Utilizing the Flexibility of the Epsilon-Skew-Normal Distribution for Tobit Regression Problems

Terry L. Mashtare Jr.

Department of Biostatistics

University at Buffalo, 249 Farber Hall, 3435 Main Street, Buffalo, NY 14214-3000, U.S.A.

Roswell Park Cancer Institute, Elm & Carlton Streets, Buffalo, NY 14263, U.S.A.

Alan D. Hutson

Department of Biostatistics

University at Buffalo, 249 Farber Hall, 3435 Main Street, Buffalo, NY 14214-3000, U.S.A.

February 11, 2008

Summary

The Tobit model was introduced by James Tobin in 1958 in order to model a specific type of discrete-continuous data commonly found in economic applications. The Tobit model is a specific case of a censored regression model and assumes that the continuous component of the data (right-tail) is normally distributed. It has been demonstrated in later research that even small departures from this assumption may lead to inconsistent estimators. One technique that is often utilized in an attempt to compensate for this weakness is to apply a log transformation to the data. How-

1

ever, we illustrate that for many biological applications this approach is oftentimes inadequate. An alternative approach is to utilize more flexible parametric models, such as the epsilon-skew-normal (ESN) developed by Mudholkar and Hutson (2000), in order to create a logical extension to the original Tobit Model. We show that this approach lends itself well towards generalizing Tobit regression models in terms of providing consistent and efficient parameter estimates.

Key Words: left-censoring; limit of detection; maximum likelihood estimation; regression modeling

1 Introduction

The Tobit model was introduced by James Tobin in 1958 in order to model a specific type of discrete-continuous data commonly found in economic applications. The Tobit model is a specific case of a censored regression model and assumes that the continuous component of the data (right-tail) is normally distributed. Early examples included modeling household expenditures of luxury goods, inheritance, and expected age of retirement. The Tobit model was first used in economics applications. More recently, Tobit models have been used for biostatistical applications such as analyzing data where measurements fall below a limit of detection. A few examples among many include cord blood IgE levels, Scirica et al (2007) and coronary artery calcification, Reilly et al (2004).

It has been demonstrated in later research that even small departures from underlying normality assumption may lead to inconsistent estimators. Arabmazar and Schmidt (1982) explored the robustness of the Tobit estimator when estimating a population mean when the assumption of normality is violated. They concluded that the bias can be quite large and that the bias is dependent on the proportion of

2

censoring.

One technique that is often utilized in an attempt to compensate for this weakness in the case of long-tailed distributions is to apply a log transformation to the data. Lorimer and Kiermeier (2007) conduct a simulation study to examine the use of Tobit models on log-transformed microbiological data. They compared the Tobit method to two other methods, using only uncensored observations, and using the limit of detection for the censored values. They concluded that the two standard methods led to biased estimates and that the Tobit model led to less biased estimates. However, their conclusions are based on the underlying assumption of normality.

In addition to log-transformations, people have considered other types of Box-Cox transformations, such as the square root transform. Han and Kronmal (2004) examine the use of Box-Cox transformations for both linear and non-linear Tobit models. They develop a method for choosing an appropriate data-based transformation.

Another approach towards improving the model fit is to consider semi-parametric approaches. Chay and Powell (2001) compared the censored least absolute deviations (CLAD) estimation, symmetrically censored least squares (SCLS) estimation, and identically censored least absolute deviations (ICLAD) estimation. Chay and Powell suggested that Tobit estimators can be biased when the normality assumption is violated.

An alternative approach, which we consider, is to utilize more flexible parametric models, such as the epsilon-skew-normal (ESN) developed by Mudholkar and Hutson (2000), in order to create a logical extension to the original Tobit model. We show that this approach lends itself well towards generalizing Tobit regression models in terms of providing consistent and efficient parameter estimates.

In section 2 we review the Tobit model and we define a new model termed the Epsilon-Skew-Normal (ESN) Tobit model. In section 3 we then use the maximum-likelihood approach to estimate the model parameters. In section 4 we then sum-

marize the results of a simulation study we conducted to examine the asymptotic properties of the maximum likelihood estimates in the ESN Tobit model and compare with the maximum likelihood estimates in the Tobit model. In section 5 we then illustrate the maximum likelihood estimation of the ESN Tobit regression model using data involving lung injury in 276 mice.

2 Background

2.1 Tobit Regression

Let Y^* denote a random variable with mean θ. Suppose we observe $Y = \max\{0, Y^* = \theta + \zeta\}$ and the error term ζ is assumed to be normally distributed with mean 0 and variance σ^2. Let ϕ and Φ represent the standard normal d.f. and c.d.f respectively.

The first two moments of Y are

$$E(Y) = \theta\Phi(\theta/\sigma) + \sigma\phi(\theta/\sigma), \tag{2.1}$$

$$E(Y^2) = (\theta^2 + \sigma^2)\Phi(\theta/\sigma) + \theta\sigma\phi(\theta/\sigma), \tag{2.2}$$

Let θ be defined as a linear combination of parameters, namely,

$$\theta = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}, \tag{2.3}$$

where, using standard notation, $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_{p-1})'$ denotes the $p \times 1$ vector of known covariates and β_j $(j = 0, 1, \ldots, p-1)$ are the p unknown regression coefficients. More compactly use standard matrix notation and denote the mean $\theta = \boldsymbol{x}'\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})$.

Let y_i, $i = 1, 2, \ldots, n$ denote observations from of a sample of size n with the corresponding vector of covariates $\boldsymbol{x}_i$ and assumed i.i.d. error terms $\zeta_i \sim N(0, \sigma^2)$. Let δ_i be an indicator variable that is 1 when $y_i > 0$ and 0 otherwise.

The log-likelihood per observation for the Tobit regression model is given by

$$l(y_i; \boldsymbol{\beta}, \sigma^2) \quad = \quad (1 - \delta_i) \log \left[\Phi \left(\frac{-\boldsymbol{x}_i' \boldsymbol{\beta}}{\sigma} \right) \right] - \frac{\delta_i}{2} \left[\log 2\pi\sigma^2 + \frac{(y_i - \boldsymbol{x}_i' \boldsymbol{\beta})^2}{\sigma^2} \right]; \ (2.4)$$

See Amemiya (1973) for detailed theoretical results for the Tobit Model.

