\right)\left(\tilde{y}_i - x_i'\beta\right)x_i + 1\left(y_i \geq 2x_i'\beta\, ;\, \tilde{y}_i > 0\right)\left(y_i - x_i'\beta\right)x_i$$

which complies with the above condition, given that $-2\sum_{i=1}^{N}\Psi\left(u_i, x_i, \beta\right)$ is equivalent to the first order conditions of the minimisation problem.

The theorem of Huber requires the definition of the following functions :

$$\lambda_N(\beta) = \frac{1}{N}\sum_{i=1}^{N}E\left[\Psi\left(u_i, x_i, \beta\right)\right]$$

$$\mu_i(\beta, d) = \sup_{\|\beta-\gamma\|\leq d}\left\|\Psi\left(u_i, x_i, \beta\right) - \Psi\left(u_i, x_i, \gamma\right)\right\|$$

complying with the following conditions :

a) the function $\mu_i(\beta, d)$ is measurable and has bounded first and second moments :

$$E\left[\mu_i(\beta, d)\right] \leq b \cdot d$$

$$E\left[\mu_i(\beta, d)\right]^2 \leq c \cdot d^2$$, for any *b*, *c*, *d* positive and $\left\|\beta - \beta_0\right\| + d \leq \varepsilon$

b) $\left\|\lambda_N(\beta)\right\| \geq a\left\|\beta - \beta_0\right\|$, for $\left\|\beta - \beta_0\right\| \leq \varepsilon$, $a > 0$ and $\lambda_N\left(\beta_0\right) = 0$.

Under those conditions it can be shown through lemma A.3 of Powell (1984) that if $\hat{\beta}_{MPT}$ is consistent for β_0, then there exist a very simple asymptotic expression for $\lambda_N\left(\hat{\beta}_{MPT}\right)$:

$$\sqrt{N}\lambda_N\left(\hat{\beta}_{MPT}\right) = -\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Psi(u_i, x_i, \beta_0) + o_p(1)$$

It is easy to prove that the required conditions over the function $\mu_i(\beta, d)$ are complied with, given that :

$$\mu_i(\beta, d) = \sup_{\|\beta-\gamma\|\le d}\left\|\Psi(u_i, x_i, \beta) - \Psi(u_i, x_i, \gamma)\right\| = \sup_{\|\beta-\gamma\|\le d}\left\|1(0 < y_i < 2x_i'\beta)(y_i - x_i'\beta)x_i\right.$$

$$+\frac{1}{2}1(y_i \ge 2x_i'\beta\,;\,\tilde{y}_i \le 0)(y_i - x_i'\beta)x_i + \frac{1}{2}1(y_i \ge 2x_i'\beta\,;\,\tilde{y}_i \le 0)(\tilde{y}_i - x_i'\beta)x_i$$

$$+1(y_i \ge 2x_i'\beta\,;\,\tilde{y}_i > 0)(y_i - x_i'\beta)x_i - 1(0 < y_i < 2x_i'\gamma)(y_i - x_i'\gamma)x_i$$

$$-\frac{1}{2}1(y_i \ge 2x_i'\gamma\,;\,\tilde{y}_i \le 0)(y_i - x_i'\gamma)x_i - \frac{1}{2}1(y_i \ge 2x_i'\gamma\,;\,\tilde{y}_i \le 0)(\tilde{y}_i - x_i'\gamma)x_i$$

$$\left.-1(y_i \ge 2x_i'\gamma\,;\,\tilde{y}_i > 0)(y_i - x_i'\gamma)x_i\right\|$$

which is bounded by $\|x_i\|^2 d$.

According to the regularity hypothesis H.2, we have that $E\left[\|x_i\|\right]^{4+\eta} < K_0$ or equivalently $E\left[\|x_i\|\right] < K_0^{1/4+\eta}$. Thus $E\left[\mu_i(\beta, d)\right] < K_0^{2/4+\eta}\, d$ and $E\left[\mu_i(\beta, d)\right]^2 < K_0^{4/4+\eta}\, d^2$ as required.

