ANNALS OF ECONOMICS AND FINANCE **4**, 125–136 (2003)

A Monte Carlo Comparison of Various Semiparametric Type-3 Tobit Estimators

Insik Min

Department of Economics, Texas A&M University
E-mail: i0m5376@neo.tamu.edu

Sheng jang Sheu

Department of Economics, University of Western Ontario

and

Zijun Wang

Private Enterprise Research Center, Texas A&M University

This paper compares recently developed semiparametric estimators of Type-3 Tobit model using Monte Carlo simulations. In particular, we examine the finite sample performance of the recently proposed method by Li and Wooldridge and compare it to some alternative semiparametric estimators. Simulation results indicated that Li and Wooldridge (2002) estimator under the independence restriction compares well relative to other alternative estimators, especially when the sample size is small or the error distribution has a thick tail.

Key Words: Type-3 Tobit model; Semiparametric estimation; Monte Carlo simulation.
JEL Classification Numbers: C10, C14, C15.

1. INTRODUCTION

This paper compares recently developed semiparametric estimators of Type-3 Tobit model using Monte Carlo simulations. Type-3 Tobit model is widely used in economics, for example we observe market wages only if an individual participates in the labor force. Because we do not observe the wage for non-participant in the labor market, samples for the wage equation are non-randomly selected. When samples are non-randomly chosen,

1529-7373/2003

estimation methods ingoring the sample selection bias may lead to inconsistent estimation results. Heckman (1979) described this problem as sample selection bias with which least square estimators are usually inconsistent.

Many examples with selectivity bias are analyzed in the applied economic fields such as modeling labor supply (Connelly 1992; Hill 1983; Riboud 1985; Gerfin 1996; Martins 2001), migration (Nakosteen and Zimmer 1980), credit scoring model (Greene 1998) and health insurance model (Van and Praag 1981), to mention only a few. Two types of sample selection models are frequently used in the empirical modelings: Type-2 and Type-3 Tobit models. In a Type-2 Tobit model, the dependent variable in the selection step is the binary variable (e.g., participant or non-participant), while in a Type-3 Tobit model the dependent variable is censored as a selection variable. It is well known that the latter uses more information about the selection variable than the former does.

In order to consistently estimate the model with possible selectivity bias, Heckman (1979) suggested a simple two-step estimation procedure in which the normal distribution of error terms is assumed. Moreover, Wooldridge (1994) proposed a two-step estimation procedure which can be generalized to non-normal error distributions. Wooldridge (1994) argued that his estimators are potentially more efficient than Heckman (1979)'s procedure especially when the X matrix is near multicollinearity. Many studies have been carried out to avoid the misspecification of distributional function and accordingly semiparametric estimation for the model with sample selection has been proposed with weaker restrictions than the normality of error terms. For a Type-3 Tobit model, Lee (1994), Chen (1997), Honore, Kyrizidou and Udry (1997, hereinafter HKU), and Li and Wooldridge (2002) suggested various semiparametric estimation methods with the independence restriction between regressors and error terms. In this paper, we use Monte Carlo experiment to compare the finite sample performances of four different Type-3 Tobit estimators.

This paper is organized as follows. In section 2, we briefly discuss the basic Type-3 Tobit model and various semiparametric estimators. Section 3 reports the Monte Carlo simulation results and compares the finite sample performance of each semiparameric estimator. Section 4 concludes the paper.

2. SEMIPARAMETRIC ESTIMATION OF TYPE-3 TOBIT MODEL

Formally, a Type-3 Tobit model consists of two equations: The first equation (the selection equation) is given by

$$y_i = \max\{x_i\beta + u_i,\ 0\}, \tag{1}$$

where y_i is known as the selection variable. This differs from a Type-2 Tobit model where one only observes the sign of y_i. In a Type-3 Tobit model, we observe y_i's actual value when $y_i > 0$. And the second equation (the main equation) is

$$w_i = \{z_i\gamma + v_i\} \cdot I(y_i > 0), \tag{2}$$

where $I(\cdot)$ is the usual indicator function which is equal to 1 when $y_i > 0$, 0 otherwise. An example of Type-3 Tobit model is that y_i is the hour worked by an individual and w_i is the wage of the same individual. We observe the wage only if the hour of work is positive.