2.2 Epsilon-Skew-Normal (ESN) distribution

The epsilon-skew-normal distribution was developed by Mudholkar and Hutson (2000). What we term the "base model" or standardized model for the epsilon-skew-normal distribution $\mathrm{ESN}(\theta, \sigma, \epsilon)$ is defined to be a unimodal distribution with the mode at θ and probability mass $(1 - \epsilon)/2$ below the mode. The probability density function (p.d.f.), the distribution function (d.f.), and quantile function (q.f.) of its canonical form $ESN(0, 1, \epsilon)$ are respectively:

$$f_0(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2(1-\epsilon)^2} \right), & \text{if } x < 0, \\[2mm] \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2(1+\epsilon)^2} \right), & \text{if } x \geq 0, \end{cases} \tag{2.5}$$

$$F_0(x) = \begin{cases} (1 - \epsilon)\Phi \left(\frac{x}{1-\epsilon} \right), & \text{if } x < 0, \\[2mm] -\epsilon + (1 + \epsilon)\Phi \left(\frac{x}{1+\epsilon} \right), & \text{if } x \geq 0, \end{cases} \tag{2.6}$$

and,

$$Q_0(u) = F_0^{-1}(u) = \begin{cases} (1 - \epsilon)\Phi^{-1} \left(\frac{u}{1-\epsilon} \right), & \text{if } 0 < u < (1 - \epsilon)/2 \\[2mm] (1 + \epsilon)\Phi^{-1} \left(\frac{u+\epsilon}{1+\epsilon} \right), & \text{if } (1 - \epsilon)/2 \leq u < 1, \end{cases} \tag{2.7}$$

where $-1 < \epsilon < 1$, and $\Phi(x)$ denotes the standard normal c.d.f.

The standard epsilon-skew-normal distribution, $ESN(0, 1, \epsilon)$, is a mixture of two half-normal distributions and reduces to the standard normal distribution when $\epsilon = 0$. The distribution is skewed right for values of $\epsilon > 0$ and skewed left for values of $\epsilon < 0$. The limiting cases of (2.5) as $\epsilon \to \pm 1$ are the well known half-normal distributions. Figure 1 gives some typical ESN probability density functions.

The p.d.f. $f_0(\cdot)$ at (2.5) has derivatives of arbitrary orders. It is differentiable once at the mode. The p.d.f. of $ESN(\theta, \sigma, \epsilon)$ is $f_0(\frac{x-\theta}{\sigma})/\sigma$, where $f_0(\cdot)$ is given by (2.5). The c.d.f. of $ESN(\theta, \sigma, \epsilon)$ is $F_0(\frac{x-\theta}{\sigma})$, where $F_0(\cdot)$ is given by (2.6). Its quantile function, $Q(u) = \theta + \sigma Q_0(u)$, can be used to generate samples from the $ESN(\theta, \sigma, \epsilon)$ population. Note the relationship $Q(\frac{1-\epsilon}{2}) = \theta$. The mean is given by

$$E(X) \quad = \quad \theta + \frac{4\sigma\epsilon}{\sqrt{2\pi}}, \tag{2.8}$$

Mudholkar and Hutson (2000) give an in depth mathematical treatment of the maximum likelihood estimation for the ESN distribution in the univariate case. Namely, they illustrate that the likelihood estimates are well behaved. Mudholkar and Hutson give detailed theoretical results pertaining to this model.

3 Epsilon-Skew-Normal Tobit Regression Model

3.1 Univariate Case

Let Y^* denote a random variable with mode θ where $Y = \max\{0, \theta + \zeta\}$. For the ESN Tobit model the error term ζ is assumed to have an $ESN(0, \sigma, \epsilon)$ distribution. The parameters of interest for the ESN Tobit regression model are similar to those in the Tobit regression framework with the exception of the additional parameter ϵ.

The mean of the ESN Tobit regression model is given by

$$E(Y) = \begin{cases} \theta[1 - K_{\theta,\sigma}] + \sigma(1 + \epsilon)^2 \kappa_{\theta,\sigma}, & \text{if } \theta \le 0, \\ \theta[1 - K_{\theta,\sigma}] + \sigma(1 - \epsilon)^2 \kappa_{\theta,\sigma} + \frac{4\sigma\epsilon}{\sqrt{2\pi}}, & \text{if } \theta > 0, \end{cases} \tag{3.1}$$

where $\kappa_{\theta,\sigma} = f_0(-\theta/\sigma)$ and $K_{\theta,\sigma} = F_0(-\theta/\sigma)$, and $f_0(\cdot)$ and $F_0(\cdot)$ are the d.f. and c.d.f. of the standard ESN distribution respectively.

Note that when $\epsilon = 0$ corresponding to normal error term, (3.1) reduces to (2.1).

The second moment of the ESN Tobit regression model is given by

$$E(Y^2) = \begin{cases} [\theta^2 + \sigma^2(1+\epsilon)^2][1 - K_{\theta,\sigma}] + \theta\sigma(1+\epsilon)^2\kappa_{\theta,\sigma}, & \text{if } \theta \leq 0, \\ \theta^2 - [\theta^2 + \sigma^2(1-\epsilon)^2]K_{\theta,\sigma} + \theta\sigma(1-\epsilon)\kappa_{\theta,\sigma} + \frac{8\theta\sigma\epsilon}{\sqrt{2\pi}} + \sigma^2(3\epsilon^2+1), & \text{if } \theta > 0, \end{cases} \tag{3.2}$$

where κ and K are defined above.

In the context of Tobit regression modeling the parameter ϵ may be interpreted as a measure of distance of the error ζ from normality with respect to ESN skew alternatives. Hence, the test

$$H_0 : \epsilon = 0$$

$$H_1 : \epsilon \neq 0 \tag{3.3}$$

may be used as a regression diagnostic tool with respect to the appropriateness of testing the appropriateness of an underlying normal model versus a broad class of skew normal alternatives.

Let $Y_1, Y_2, \ldots, Y_n$ denote i.i.d. $\zeta_i \sim ESN(0, \sigma, \epsilon)$, $i = 1, 2, \ldots, n$. Let δ_i be an indicator variable that is 1 when $y_i > 0$ and 0 otherwise.

The log-likelihood per observation for the ESN Tobit regression model is given by

$$l(y_i; \theta, \sigma^2, \epsilon) = (1 - \delta_i) \log\left[F_0\left(\frac{-\theta}{\sigma}\right)\right] + \delta_i \log\left[\frac{1}{\sigma}f_0\left(\frac{y_i - \theta}{\sigma}\right)\right]. \tag{3.4}$$

More specifically, we need to consider the cases where $\theta \leq 0$, $\theta > 0$ and $y_i < \theta$, and $0 < \theta \leq y_i$.

