Econometrics Journal (2007), volume **10**, pp. 426–438.
doi: 10.1111/j.1368-423X.2007.00214.x

Two-stage estimation of limited dependent variable models with errors-in-variables

LIQUN WANG AND CHENG HSIAO

Department of Statistics, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2
E-mail: liqun_wang@umanitoba.ca

Department of Economics, University of Southern California, Los Angeles, CA 90089-0253, USA

First version received: May 2006; final version accepted: February 2007

Summary This paper deals with censored or truncated regression models where the explanatory variables are measured with additive errors. We propose a two-stage estimation procedure that combines the instrumental variable method and the minimum distance estimation. This approach produces consistent and asymptotically normally distributed estimators for model parameters. When the predictor and instrumental variables are normally distributed, we also propose a maximum likelihood based estimator and a two-stage moment estimator. Simulation studies show that all proposed estimators perform satisfactorily for relatively small samples and relatively high degree of censoring. In addition, the maximum likelihood based estimators are fairly robust against non-normal and /or heteroskedastic random errors in our simulations. The method can be generalized to panel data models.

Key words: *Censored regression, Errors in variables, Instrumental variables, Limited dependent variables, Measurement errors, Minimum distance estimation, Tobit model, Truncated outcome.*

1. INTRODUCTION

Censored or truncated regression models are widely used in econometrics as well as in biomedical, health and many other scientific fields. They are special cases of more general limited dependent variable models. Econometric applications of Tobit models include Heckman and MaCurdy (1986) and Killingsworth and Heckman (1986), among others (see also Amemiya 1984, 1985a and Maddala 1985). Another important issue that arises in econometrics and many other applied fields is the problem of errors-in-variables (Hsiao 1983, 1989; Fuller 1987; Carroll *et al.* 1995; Gustafson 2004). The combined problem of censored data and measurement errors arises also in more recent bioinformatics (Dood *et al.* 2004).

The censored regression or other limited dependent variable models with measurement errors have been considered by several authors. Hsiao (1991) studied a class of binary choice models where the explanatory variables are measured with errors. Weiss (1993) investigated the least absolute deviation estimators of a censored linear errors-in-variables model when certain instrumental variables are available. Wang (1998, 2002) derived consistent moment estimators and maximum likelihood estimators for a class of limited dependent variable models under the assumption of known reliability ratio which is equivalent to known measurement error variance. While this assumption is satisfied in many cases in natural sciences where validation or repeated measurement data are available, many data sets in economics and social sciences do not appear to possess this knowledge. An alternative approach to dealing with measurement error problems in regression models is to use instrumental variables (e.g. Amemiya 1985b, 1990; Weiss 1993; Wang and Hsiao 1995). When the response variables are censored or truncated, their conditional expectations are no longer linear in the predictor variables or parameters. Hence, the censored or truncated regression models are non-linear models. Recently, Schennach (2004) studied a general non-linear errors-in-variables model and showed that the model can be identified and consistently estimated using instrumental variables. However, in its general form the proposed estimation method, while based on Fourier transforms and generalized function theory, is quite technical and computationally intensive.

In this paper, we exploit the specific nature of non-linearity due to censoring and propose estimators which are easy to implement and possess fairly good finite sample properties. The results demonstrate that the existence of instrumental variables can provide consistent estimators of the censored or truncated linear models with additive measurement errors. Our procedure consists of two stages. The first stage involves substituting the instrumental variable in lieu of the variables measured with errors and obtain consistent estimators for the transformed model. The second stage retrieves the parameters of interest by the minimum distance method. The numerical computation is easy and straightforward. When the predictor and instrumental variables are normally distributed, we also propose a maximum likelihood and a two-step moment estimator. While the maximum likelihood estimator is inconsistent when normality assumption is violated, the two-stage minimum distance estimation method remains consistent and asymptotically normally distributed. Simulation studies demonstrate that all proposed estimators perform satisfactorily even for relatively small samples and relatively high degree of censoring in the observations of the response variable. In addition, the maximum likelihood based estimator is fairly robust against some non-normal but symmetric random error distributions.