Without the specific distribution assumption, the first step (selection) equation can be estimated by Censored Least Absolute Deviation (CLAD) method proposed by Powell (1984). Powell (1984) established the $\sqrt{n}$-consistency and the asymptotic normality for the CLAD estimator. Let $\hat{\beta}_{CLAD}$ denote the CLAD estimator of β. $\hat{\beta}_{CLAD}$ is obtained by minimizing the following objective function,

$$\hat{\beta}_{CLAD} = \arg\min_{\beta} \frac{1}{n} \sum_{i=1}^{n} |y_i - \max\{0, x_i\beta\}| \tag{3}$$

For the second step estimation, take the expectation of (2), conditioning on $y_i > 0$, gives

$$E(w_i|y_i > 0) = z_i\gamma + E(v_i|y_i > 0) = z_i\gamma + E(v_i|u_i > -x_i\beta) \tag{4}$$

Because $E(v_i|u_i > -x_i\beta)$ does not have zero mean, the least squares estimator of regressing w_i on z_i will lead to inconsistent estimation of γ. Under the joint normality assumption of (u_i, v_i), one can add a bias correction term which can be written as the inverse Mills ratio to restore a zero mean condition in (4). However, this estimator based on a parametric distribution assumption of the error terms is consistent only when the error distribution is correctly specified. As an alternative, various semiparametric approaches without assuming the underlying error distribution are proposed. In the following subsection, we briefly discuss some semiparametric estimators which are used in our simulation.

2.1. Lee's Estimator (Lee 1994)

Lee (1994) suggested using a nonparametric term to correct the sample selection bias in the main equation. He employed the 'index property' suggested by Ichimura and Lee (1991). For the identification of γ in (4), the index property that additional term generated from the non-random sample selection is a function of $x_i\beta$. The selection bias term can be rewritten as

$$E(v|u > -x\beta) = \frac{\int\limits_{-\infty}^{\infty}\int\limits_{x\beta}^{\infty}\int\limits_{-x\beta}^{\infty} v f(v,u)h(k)du\, dk\, dv}{\int\limits_{x\beta}^{\infty}\int\limits_{-x\beta}^{\infty} f_u(t)h(k)dt\, dk} \tag{5}$$

where $f(v,u)$ is a joint density of v and u, $f_u(\cdot)$ is a marginal density of u and $h(\cdot)$ is the density function of $x\beta$. The distribution of $x\beta$ should be absorbed in the truncated sample selection term. Nonparametric estimate for the bias correction term is obtained by

$$\hat{E}(v|u > -x_i\beta,\, x\beta > x_i\beta) = \frac{\int_{x_i\beta}^{\infty}\int_{-x_i\beta}^{\frac{1}{na_n^2}\sum_{j=1,j\neq i}^{n} v_j K\left(\frac{u_j-u_i}{a_1}\right)K\left(\frac{x_j\beta-x_i\beta}{a_2}\right)}}{\int_{x_i\beta}^{\infty}\int_{-x_i\beta}^{\frac{1}{na_n^2}\sum_{j=1,j\neq i}^{n} K\left(\frac{u_j-u_i}{a_1}\right)K\left(\frac{x_j\beta-x_i\beta}{a_2}\right)}} \tag{6}$$

where $(a_1,\, a_2)$ are smoothing parameters[1] and $K(\cdot)$ is a kernel function.[2] In practice, the conditions $x\beta > x_i\beta$ and $u > -x_i\beta$ can be reduced to $y = x\beta + u > 0$.

Subtracting (4) from (2), the regression model for the main equation is given by

$$w_i - \hat{E}(w|u > -x_i\beta,\, x\beta > x_i\beta) = \left[z - \hat{E}(z|x\beta > x_i\beta)\right]\cdot\gamma + \varepsilon_i \tag{7}$$

Because the error term (ε_i) in (7) satisfies the condition $E(\varepsilon_i|x_i) = 0$, γ can be consistently estimated by the least squares procedure. Lee (1994) also showed that his suggested semiparametric estimator has an asymptotic normal distribution.