For $\theta \leq 0$ the log-likelihood per observation takes the form

$$l(y_i; \theta, \sigma^2, \epsilon) = (1 - \delta_i) \log\left[-\epsilon + (1+\epsilon)\Phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)\right]$$
$$- \delta_i\left[\frac{1}{2}\log(2\pi\sigma^2) + \frac{(y_i - \theta)^2}{2\sigma^2(1+\epsilon)^2}\right], \tag{3.5}$$

7

For $\theta > 0$ and $y_i < \theta$ the log-likelihood per observation takes the form

$$l(y_i; \theta, \sigma^2, \epsilon) = (1 - \delta_i) \log \left[(1 - \epsilon) \Phi \left(\frac{-\theta}{\sigma(1 - \epsilon)} \right) \right]$$
$$- \delta_i \left[\frac{1}{2} \log(2\pi\sigma^2) + \frac{(y_i - \theta)^2}{2\sigma^2(1 - \epsilon)^2} \right], \tag{3.6}$$

and for $0 < \theta \leq y_i$ the log-likelihood per observation takes the form

$$l(y_i; \theta, \sigma^2, \epsilon) = -\tfrac{1}{2} \log(2\pi\sigma^2) - \frac{(y_i - \theta)^2}{2\sigma^2(1 + \epsilon)^2}. \tag{3.7}$$

The score functions are given in Appendix A. Note however direct maximum likelihood solutions for the estimates of the vector of parameters β are analytically intractable. An alternative approach is to utilize numerical nonlinear maximization routines such as SAS's PROC NLMIXED and PROC NLP. The basic requirements are an expression for the loglikelihood per observation, reasonable starting values for the parameters β, ϵ, and σ, and programming statements corresponding to any parameter bounds.

The general approach for determining starting values for the maximum likelihood estimation method is to first fit the standard Tobit model ($\epsilon = 0$). The parameter estimates from the Tobit model are then used as the starting values for θ and σ with the ϵ set to zero as the starting value. Equations (3.1) and (3.2) provide alternative mechanics for parameter estimation and starting values via the method of moments.

Covergence of the fitting algorithm is given by relative gradient convergence criterion GCONV $< 1E-8$. Error estimates are computed using the inverse of the Hessian matrix. The SAS/STAT 9.1 user guide (2004) has detailed information regarding convergence criteria and standard error estimation when using the NLMIXED procedure.

3.2 Multivariate Regression Case

We extended the univariate case to the multivariate regression case through the parameter θ. We define θ as a linear combination of parameters, namely,

$$\theta = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1}, \tag{3.8}$$

where, using standard notation, $\boldsymbol{x} = (1, x_1, x_2, \ldots, x_{p-1})'$ denotes the $p \times 1$ vector of known covariates and β_j $(j = 0, 1, \ldots, p-1)$ are the p unknown regression coefficients. More compactly use standard matrix notation and denote the mode $\theta = \boldsymbol{x}'\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})$.

Note that when $\epsilon = 0$ this reduces to the standard Tobit regression and further reduces to classic least squares regression when all δ_i, $i = 1, 2, \ldots, n$.

The maximum likelihood estimation for this model is a straightforward extension to the univariate ESN Tobit model discussed in the previous section. In section 5, we look at examples fitting the ESN Tobit model for both the univariate and the multivariate regression case.

4 Simulation

We conducted a simulation study in order to examine the behavior of the maximum likelihood estimates of the location parameter for the Tobit model over a variety of scenarios.

For the univariate case, we examined the estimated coverage probability for the normal model fit versus the ESN model fit for samples of size 25, 50, and 100 for the left censored regression model $Y_i = \max\{0, \theta + \zeta_i\}$ over the values of $\epsilon = 0$ to 0.75 by 0.25 using 10,000 replications, where $\zeta_i \sim \text{ESN}(0, 1, \epsilon)$.

Let $Y^* \sim ESN(\theta, \sigma, \epsilon)$ and let $c = P(Y^* \le 0)$. The range of simulation values for θ were chosen by solving $\theta + \sigma Q_0(c) = 0$ for θ, where Q_0 is given by equation

9

(2.7). Letting $c = 0.2, 0.3$, and 0.4 gives values for θ that simulate samples with an average of 20%, 30%, and 40% censoring respectively.

The parameter of interest for our simulation study was $\mu = E(Y^*)$, where $\mu = \theta$ for the normal model ($\epsilon = 0$), and $\mu = \theta + \frac{4\sigma\epsilon}{\sqrt{2\pi}}$ for the ESN model. Under each model, 95% confidence intervals for μ were computed. The coverage probabilities were estimated by determining the proportion of times the computed confidence interval captured the true value of μ in the 10,000 replications.

Tables 1 and 2 gives $\hat{\mu} \pm \hat{\sigma}_{\hat{\mu}}$, where $\hat{\sigma}_{\hat{\mu}}$ is the standard deviation of the estimates for μ, along with the estimated coverage probability for each simulation. When simulating under the normal distribution ($\epsilon = 0$), both the Tobit and ESN Tobit models have comparable and acceptable coverage probabilities. As we varied the skewness parameter ϵ in an increasing fashion, the Tobit model began to have coverage probabilities below the 95% level for larger samples, falling as low as 81%. This is similar to behavior seen earlier by Arabmazar and Schmidt (1982) and Chay and Powell (2001). It is interesting to note here that the estimated coverage probability for the normal model decreases as the sample size increases. The implication is that potentially the estimate of the mean may be inconsistent.

For the multivariate regression case, we fit the model $Y_i = \max\{0, \beta_0 + \beta_1 x_i + \zeta_i\}$ where $\zeta_i \sim \text{ESN}(0, 1, \epsilon)$ and $\beta_0 = 0.5$. The value of x_i was defined as an indicator variable given by 0 or 1 divided evenly by sample size. For $n = 25$, $n_1 = 12$ had a covariate level of 0, and $n_2 = 13$ had a covariate level of 1.