This paper is organized as follows. Section 2 introduces the censored linear regression model with errors in variables and instrumental variables. Section 3 describes the two-stage procedure for consistent estimation of model parameters. A maximum likelihood and a method of moment-based estimator are proposed in Sections 4 and 5, respectively. The finite sample behaviour of the proposed estimators is investigated through simulation studies in Section 6. Conclusions and discussion are given in Section 7.

2. THE MODEL

Consider a linear latent relationship

$$\eta = \alpha_1 + \alpha_2' \xi + \varepsilon, \tag{1}$$

where $\eta \in I\!R$ and $\xi \in I\!R^k$ are latent response and predictor variables, respectively, α_1 and α_2 are unknown parameters and ε is a random error. The observed predictor variable is

$$x = \xi + \delta, \tag{2}$$

where δ is the random measurement error. We assume that ξ, δ and ε are mutually uncorrelated and have means μ_ξ, 0, 0 and covariances Σ_ξ, Σ_δ, σ_ε^2, respectively, where Σ_ξ has full rank but Σ_δ can be singular to allow some components of ξ be measured without error. In addition, suppose that an instrumental variable (IV) $z \in \mathbb{R}^\ell$ is available which is correlated with ξ but uncorrelated with δ and ε. Then, it follows from (2) that $\Sigma_{z\xi} = E(z - \mu_z)(\xi - \mu_\xi)' = \Sigma_{zx}$ and $\mu_\xi = \mu_x$. Throughout the paper we assume that the $\ell \times k$ matrix Σ_{zx} has full column rank k. It will be clear later that this assumption ensures the identifiability and consistent estimation of α_1 and α_2 in the model. This assumption implies $\ell \geq k$ and is easy to check once the data are available.

If there were no censoring, then $\eta = y$ is fully observed and (1) and (2) give a conventional linear errors-in-variables model (e.g. Aigner *et al.* 1984)

$$y = \alpha_1 + \alpha_2' x + \varepsilon - \alpha_2' \delta.$$

It is well known that the ordinary least-squares estimators for α_1 and α_2 based on (x, y) will be inconsistent, because x is correlated with the error term $\varepsilon - \alpha_2'\delta$. Instead, the usual IV estimation procedure based on x, y and z will yield consistent estimators.

However, in the censored model the observed response variable is

$$y = \eta \mathbf{1}(\eta > 0), \tag{3}$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. Substituting x for ξ into (3) yields

$$y = \alpha_1 + \alpha_2' x + \varepsilon^*, \tag{4}$$

where $\varepsilon^* = \varepsilon - \alpha_2'\delta - \eta\mathbf{1}(\eta \leq 0)$. Applying IV approach to (4) directly will not lead to consistent estimators of α_1 and α_2, because z is correlated with ε^* through η.

In this paper, our goal is to find consistent estimators for parameters α_1, α_2 and σ_ε^2 in model (1)–(3). The observed data are (x_i, z_i, y_i), $i = 1, 2, \ldots, n$, which are supposed to be independent but not necessarily identically distributed. Though we treat the censored model explicitly, it is easy to see that our method applies also to truncated model as well as other limited dependent variable models.