2.2. Chen's Estimator (Chen 1997)

Chen (1997) proposed an alternative semiparametric estimator under the independence assumption between error terms and regressors. With consistently estimated $\hat{\beta}_{CLAD}$ in the first step, a bias correction term is converted to a constant term through trimming the sample in the second step. By switching the bias term into a constant, the net effect on the subsample of the selection is only to shift the intercept term in the main equation without affecting the slope coefficients.

---

[1] Lee (1994) showed in the simulation that his estimators are not very sensitive to the choice of the smoothing parameter.

[2] In this article, the standard normal density function is used as the kernel function, while Lee (1994) adopted the following density function: $K(t) = (15/16)(1-t^2)^2$ if $|t| < 1$, $K(t) = 0$ otherwise.

Chen (1997) suggested a constant censoring point instead of $-x\beta$ through trimming the original sample so that the conditional expectation $E(v_i|u_i > -x_i\beta)$ becomes zero in the trimmed sample. In practice, he set up K different subsamples in which $(u \in (c_{k-1}, c_k),\, x_i\beta > -c_{k-1})$ holds in the kth subsample, and $c_0 < c_1 < \cdots < c_K$. Define $\alpha_k = E(v|c_k > u > c_{k-1},\, x\beta > -c_{k-1},\, y > 0)$ as the conditional mean of v in the kth subsammple. As a consequence, α_k becomes a constant under the assumption that regressors and error terms are independent. Hence, equation (4) can be rewritten as

$$E(w|c_k > u > c_{k-1},\, x\beta > -c_{k-1},\, y > 0) = z\gamma + \alpha_k \tag{8}$$

Chen's estimator for γ is constructed by pooling K subsamples:

$$\begin{aligned}\hat{\gamma}_{p1} \;=\; & \min_{\gamma,\alpha_1,\alpha_2,\cdots\alpha_K} \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K} I(c_k > u_i > c_{k-1},\, x_i\beta > -c_{k-1},\, y_i > 0) \\ \times\; & (w_i - z_i\gamma - \alpha_k)^2 \end{aligned} \tag{9}$$

where $I(\cdot)$ is an indicator trimming function and $\hat{\gamma}_{p1}$ denotes a least-squares-type estimator. Chen has shown that this estimator is consistent and asymptotically normal under some regularity conditions.

2.3. HKU's Estimator (Honore, Kyriazidou and Udry 1997)

HKU (1997) considered two semiparmetric estimators for Type-3 Tobit model: One is under the assumption of conditional symmetry, the other is under the assumption of independence between error terms and regressors to restore a zero conditional mean. In this paper, we focus on the second one based on the independence condition to keep same assumption with other estimators. Under this assumption, a pairwise comparison approach is employed to correct the sample selection bias in the main equation. Pairwise difference approach is based on the fact that the difference of independently and identically distributed random variables is distributed around zero. Consider the following two pairwise equations:

$$E(w_i|u_i > -x_i\beta) = z_i\gamma + E(v_i|u_i > -x_i\beta) \tag{10}$$

$$E(w_j|u_j > -x_j\beta) = z_j\gamma + E(v_j|u_j > -x_j\beta) \tag{11}$$

Subtracting (11) from (10), the pairwise difference regression model is

$$\begin{aligned}&E(w_i|u_i > -x_i\beta) - E(w_j|u_j > -x_j\beta) \\ &= (z_i - z_j)\gamma + [E(v_i|u_i > -x_i\beta) - E(v_j|u_j > -x_j\beta)]\end{aligned} \tag{12}$$

If the conditional mean of the error terms inside the bracket of (12) is zero, then (12) can be estimated by the least squares method. In order to restore a zero conditional mean, a truncated point should be used which gives the same bias for v_i and v_j. They suggested the following trimmings: $u_i > \max\{-x_i\beta, -x_j\beta\}$ and $u_j > \max\{-x_i\beta, -x_j\beta\}$. Therefore, by replacing the pairwise trimming, $(v_i - v_j)$ is distributed symmetrically around zero.