Choosing the range of simulation values for β_1 is more complex than in the univariate case. For this purpose we used a two step process. First, simply let $\theta_i = \beta_0 + \beta_1 x_i$ and then average both side sides to arrive at $\bar{\theta} = \beta_0 + \beta_1 \bar{x}$. We then chose values of $\bar{\theta}$ using the same method as in the univariate case. Since we fixed $\beta_0 = 0.5$ and the average of the covariate indicator variables, $\bar{x}$ is 0.5, we then choose β_1 using the relation $\beta_1 = \frac{\bar{\theta} - 0.5}{0.5}$. These values of β_1 simulate samples with an average

censoring proportion near to 20%, 30%, and 40% and were applied consistently across model comparisons.

Tables 3 and 4 gives $\hat{\beta}_1 \pm \hat{\sigma}_{\hat{\beta}_1}$, where $\hat{\sigma}_{\hat{\beta}_1}$ is the standard deviation of the estimates for β_1, along with the estimated coverage probability for each simulation. When simulating under the Tobit assumptions the estimated coverage probability was slightly below the expected 95% level for the ESN Tobit model for small samples. As expected, the estimated coverage probability for β_1 did approach 95% as the sample size increased.

As ϵ increased, the estimated coverage probability β_1 using Tobit model was slightly below the 95% level. Interestingly, we did not see as much of a decrease in the coverage probability for β_1 as the sample size increased as we did in the location case.

In general, when $\epsilon = 0$ both the Tobit and ESN Tobit models performed equally well. There was no substantial penalty for using ESN Tobit model when Tobit model assumptions held. When $\epsilon > 0$ the Tobit model did not perform as well as the ESN Tobit model.

5   Application

We illustrate the maximum likelihood estimation of the ESN Tobit regression model using data involving lung injury in 276 mice, see Raghavendran et al. (2005). There were two groups of mice, wild type (WT), and MCP-1$(-, -)$ mice. There were 145 WT mice and 131 MCP-1$(-, -)$ mice. The mice were sacrificed at 5, 24, or 48 hours to assess lung injury. For purpose of illustration we chose two of the set of cytokines used in this study as our response variables, namely IL-1β and MIP-2. The limit of detection for both IL-1β and MIP-2 was 32 pg/ml. Note that we subtracted the limit of detection from all the observations before fitting the Tobit and ESN Tobit models.

We then simply added the limit of detection to the location parameter estimates to arrive at the final fitted results. The histograms of the data are given in figures 2 and 3. The percent of values falling below the limit of detection was 45.7% for IL-1β and 21.7% for MIP-2.

We first estimated the mean IL-1β level using a univariate model and a multivariate model controlling for mouse type and time. We the fit the OLS model, the Tobit Model, and the ESN Tobit Model. The results are summarized in tables 5 and 7.

It is interesting to note the p-value for testing $H_0 : \epsilon = 0$, see (3.3) for details. For the univariate model for the mean IL-1β level, the skewness parameter in the ESN Tobit model is not significantly different from zero (p = 0.69). This suggests that the Tobit model may be appropriate for modeling the mean IL-1β levels. In the Tobit model, we conclude that the mean IL-1β level is significantly different from zero (p = 0.0058) while in the ESN Tobit model, we conclude that the mean IL-1β level is not significantly different from zero (p = 0.2675), at $\alpha = 0.05$.

For the multivariate model for the mean IL-1β levels with covariates mouse type and time, the skewness parameter in the ESN Tobit model is not significantly different from zero (p = 0.6217). Again, this suggests that the Tobit model may be appropriate. In both the Tobit and ESN Tobit models, we conclude that there is no significant difference in the IL-1β levels between WT mice and MCP-1$(-,-)$ mice (p = 0.6105 and p= 0.5234 respectively). In both the Tobit and ESN Tobit models, we conclude that there is a significant time effect (p = 0.0013 and p = 0.0014 respectively).

Next, we estimated the mean MIP-2 levels using a univariate model and a multivariate model again with covariates mouse type and time. The results are summarized in tables 6 and 8.

For the univariate model for the mean MIP-2 level, the skewness parameter in

the ESN Tobit model is highly significantly different from zero (p < .0001). This suggests that the Tobit model likely a poor choice for modeling mean MIP-2 level. We see in table 6 the confidence interval for the mean MIP-2 level in the ESN Tobit model suggests the mean MIP-2 level is higher than the mean MIP-2 level estimated in the Tobit model.

For the multivariate model for the mean MIP-2 level, the skewness parameter in the ESN Tobit model is highly significantly different from zero (p < .0001). Again, this suggests that the Tobit model may not be appropriate. In the Tobit model, we conclude that there is a significant difference in the MIP-2 level between WT mice and MCP-1$(-,-)$ mice (p = 0.0318). In the ESN Tobit model, we conclude that there is no significant difference in the MIP-2 level between WT mice and MCP-1$(-,-)$ mice (p = 0.7869).

We see that when fitting IL-1β using the ESN Tobit model, we fail to reject the hypothesis that $\epsilon = 0$. Both the Tobit and ESN Tobit models yielded similar results. When fitting MIP-2 using the ESN Tobit model, we strongly rejected the hypothesis that $\epsilon = 0$. Thus the Tobit model potentially gives the misleading conclusion that there is no difference in the MIP-2 levels between WT mice and MCP-1$(-,-)$ mice. This example illustrates how the Tobit model may lead to false conclusions when the underlying assumptions are not met.

Acknowledgements

Appendix A

The score equation per observation for θ, σ, and ϵ are as follows:

$$\frac{\partial l(y_i; \theta, \sigma^2, \epsilon)}{\partial \theta} = \begin{cases} \dfrac{-(1-\delta_i)\phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)}{\sigma\left[-\epsilon+(1+\epsilon)\Phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)\right]} + \dfrac{\delta_i(y_i-\theta)}{\sigma^2(1+\epsilon)^2} & \text{if } \theta \le 0, \\[4ex] \dfrac{-(1-\delta_i)\theta\phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)}{\sigma(1-\epsilon)\Phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)} & \text{if } \theta > 0 \text{ and } y_i < \theta, \\[4ex] \dfrac{(y_i-\theta)}{\sigma^2(1+\epsilon)^2} & \text{if } 0 < \theta \le y_i. \end{cases} \qquad (5.1)$$