3. TWO-STAGE ESTIMATION

Although applying IV method directly will not yield consistent estimators of α_1 and α_2 if η is censored, there exists an indirect method that can yield consistent estimators. First, since $\Sigma_{z\xi} \neq 0$, let $\beta_2 = \Sigma_z^{-1}\Sigma_{z\xi}$, $\beta_1 = \mu_\xi - \beta_2'\mu_z$ and $\tau = \xi - \beta_1 - \beta_2' z$. Then we have

$$\xi = \beta_1 + \beta_2' z + \tau, \tag{5}$$

where τ is uncorrelated with δ, ε and satisfies $E(z\tau') = 0$ by construction. Further, because z is uncorrelated with τ and δ, substituting (5) into (2) results in a standard linear regression equation

$$x = \beta_1 + \beta_2' z + \tau + \delta. \tag{6}$$

It follows that β_1 and β_2 can be consistently estimated using the least-squares method and the sample moments of (x_i, z_i), as $\hat\beta_2 = \hat\Sigma_z^{-1}\hat\Sigma_{zx}$ and $\hat\beta_1 = \bar x - \hat\beta_2' \bar z$. On the other hand, substituting (5) into (1) we obtain

$$\eta = \gamma_1 + \gamma_2' z + u, \tag{7}$$

where

$$\gamma_1 = \alpha_1 + \alpha_2'\beta_1, \tag{8}$$

$$\gamma_2 = \beta_2\alpha_2 \tag{9}$$

and $u = \varepsilon + \alpha_2'\tau$ is uncorrelated with z. Hence, (7) together with (3) forms an error-free Tobit model and consequently γ_1 and γ_2 can be consistently estimated using the data (z_i, y_i). More specifically, if ε and τ are normally distributed, then u is also normal and consistent and efficient estimators of γ_1 and γ_2 can be obtained by the maximum likelihood method (e.g. Amemiya 1973; Wang 1998). In this case, all uncorrelatedness among various variables assumed before becomes independence. On the other hand, if the distribution of u is not normal but symmetric about zero, then consistent and asymptotically normally distributed estimators of these parameters can still be obtained by applying the symmetrically trimmed least-squares estimator of Powell (1986a). More generally, under conditional median or quantile restrictions for the distribution of u, the least absolute deviations (LAD) estimator of Powell (1984) or the regression quantiles estimator of Powell (1986b) give consistent estimators for γ_1 and γ_2 (see also Khan and Powell (2001) for other semiparametric methods). Note that the later two procedures do not require u to be uncorrelated with z and apply to heteroskedastic data (z_i, y_i) as well.

Now, let $\gamma = (\gamma_1, \gamma_2')'$, $\alpha = (\alpha_1, \alpha_2')'$ and

$$B = \begin{pmatrix} 1 & \beta_1' \\ 0 & \beta_2 \end{pmatrix}.$$

Then (8) and (9) can be written jointly as

$$\gamma = B\alpha. \tag{10}$$

Note that (10) contains $\ell + 1$ equations. Given γ and B, a necessary and sufficient condition for the unique solution of α is $\text{rank}(B) = k + 1$ (e.g. Hsiao 1983), which holds if and only if

$$\text{rank}(\beta_2) = k. \tag{11}$$

Since $\beta_2 = \Sigma_z^{-1}\Sigma_{zx}$, (11) is equivalent to rank $(\Sigma_{zx}) = k$. Therefore, if consistent estimators of γ and B exist, we can obtain consistent estimator of α by solving (10) provided the rank condition for identification holds. To simplify notation, denote

$$X = \begin{pmatrix} 1 & x_1' \\ 1 & x_2' \\ \vdots & \vdots \\ 1 & x_n' \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & z_1' \\ 1 & z_2' \\ \vdots & \vdots \\ 1 & z_n' \end{pmatrix}, \quad U = \begin{pmatrix} 0 & \tau_1' + \delta_1' \\ 0 & \tau_2' + \delta_2' \\ \vdots & \vdots \\ 0 & \tau_n' + \delta_n' \end{pmatrix}.$$

Then (6) can be written as $X = ZB + U$ and, therefore, the LSE for B is given by

$$\hat{B} = (Z'Z)^{-1}Z'X = \begin{pmatrix} 1 & \hat{\beta}_1' \\ 0 & \hat{\beta}_2 \end{pmatrix}. \tag{12}$$