Using the pairwise difference approach, HKU estimator based on the main equation is

$$\begin{aligned}\hat{\gamma}_{h2} &= \min_{\gamma} \sum_{i<j} I\left[y_i > \max\{0, (x_i - x_j)\beta\},\, y_j > \max\{0, (x_i - x_j)\beta\}\right] \\ &\times \{w_i - w_j - (z_i - z_j)\gamma\}^2 \qquad (13)\end{aligned}$$

where $I(\cdot)$ is the usual indicator function.

2.4. Li-Wooldridge's Estimator

Wooldridge (1994) proposed an alternative parametric two-step estimation procedure for a sample selection model. He argued that his estimator may be more efficient than the one suggested by Heckman (1979). Moreover, Wooldridge (1994) suggested that his method can be generalized to the non-normal error case. Li and Wooldridge (2002) proposed a new semiparametric estimator in the second step under the assumption that error terms are independent with regressors. Using the data with $y_i > 0$, equation (4) can be written as

$$E(w_i|x_i, u_i, y_i > 0) = z_i\gamma + E(v_i|x_i, u_i, y_i > 0) = z_i\gamma + E(v_i|u_i) = z_i\gamma + g(u_i) \qquad (14)$$

where the second equality used the fact that $E(v_i|x_i, u_i, y_i > 0) = E(v_i|u_i)$ by the independence assumption. $g(u_i) = E(v_i|u_i)$ is an unknown function. Li and Wooldridge showed that γ can be consistently estimated from the following partially linear model, using observations with $y_i > 0$,

$$w_i = z_i\gamma + g(u_i) + \varepsilon_i \qquad (15)$$

Robinson (1988) considered the consistent estimation of γ in the case u_i is observable. When u_i is not observable but can be consistently estimated (e.g., generated regressor), Li and Wooldridge (2002) established the $\sqrt{n}$-consistency and asymptotic normal distribution of $\hat{\gamma}$ based on estimating (15) with u_i being a generated regressor.

3. SIMULATION RESULTS

In this section, we compare the finite sample performances of the four different Type-3 Tobit estimators discussed in section 2. In order to compare with the results of Chen (1997), the data generating process in this paper is similar to the ones used in Chen (1997).

The selection equation of the Type-3 Tobit model in our experiment is

$$y_i = \max\{\beta_1 x_{1i} + \beta_2 x_{2i} + u_i,\, 0\} \tag{16}$$

And the main equation is given by

$$w_i = \{\gamma_1 z_{1i} + \gamma_2 z_{2i} + v_i\} \cdot I(y_i > 0) \tag{17}$$

The true parameters are the same as in Chen (1997), which are $\{\beta_1, \beta_2\} = \{1, 1\}$ and $\{\gamma_1, \gamma_2\} = \{1, 2\}$. The independent variables, $x_1 = z_1$ and $x_2 = z_2$ are designed for both equations. x_1 and x_2 follow a Normal (0,1) distribution and a Uniform (-2,2) distribution, respectively. The error term u_i in the selection equation follows three different distributions: (i) u_i is drawn from a Normal (0,1) distribution, (ii) u_i is drawn as a mixed gamma and normal, which is ($\sqrt{0.8}$·Standardized $\chi^2_{(8)}$ + $\sqrt{0.2}$·Normal (0,1)),[3] and (iii) u_i follows a Cauchy (0,1) distribution that has symmetric and much heavier tails than the normal distribution. The error term v_i in the main equation is constructed by mixing u_i and a normal distribution: $v_i = \sqrt{0.5} \cdot u_i + \sqrt{0.5} \cdot Normal(0, 1)$. For $\hat{\gamma}_{p1}$ of Chen (1997), the censored intervals are chosen with K=20 with $c_0 = -4$ and $c_K = 4$. We do not trim the data set when using $\hat{\gamma}_l$ of Lee (1994). We conduct 2,000 replications with the sample size equal to 50, 100, 200 and 500. We select three different error term distributions. Because similar simulation results are verified for the coefficient $\hat{\gamma}_2$ on z_{2i} in (17), only results for the coefficient $\hat{\gamma}_1$ on z_{1i} are reported. The tables 1∼3 present the mean, standard deviation, and Root Mean Squared Error (RMSE) of the estimates in each case.