Condition b) required over $\lambda_N(\beta)$ is not obvious. In order to show it let us prove that $\lambda_N(\beta) = -C_N(\beta - \beta_0) + O(\|\beta - \beta_0\|)$.

$$\left\|\lambda_N(\beta) + C_N(\beta - \beta_0)\right\| = E_x\left\|\frac{1}{N}\sum_{i=1}^{N}\left\{1(x_i'\beta > 0)\int_{-x_i'\beta_0}^{2x_i'\beta - x_i'\beta_0}(y_i - x_i'\beta)x_i f(u)du\right.\right.$$

$$+\frac{1}{2}\int_{2x_i'\beta-x_i'\beta_0}^{+\infty}\left(y_i-x_i'\beta\right)x_i\,f(u)du+1\left(x_i'\beta_0>0;x_i'\beta<x_i'\beta_0\right)\frac{1}{2}\int_{2x_i'\beta-x_i'\beta_0}^{x_i'\beta_0}\left(y_i-x_i'\beta\right)x_i\,f(u)du$$

$$+1\left(x_i'\beta_0>0;x_i'\beta<x_i'\beta_0\right)\frac{1}{2}\int_{x_i'\beta_0}^{+\infty}\left(\tilde{y}_i-x_i'\beta\right)x_i\,f(u)du$$

$$+1\left(x_i'\beta_0<0;x_i'\beta<x_i'\beta_0\right)\frac{1}{2}\int_{2x_i'\beta-x_i'\beta_0}^{+\infty}\left(\tilde{y}_i-x_i'\beta\right)x_i\,f(u)du$$

$$+1\left(x_i'\beta_0<x_i'\beta\right)\frac{1}{2}\int_{2x_i'\beta-x_i'\beta_0}^{+\infty}\left(\tilde{y}_i-x_i\beta\right)x_i\,f(u)\,du\Bigg\}+\frac{1}{N}\sum_{i=1}^{N}\left\{\int_{-x_i'\beta_0}^{+\infty}x_ix_i'\left(\beta-\beta_0\right)f(u)\,du\right\}\Bigg\|.$$

Considering hypothesis H.1-H.3 on the regressors and the error terms above we arrive to the following result :

$$\left\|\lambda_N(\beta)+C_N\left(\beta-\beta_0\right)\right\|\leq 2\|\delta\|K_0^{2/4+\eta}\left(1+L_0K_0\|\beta\|\right).$$

Given that condition b), the lemma A.3 of Powell guarantees that :

$$-\frac{1}{N}\sum_{i=1}^{N}\Psi\left(u_i\,,x_i\,,\,\beta_0\right)+o_p(1)=-C_N\left(\beta-\beta_0\right)+O\left(\left\|\beta-\beta_0\right\|\right)$$

thus we have $C_N\sqrt{N}\left(\beta-\beta_0\right)=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Psi\left(u_i\,,x_i\,,\,\beta_0\right)+o_p(1)$, and the application of the Central Limit Theorem of Liapunov to this expression lead us to asymptotic normality, as we wanted to prove.

REFERENCES

AMEMIYA, T. (1973): “Regression Analysis when the Dependent Variable is Truncated Normal”, Econometrica, 41, 997-1016.

POWELL, J. L. (1984): “Least Absolute Deviations Estimation for the Censored Regression Model”, Journal of Econometrics, 25, 303-325.

POWELL, J. L. (1986): “Symmetrically Trimmed Least Squares Estimation for Tobit Models”, Econometrica, 54, 1435-1460.

WHITE, H. (1980a): “Nonlinear Regression on Cross-Section Data”, Econometrica, 48, 721-746.

WHITE, H. (1980b): “A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, Econometrica, 48, 817-838.