Given the consistent estimators $\hat{\gamma}$ and $\hat{B}$, consistent estimator for α can be obtained by minimizing $(\hat{\gamma} - \hat{B}\alpha)' A_n (\hat{\gamma} - \hat{B}\alpha)$, where A_n is a non-negative definite weighting matrix which may depend on the data. The minimum distance estimator (MDE) is given by

$$\hat{\alpha} = (\hat{B}' A_n \hat{B})^{-1} \hat{B}' A_n \hat{\gamma}. \tag{13}$$

Further, if $\sqrt{n}(\hat{\gamma} - \gamma) \xrightarrow{L} N(0, \Sigma_{\hat{\gamma}})$, then the delta-method implies

$$\sqrt{n}\,(\hat{\alpha} - \alpha) \xrightarrow{L} N(0, (B'AB)^{-1} B' A \Sigma_{\hat{\gamma}} AB (B'AB)^{-1}),$$

where $A = \text{plim}(A_n/n)$. The lower bound of the above asymptotic covariance matrix is $(B'\Sigma_{\hat{\gamma}}^{-1} B)^{-1}$, which is attained by taking $A = \Sigma_{\hat{\gamma}}^{-1}$. The corresponding efficient MDE is obtained by using the weight $A_n = \hat{\Sigma}_{\hat{\gamma}}^{-1}$ which is a consistent estimator for $\Sigma_{\hat{\gamma}}^{-1}$. However, there are some other interesting and practical choices of A_n. For example, the identity weight $A_n = I$ results in $\hat{\alpha} = (\hat{B}'\hat{B})^{-1} \hat{B}' \hat{\gamma}$. Another choice is $A_n = (Z'Z)^2$ which leads to $\hat{\alpha} = (X'ZZ'X)^{-1} X'ZZ'Z\hat{\gamma}$. A particularly interesting choice is $A_n = Z'Z$, which gives

$$\hat{\alpha} = (X'Z(Z'Z)^{-1} Z'X)^{-1} X'Z\hat{\gamma} = (\hat{X}'\hat{X})^{-1} \hat{X}' Z \hat{\gamma}, \tag{14}$$

where $\hat{X} = Z\hat{B}$. The above $\hat{\alpha}$ has the same form as the two-stage least-squares (2SLS) estimator in simultaneous equations models. The only difference is that $\hat{\gamma}$ is not the least-squares estimator because of censoring. In this sense, the second-stage MDE can be regarded as a generalization of 2SLS estimator.

In the rest of this section, we derive the consistent estimator for σ_ε^2 under the additional assumption that $\Sigma_{x\eta}$ can be consistently estimated. First, from (1) and (2) we have $\Sigma_{x\eta} = \Sigma_{x\xi}\alpha_2 = \Sigma_\xi \alpha_2$ and hence

$$\begin{aligned} \sigma_\eta^2 &= \alpha_2' \Sigma_\xi \alpha_2 + \sigma_\varepsilon^2 \\ &= \Sigma_{x\eta}' \alpha_2 + \sigma_\varepsilon^2. \end{aligned} \tag{15}$$

On the other hand, from (7) we have

$$\sigma_\eta^2 = \gamma_2' \Sigma_z \gamma_2 + \sigma_u^2. \tag{16}$$

It follows then from (15) and (16) that

$$\sigma_\varepsilon^2 = \sigma_u^2 + \gamma_2' \Sigma_z \gamma_2 - \Sigma_{x\eta}' \alpha_2. \tag{17}$$