The finite sample performance of RMSE is mixed for various cases. In Table 1, Li and Wooldridge's estimator performs better than other estimators in the Normal distribution case, especially when the sample size is small ($n = 50$ and 100). However, as the sample size increases, the RMSE of HKU (n=200) and Chen (n=500) estimators tend to be smaller than that of Li-Wooldridge's method. Also note that the bias and variance of four estimators are significantly improved as the sample size increases. In particular, when the sample size is 500, most seimparametric methods (except for Lee) sharply estimate the true parameter. For Li-Wooldridge's estimator, we use two different methods to choose the smoothing parameter; one is the fixed bandwidth selection (denoted by LW1 and LW2 in Table 1) where the smoothing parameter is chosen via $h = c \cdot sd(\hat{u}_i) \cdot n^{-1/5}$, and c=1 (LW1) and c=2 (LW2) are selected for the constant c. The other is

[3] Standardized $\chi^2_{(8)}$ has the same density function with the gamma (0,1) distribution in Chen (1997).

TABLE 1.

$\mu_i \sim N(0,1)$

Sample size	Estimator	True value	Mean	Standard deviation	RMSE
N=50	Lee	1.000	0.9780	0.3072	0.3079
	Chen	1.000	0.9793	0.3234	0.3239
	HKU	1.000	0.9696	0.2756	0.2772
	LW1	1.000	0.9703	0.2791	0.2805
	LW2	1.000	0.9343	0.2613	0.2693
	LWCV	1.000	0.9682	0.2824	0.284
$N = 100$	Lee	1.000	1.0287	0.2222	0.2240
	Chen	1.000	0.996	0.2068	0.2068
	HKU	1.000	0.9826	0.1921	0.1928
	LW1	1.000	0.9900	0.1956	0.1958
	LW2	1.000	0.9612	0.1843	0.1882
	LWCV	1.000	0.9885	0.1983	0.1986
$N = 200$	Lee	1.000	1.0575	0.1682	0.1777
	Chen	1.000	0.9960	0.1327	0.1327
	HKU	1.000	0.9912	0.1309	0.1311
	LW1	1.000	0.9918	0.1373	0.1375
	LW2	1.000	0.9735	0.1306	0.1332
	LWCV	1.000	0.9918	0.1372	0.1374
$N = 500$	Lee	1.000	1.1072	0.1160	0.1579
	Chen	1.000	1.0036	0.0829	0.0829
	HKU	1.000	1.0014	0.0837	0.0837
	LW1	1.000	0.9999	0.0857	0.0857
	LW2	1.000	0.9895	0.083	0.0837
	LWCV	1.000	10000	0.086	0.0859

1. Because similar simulation results are found for $\hat{\gamma}_2$, we report only results for $\hat{\gamma}_1$in equation (17).
2. LW1 and LW2 have the fixed bandwidth with c=1 and c=2, respectively. LWCV has the bandwidth obtained from cross validation method.

the local constant cross validation method (denoted by LWCV) which leads to the optimal bandwidth selection. Even though LW2 performs slightly better than LW1 and LWCV in the normal error case, overall the RMSE is very similar with respect to other choices of h.[4] This finding also holds true for the case of the mixed gamma and normal distribution.

Table 2 reports simulation results of four semiparametric estimators when u_i is the mixed gamma and normal distribution, and thus is an asym-

[4]Christofides et al. (2003) pointed out that the semiparametric estimator of γ depends on the average of the nonparametric estimates which is less sensitive to different values of the smoothing parameter.

TABLE 2.