$$\frac{\partial l(y_i; \theta, \sigma^2, \epsilon)}{\partial \sigma^2} = \begin{cases} \dfrac{(1-\delta_i)\theta\phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)}{2\sigma^3\left[-\epsilon+(1+\epsilon)\Phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)\right]} + \delta_i\left[\frac{-1}{2\sigma^2} + \frac{(y_i-\theta)^2}{2\sigma^4(1+\epsilon)^2}\right] & \text{if } \theta \le 0, \\[4ex] \dfrac{(1-\delta_i)\theta\phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)}{2\sigma^3(1-\epsilon)\Phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)} + \delta_i\left[\frac{-1}{2\sigma^2} + \frac{(y_i-\theta)^2}{2\sigma^4(1-\epsilon)^2}\right] & \text{if } \theta > 0 \text{ and } y_i < \theta, \\[4ex] \dfrac{-1}{2\sigma^2} + \dfrac{(y_i-\theta)^2}{2\sigma^4(1+\epsilon)^2} & \text{if } 0 < \theta \le y_i. \end{cases}$$
$$(5.2)$$

$$\frac{\partial l(y_i; \theta, \sigma^2, \epsilon)}{\partial \epsilon} = \begin{cases} (1-\delta_i)\left[\dfrac{-1+\Phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)+\frac{\theta}{\sigma(1+\epsilon)}\phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)}{-\epsilon+(1+\epsilon)\Phi\left(\frac{-\theta}{\sigma(1+\epsilon)}\right)}\right] + \delta_i\left[\frac{(y_i-\theta)^2}{\sigma^2(1+\epsilon)^3}\right] & \text{if } \theta \le 0, \\[4ex] (1-\delta_i)\left[\dfrac{-1}{1-\epsilon} - \dfrac{\theta\phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)}{\sigma(1-\epsilon)^2\Phi\left(\frac{-\theta}{\sigma(1-\epsilon)}\right)}\right] + \delta_i\left[\frac{(y_i-\theta)^2}{\sigma^2(1-\epsilon)^3}\right] & \text{if } \theta > 0 \text{ and } y_i < \theta, \\[4ex] \dfrac{(y_i-\theta)^2}{\sigma^4(1+\epsilon)^3} & \text{if } 0 < \theta \le y_i. \end{cases}$$
$$(5.3)$$

References

Amemiya, T. (1973) Regression Analysis When the Dependent Variable is Truncated Normal. *Econometrica* **41**, 997-1016.

Arabmazar, A. and Schmidt, P. (1982) An Investigation of the Robustness of the Tobit Estimator to Non-Normality. *Econometrica* **50**, 1055-1064.

Chay K.Y. and Powell J.L. (2001) Semiparametric censored regression models. *Journal of Economic Perspectives* **15**, 29-42.

Han, C. and Kronmal, R. (2004) Box-Cox transformation of left-censored data with application to the analysis of coronary artery calcification and pharmacokinetic data. *Statistics in Medicine* **23**, 3671-3679.

Lorimer, M.F. and Kiermeier, A. (2007) Analysing microbiological data: Tobit or not Tobit? *International Journal of Food Microbilogy* **116**, 313-318.

Mudholkar, G.S. and Hutson, A.D. (2000) The Epsilon-Skew-Normal distribution for Analyzing Near-Normal data. *Journal of Statistical Planning and Inference* **83**, 291-309.

SAS (v9.1) SAS Institute Inc., Cary, NC, USA.

SAS Institute Inc. 2004. SAS/STAT 9.1 User's Guide. Cary, NC: SAS Institute Inc.

Raghavendran, K., Davidson, B.A., Mullan, B.A, et al. (2005) Acid and particulate-induced aspiration lung injury in mice: importance of MCP-1. *Am J Physiol Lung Cell Mol Physiol* **289**, 134-143.

Reilly, M.P., Wolfe, M.L., Localio, A.R., and Rader D.J. (2004) Coronary artery calcification and cardiovascular risk factors: impact of the analytic approach. *Atherosclerosis* **173**, 69-78.

Scirica, C.V., Gold, D.R., Ryan, L., et al. (2007) Predictors of cord blood IgE levels in children at risk for asthma and atopy. *Journal of Allergy and Clinical Immunology* **111**, 81-88.