$u_i \sim$ Mixed Gamma and Normal distribution

Sample size	Estimator	True value	Mean	Standard deviation	RMSE
$N = 50$	Lee	1.000	0.9632	0.3070	0.3090
	Chen	1.000	0.9754	0.3248	0.3255
	HKU	1.000	0.9696	0.2814	0.2829
	LW1	1.000	0.9619	0.2804	0.2828
	LW2	1.000	0.9181	0.2571	0.2697
	LWCV	1.000	0.9555	0.2825	0.2859
$N = 100$	Lee	1.000	0.9758	0.2222	0.2234
	Chen	1.000	0.961	0.1966	0.2003
	HKU	1.000	0.9603	0.2032	0.2070
	LW1	1.000	0.9534	0.1916	0.1971
	LW2	1.000	0.9203	0.1810	0.1977
	LWCV	1.000	0.9536	0.1928	0.1982
$N = 200$	Lee	1.000	1.0171	0.1635	0.1643
	Chen	1.000	0.9681	0.1353	0.139
	HKU	1.000	0.9791	0.1418	0.1433
	LW1	1.000	0.9622	0.1377	0.1428
	LW2	1.000	0.9388	0.1320	0.1454
	LWCV	1.000	0.963	0.1386	0.1434
$N = 500$	Lee	1.000	1.0504	01133	0.1239
	Chen	1.000	0.9592	0.0812	0.0908
	HKU	1.000	0.9724	0.0889	0.0930
	LW1	1.000	0.9587	0.0855	0.095
	LW2	1.000	0.9441	0.0829	0.100
	LWCV	1.000	0.9591	0.0858	0.095

1. Because similar simulation results are found for $\hat{\gamma}_2$, we report only results for $\hat{\gamma}_1$in equation (17).
2. LW1 and LW2 have the fixed bandwidth with c=1 and c=2, respectively. LWCV has the bandwidth obtained from cross validation method.

metric distribution. Similarly to the results of Table 1, Li-Wooldridge's estimator slightly dominates other semiparametric estimators when the sample size is 50 and 100. It is noted that Chen's estimator performs better than other estimators including Li-Woodridge method when the sample size is large. However, as a whole the finite sample performances of semiparametric estimators are slightly deteriorating when compared with the normally distributed error case in that the estimators reveal the larger variance and bias as reported in Table 2 . The asymmetric error distribution seems to affect the precise parameter estimate of a Type-3 Tobit model.

TABLE 3.

$u_i \sim Cauchy\,(0, 1)$

Sample size	Estimator	True value	Mean	Standard deviation	RMSE
$N = 50$	Lee	1.000	0.8459	0.7403	0.7558
	Chen	1.000	0.9879	0.7208	0.7206
	LW1	1.000	0.8213	0.6483	0.6678
	LW2	1.000	0.9880	0.2724	0.2726
	LWCV	1.000	0.8430	0.6503	0.6687
$N = 100$	Lee	1.000	0.8778	0.5356	0.5491
	Chen	1.000	0.9614	0.4951	0.4963
	LW1	1.000	0.8050	0.4576	0.4972
	LW2	1.000	1.0012	0.1721	0.1721
	LWCV	1.000	0.8317	0.4913	0.5191
$N = 200$	Lee	1.000	0.8850	0.5275	0.5396
	Chen	1.000	0.9560	0.4142	0.4163
	LW1	1.000	0.7780	0.3420	0.4076
	LW2	1.000	1.0121	0.1281	0.1286
	LWCV	1.000	0.7907	0.4001	0.4513
$N = 500$	Lee	1.000	0.9157	0.3512	0.3610
	Chen	1.000	0.9897	0.3631	0.3631
	LW1	1.000	0.7544	0.2858	0.3767
	LW2	1.000	1.0262	0.0816	0.0857
	LWCV	1.000	0.7441	0.3254	0.4139

1. Because similar simulation results are found for $\hat{\gamma}_2$, we report only results for $\hat{\gamma}_1$in equation (17).
2. LW1 and LW2 have the fixed bandwidth with c=1 and c=2, respectively. LWCV has the bandwidth obtained from cross validation method.

Table 3 presents simulation results with a Cauchy error distribution that has no finite moments of any order (a thick tail distribution). In the case of the Cauchy errors, McDonald and Xu (1996) examined that Powell's CLAD estimator which is employed in our analysis dominates other estimators such as Tobit and Semiprametric Maximum Likelihood (SP-ML) estimators. Hence we expect that there is no serious bias in estimating the parameters in the first step even with a Cauchy error distribution. Under this design, LW2 estimator outperforms the other alternatives either the sample size is large or small. Since the HKU estimator is quite distorted in this case, we do not report the RMSE for HKU estimator. As a consequence, the thick tail (no finite moments of any order) critically affects the performance of the pairwise difference approach. Compared to the two previous experiments, the bias and variance estimated from the four estimators became relatively larger in the Cauchy distribution case.

In summary, Li and Wooldridge's estimator based on the semiparametric partially linear model compares favorably with other existing methods, and especially when the sample size is small.

4. CONCLUSION

We have investigated and compared the finite sample performances of the four different semiparametric estimators recently developed for a Type-3 Tobit model. In general, the Monte Carlo simulation results suggest that all the semiparametric methods lead to similar results. Specifically we found that the asymmetry or thick tails (no finite moments) error distribution moderately affect the bias and variance of semiparamtric estimator. Simulation results also indicated that Li and Wooldridge's estimator under the independence restriction performs relatively well compared with other alternative approaches, especially when the sample size is small and the error has a thick tail distribution. Our simulations reinforce the fact that semiparametric estimators of Type-3 Tobit models can be widely employed in empirical research without the standard assumption of normal error distributions.

REFERENCES

Chen, S., 1997, Semiparametric esitmation of the type-3 tobit model. *Journal of Econometrics* **80**, 1-34.

Christofides, L. N., Q. Li, Z. Liu, and I. Min., 2003, Recent two-stage sample selection procedures with an application to the gender wage gap. *Journal of Business and Economic Statistics*. Forthcoming.

Connelly, R., 1992, The effect of child care costs on married women's labor force participation. *The Review of Economics and Statistics* **74**, 83-90.

Gerfin, M., 1996, Parametric and semiparametric estimation of the binary response model of labor market participation. *Journal of Applied Econometrics* **11**, 321-339.

Greene, H. W., 1998, Sample selection in credit-scoring models. *Japan and the World Economy* **10**, 299-316.

Heckman, J., 1979, Sample selection as a specification error. *Econometrica* **47**, 153-161.

Hill, M. A., 1983, Female labor force participation in developing and developed countries – Consideration of the informal sector. *The Review of Economics and Statsitics* **65**, 459-468.

Honore, B. E., E. Kyriazidou, and C. Udry, 1997, Estimation of type-3 tobit models using symmetric trimming and pairwise comparisons. *Journal of Econometrics* **76**, 104-128.

Ichimura, H. and L. F. Lee., 1991, Semiparametric least squares estimation of multiple index models. In: W. A. Barnett, J. L. Powell and G. Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics*: Cambridge Univeristy Press.

Lee, L., 1994, Semiparametric two-stage estimation of sample selection models subject to tobit-type selection rules. *Journal of Econometrics* **61**, 305-344.

Li, Q. and J. M. Wooldridge, 2002, Semiparametric estimation of partially linear models for dependent data with generated regressors. *Econometric Theroy* **18**, 625-645.

Martins, M. F., 2001, Parametric and semiparametric estimation of sample selection models: An empirical application to the female labor force in Portugal. *Journal of Applied Econometrics* **16**, 23-29.

McDonald, J. B. and Y. J. Xu, 1996, A comparison of semi-parametric and partially adaptive estimators of the censored regression model with possibly skewed and leptokurtic error distributions. *Economics Letter* **51**, 153-159.

Nakosteen, R. A. and M. Zimmer, 1980, Migration and income: The question of self-selection. *Southern Economic Journal* **46**, 840-851.

Powell, J. L., 1984, Least sbsolute deviations estimation for the censored regression model. *Journal of Econometrics* **25**, 303-325.

Riboud, M., 1985, An analysis of women's labor force participation in France: Cross-section estimates and time-series evidence. *Journal of Labor Economics* **3**, S177-S200.

Robinson, P., 1988, Root-N consistent semiparametric regression. *Econometrica* **56**, 931-954.

Van de Ven, W. P. and B. M. Van Praag, 1981, The demand for deductibles in private health insurance. *Journal of Econometrics* **17**, 229-252.

Wooldridge, J. M., 1994, A simple specification test for the predictive ability of transformation models. *The Review of Economics and Stastistics* **76**, 59-65.