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

Figure 1: Some typical $ESN(0, 1, \epsilon)$ probability density functions.
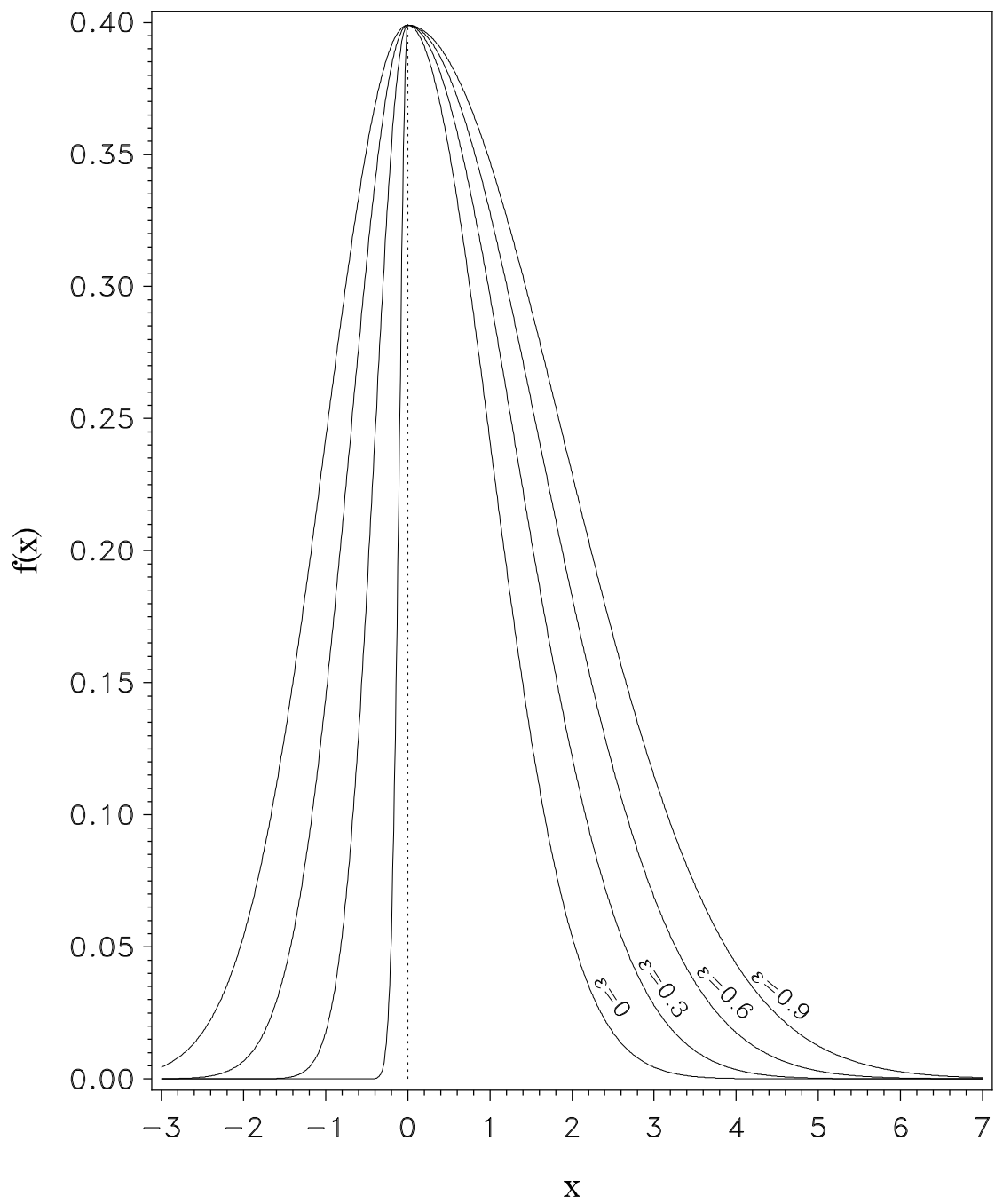
Figure 2: Histogram of IL-1β data. The first bin corresponds to the IL-1β values below the limit of detection.
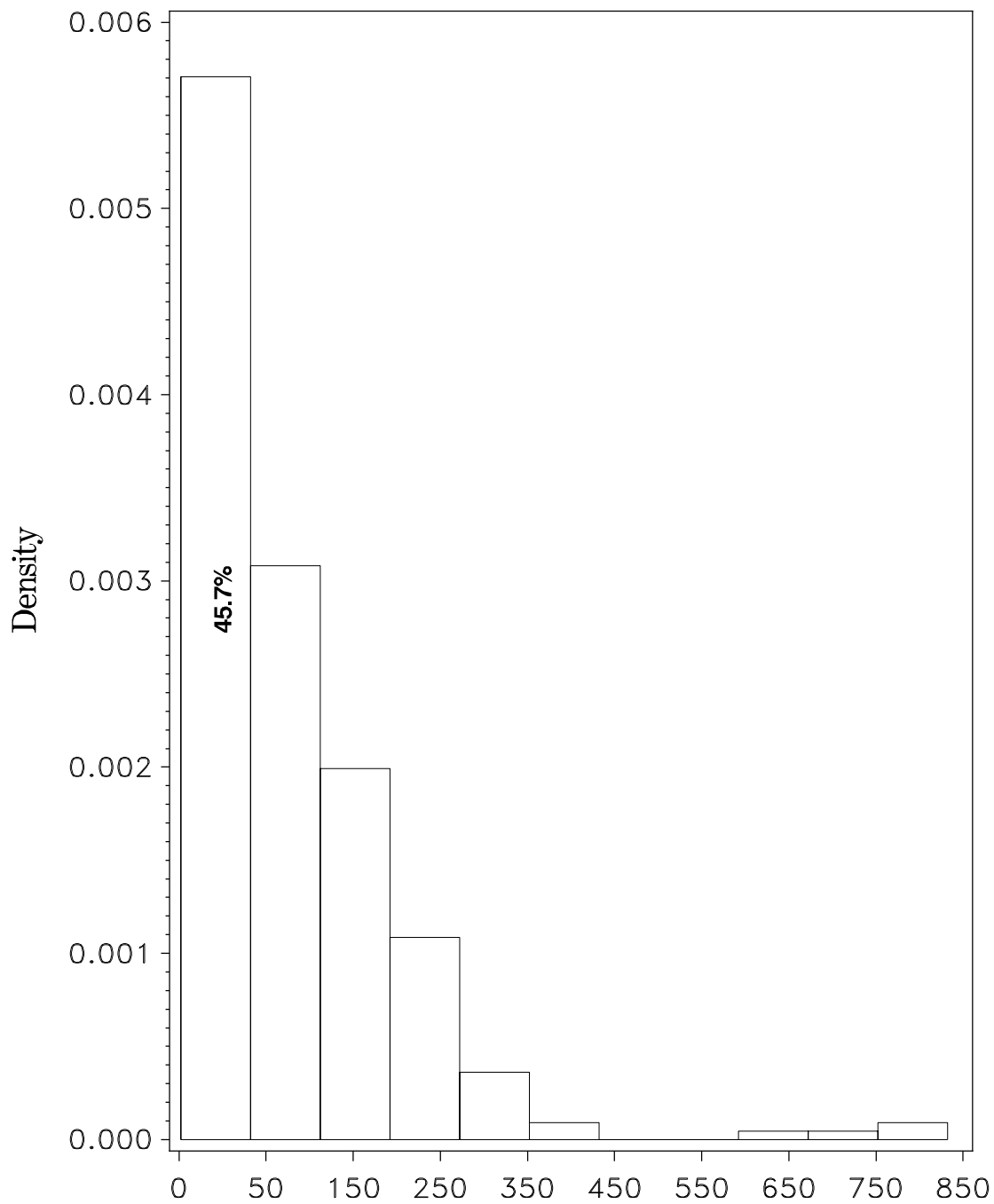
Figure 3: Histogram of MIP-2 data. The first bin corresponds to the MIP-2 values below the limit of detection.
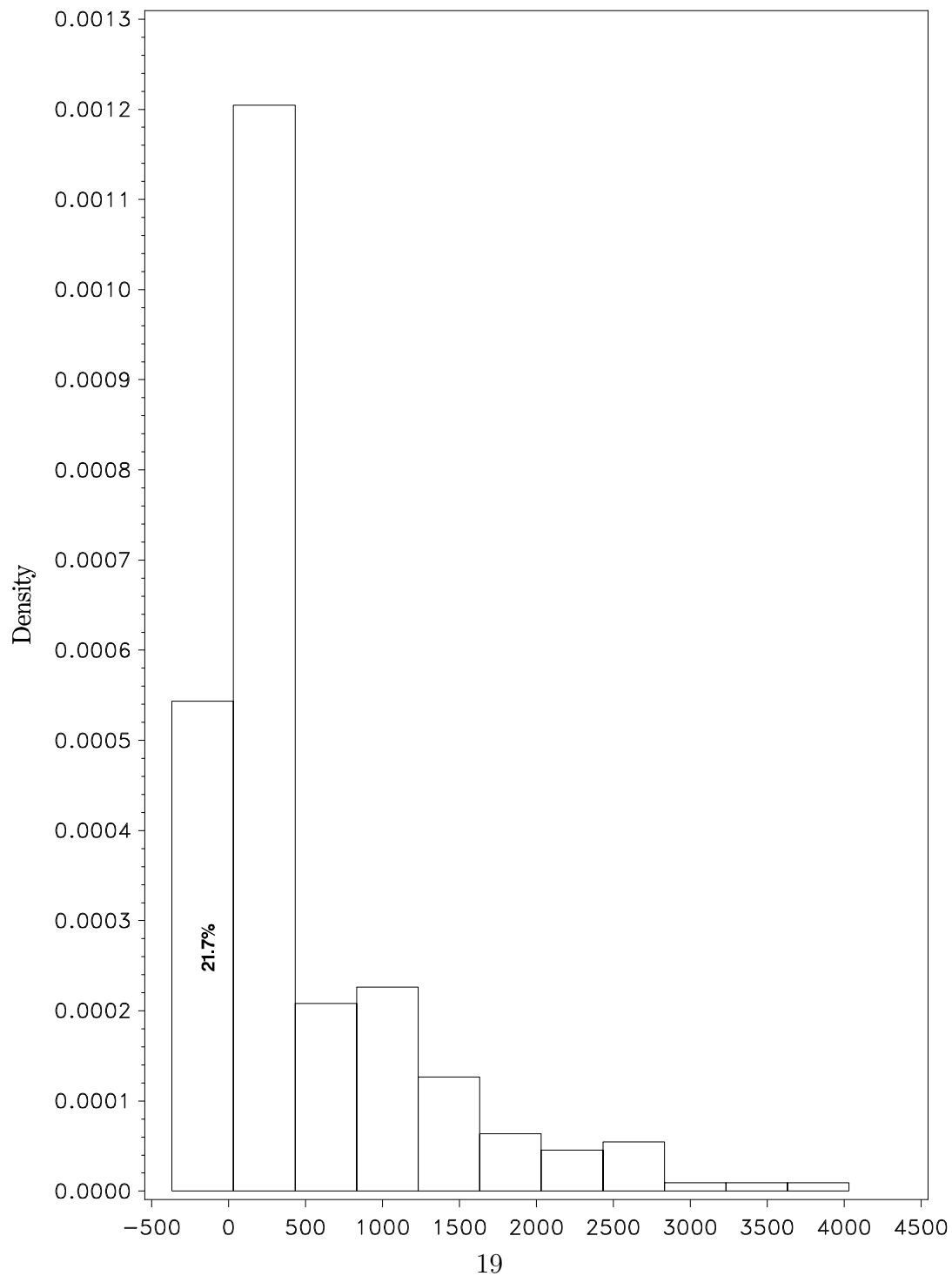
Table 1: $\hat{\mu} \pm \hat{\sigma}_{\hat{\mu}}$ and Estimated Coverage Probability for $\epsilon = 0$ and 0.25

ϵ	Censor %	μ	$n = 25$		$n = 50$		$n = 100$	
			Normal	ESN	Normal	ESN	Normal	ESN
0.00	20%	0.84	0.83 ± 0.21 95%	0.82 ± 0.23 96%	0.84 ± 0.15 95%	0.84 ± 0.16 96%	0.84 ± 0.10 95%	0.84 ± 0.11 96%
	30%	0.52	0.52 ± 0.22 96%	0.49 ± 0.27 96%	0.52 ± 0.15 95%	0.51 ± 0.18 96%	0.52 ± 0.11 95%	0.52 ± 0.12 96%
	40%	0.25	0.24 ± 0.23 96%	0.19 ± 0.33 95%	0.25 ± 0.17 95%	0.23 ± 0.23 95%	0.25 ± 0.11 95%	0.24 ± 0.16 94%
0.25	20%	0.87	0.83 ± 0.22 95%	0.84 ± 0.23 96%	0.84 ± 0.15 95%	0.86 ± 0.16 96%	0.83 ± 0.11 95%	0.86 ± 0.11 96%
	30%	0.59	0.53 ± 0.23 96%	0.54 ± 0.28 97%	0.54 ± 0.16 95%	0.56 ± 0.18 97%	0.54 ± 0.11 94%	0.58 ± 0.12 96%
	40%	0.34	0.26 ± 0.26 97%	0.24 ± 0.37 97%	0.26 ± 0.18 96%	0.27 ± 0.24 97%	0.26 ± 0.13 94%	0.30 ± 0.17 95%

Table 2: $\hat{\mu} \pm \hat{\sigma}_{\hat{\mu}}$ and Estimated Coverage Probability for $\epsilon = 0.5$ and 0.75

ϵ	Censor %	μ	$n = 25$		$n = 50$		$n = 100$	
			Normal	ESN	Normal	ESN	Normal	ESN
0.50	20%	0.92	0.86 ± 0.23 94%	0.89 ± 0.24 96%	0.86 ± 0.16 94%	0.91 ± 0.16 96%	0.86 ± 0.12 92%	0.92 ± 0.11 96%
	30%	0.67	0.57 ± 0.25 95%	0.60 ± 0.29 97%	0.57 ± 0.18 93%	0.62 ± 0.19 97%	0.57 ± 0.12 90%	0.64 ± 0.13 96%
	40%	0.42	0.27 ± 0.28 97%	0.27 ± 0.39 98%	0.28 ± 0.20 93%	0.33 ± 0.26 97%	0.28 ± 0.14 87%	0.36 ± 0.17 96%
0.75	20%	1.01	0.91 ± 0.26 93%	0.96 ± 0.27 95%	0.91 ± 0.18 92%	0.98 ± 0.18 95%	0.91 ± 0.13 89%	0.99 ± 0.12 95%
	30%	0.75	0.61 ± 0.27 94%	0.66 ± 0.32 97%	0.61 ± 0.19 92%	0.69 ± 0.21 97%	0.61 ± 0.14 85%	0.71 ± 0.14 96%
	40%	0.49	0.29 ± 0.31 96%	0.32 ± 0.42 98%	0.30 ± 0.21 91%	0.37 ± 0.28 97%	0.30 ± 0.15 81%	0.41 ± 0.19 96%

Table 3: $\hat{\beta}_1 \pm \hat{\sigma}_{\hat{\beta}_1}$ and Estimated Coverage Probability for $\epsilon = 0$ and 0.25

ϵ	Censor %	β_1	n = 25		n = 50		n = 100	
			Normal	ESN	Normal	ESN	Normal	ESN
0.00	20%	0.68	0.68 ± 0.41 94%	0.69 ± 0.48 90%	0.69 ± 0.29 95%	0.69 ± 0.31 92%	0.68 ± 0.21 95%	0.68 ± 0.21 94%
	30%	0.05	0.05 ± 0.42 95%	0.05 ± 0.49 91%	0.05 ± 0.29 95%	0.05 ± 0.32 93%	0.05 ± 0.21 95%	0.05 ± 0.22 94%
	40%	-0.49	-0.49 ± 0.45 95%	-0.51 ± 0.53 91%	-0.49 ± 0.31 95%	-0.50 ± 0.34 92%	-0.49 ± 0.22 95%	-0.50 ± 0.23 93%
0.25	20%	-0.07	-0.07 ± 0.43 95%	-0.07 ± 0.48 93%	-0.07 ± 0.31 95%	-0.07 ± 0.32 93%	-0.07 ± 0.22 95%	-0.07 ± 0.21 94%
	30%	-0.62	-0.64 ± 0.46 95%	-0.64 ± 0.52 91%	-0.65 ± 0.33 95%	-0.64 ± 0.34 92%	-0.66 ± 0.23 95%	-0.64 ± 0.23 94%
	40%	-1.13	-1.18 ± 0.53 95%	-1.19 ± 0.59 92%	-1.17 ± 0.36 95%	-1.16 ± 0.38 93%	-1.16 ± 0.25 95%	-1.14 ± 0.25 95%

Table 4: $\hat{\beta}_1 \pm \hat{\sigma}_{\hat{\beta}_1}$ and Estimated Coverage Probability for $\epsilon = 0.5$ and 0.75

ϵ	Censor %	β_1	$n = 25$		$n = 50$		$n = 100$	
			Normal	ESN	Normal	ESN	Normal	ESN
0.50	20%	-0.75	-0.82 ± 0.48 95%	-0.82 ± 0.51 93%	-0.82 ± 0.34 94%	-0.79 ± 0.33 93%	-0.82 ± 0.24 94%	-0.77 ± 0.22 94%
	30%	-1.25	-1.34 ± 0.53 95%	-1.35 ± 0.56 94%	-1.33 ± 0.37 94%	-1.29 ± 0.37 94%	-1.33 ± 0.26 94%	-1.27 ± 0.25 94%
	40%	-1.76	-1.87 ± 0.78 94%	-1.90 ± 0.87 94%	-1.82 ± 0.41 94%	-1.80 ± 0.43 94%	-1.83 ± 0.29 94%	-1.79 ± 0.30 94%
0.75	20%	-1.38	-1.50 ± 0.54 94%	-1.52 ± 0.53 96%	-1.51 ± 0.39 93%	-1.45 ± 0.37 95%	-1.51 ± 0.26 92%	-1.40 ± 0.25 96%
	30%	-1.89	-2.02 ± 0.64 94%	-2.07 ± 0.69 96%	-2.00 ± 0.42 94%	-1.97 ± 0.42 96%	-2.00 ± 0.29 92%	-1.92 ± 0.30 95%
	40%	-2.41	-2.57 ± 1.07 93%	-2.65 ± 1.20 95%	-2.47 ± 0.50 93%	-2.48 ± 0.55 95%	-2.46 ± 0.32 93%	-2.44 ± 0.36 95%

Table 5: Univariate Model IL-1β

Method	Parameter	Estimate	S.E.	p-value	95% CI
OLS	mean	97.94	6.73	$< .0001$	$(84.69, 111.19)$
Tobit	mean	34.68	12.47	0.0058	$(11.87, 59.22)$
ESN Tobit	mean	42.61	38.36	0.2675	$(-32.90, 118.12)$
	ϵ	0.1718	0.4302	0.6900	$(-0.68, 1)$

Table 6: Univariate Model MIP-2

Method	Parameter	Estimate	S.E.	p-value	95% CI
OLS	mean	466.13	41.59	$< .0001$	$(384.25, 548.01)$
Tobit	mean	369.36	43.11	$< .0001$	$(284.48, 454.23)$
ESN Tobit	mean	575.03	56.37	$< .0001$	$(464.06, 686.01)$
	ϵ	0.6714	0.1182	$< .0001$	$(0.44, .90)$

Table 7: Multivariate Regression IL-1β

Method	Parameter	Estimate	S.E.	p-value	95% CI
OLS	slope	113.80	23.26	0.0003	$(72.02, 163.59)$
	mouse type	0.98	13.41	0.9418	$(-25.42, 27.38)$
	time	-0.85	0.39	0.0288	$(-1.62, -0.09)$
Tobit	slope	71.53	39.65	0.3197	$(-6.53, 149.58)$
	mouse type	11.65	22.85	0.6105	$(-33.34, 56.64)$
	time	-2.23	0.68	0.0013	$(-3.57, -0.88)$
ESN Tobit	slope	80.53	50.14	0.1094	$(-18.7, 179.23)$
	mouse type	15.83	24.78	0.5234	$(-32.94, 64.60)$
	time	-2.46	0.76	0.0014	$(-3.96, -0.95)$
	ϵ	0.2075	0.42	0.6217	$(-.62, 1)$

Table 8: Multivariate Regression MIP-2

Method	Parameter	Estimate	S.E.	p-value	95% CI
OLS	slope	576.90	135.28	$< .0001$	$(310.57, 843.23)$
	mouse type	154.51	78.02	0.0487	$(0.92, 308.11)$
	time	-13.55	2.26	$< .0001$	$(-17.99, -9.11)$
Tobit	slope	512.69	138.74	0.0003	$(239.56, 753.83)$
	mouse type	173.86	80.58	0.0318	$(15.22, 332.50)$
	time	-15.98	2.34	$< .0001$	$(-20.58, -11.38)$
ESN Tobit	slope	787.49	146.36	$< .0001$	$(499.36, 1075.62)$
	mouse type	20.61	76.14	0.7869	$(-129.29, 170.50)$
	time	-11.45	2.51	$< .0001$	$(-16.40, -6.50)$
	ϵ	0.6796	0.09	$< .0001$	$(0.51, 0.85)$