Therefore, σ_ε^2 can be consistently estimated as long as $\Sigma_{x\eta}$ does. One sufficient condition for consistent estimation of $\Sigma_{x\eta}$ is that the joint distribution of x and η satisfies

$$E(x|\eta) = \mu_x + \Sigma_{x\eta}(\eta - \mu_\eta)/\sigma_\eta^2. \tag{18}$$

For instance, if x and η are jointly normally distributed, then (18) holds and it further implies

$$\Sigma_{x\eta} = E(xy|y > 0) - \mu_x E(y|y > 0)$$

which can be consistently estimated using the uncensored sample points of (x_i, y_i). More generally, (18) remains true if x and η have an elliptically contoured distribution. In this case, (18) implies

$$\Sigma_{x\eta} = \frac{\sigma_\eta^2 \,[E(x|y > 0) - \mu_x]}{E(y|y > 0) - \mu_\eta}$$

which can be consistently estimated because $\mu_\eta = \gamma_1 + \gamma_2'\mu_z$ and $\sigma_\eta^2 = \gamma_2'\Sigma_z\gamma_2 + \sigma_u^2$. It is well known that the family of elliptically contoured distributions contains many commonly seen distributions such as uniform, normal, Cauchy, Student's t, double exponential and many other distributions (see, e.g. Fang and Zhang 1990; or Gupta and Varga 1993).

Finally, using the delta method it is straightforward to derive the asymptotic variance for the estimator $\hat{\sigma}_\varepsilon^2$ through the variances and covariance of $\hat{\gamma}$ and $\hat{\sigma}_u^2$. In the next two sections, two special instances of the two-stage estimation procedure of this section are investigated in more details.

4. MAXIMUM LIKELIHOOD BASED ESTIMATORS

In the previous section, we have derived the two-stage MDE for α without assuming the normal distribution for (x_i, z_i, η_i). If these variables are indeed normally distributed, then more efficient estimators can be obtained based on the maximum likelihood method. In this and the next section, we give some details for two estimation procedures, under the assumption that (x_i, z_i, η_i) are normally distributed.

Let $\hat{\gamma}$ denote the maximum likelihood estimator (MLE) for Tobit model (7) and (3), and let $\hat{\alpha}$ be the second-stage MDE (13) with $A_n = \hat{\Sigma}_{\hat{\gamma}}^{-1}$. Consistent estimate $\hat{\Sigma}_{\hat{\gamma}}^{-1}$ can be obtained using the second derivative of the log-likelihood with respect to the parameters (Wang 1998). In the following, we find the exact expression for $\Sigma_{\hat{\gamma}}^{-1}$. Wang (1998) derived the asymptotic covariance matrix of the MLE for model (1) and (2) under the condition that $\Sigma_\xi^{-1}\Sigma_\delta$ is known. The result can be applied to model (7) (without measurement error) by simply setting the measurement error covariance Σ_δ to zero in the covariance formula of Wang (1998). To simplify notation, denote $\tilde{z} = (1, z')'$, $\lambda = \phi(\gamma'\tilde{z}/\sigma_u)/\Phi(-\gamma'\tilde{z}/\sigma_u)$ and $\Psi = \Phi(\mu_\eta/\sigma_\eta)$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the standard normal density and distribution function. Then using the matrix partition and after some tedious matrix manipulation, we have

$$\Sigma_{\hat{\gamma}}^{-1} = \frac{\Psi}{\sigma_u^2}\left[C - \frac{(C\gamma - D)(C\gamma - D)'}{\gamma'(C\gamma - 2D) + E(y^2 \mid y > 0) + \sigma_u^2}\right], \tag{19}$$

where $C = E(\tilde{z}\tilde{z}' \mid y > 0) + (1/\Psi - 1)E[\lambda(\lambda - \gamma'\tilde{z})\tilde{z}\tilde{z}' \mid y = 0]$ and $D = E(y\tilde{z}' \mid y > 0)$. From (7), it is easy to see that $\mu_\eta = \gamma_1 + \gamma_2'\,\mu_z$ and $\sigma_\eta^2 = \gamma_2'\,\Sigma_z\gamma_2 + \sigma_u^2$.

5. TWO-STEP MOMENT ESTIMATORS

The MLE-based estimation of last section requires numerical maximization of a multi-dimensional likelihood function. Wang (1998) also proposes a two-step moment estimation procedure, which is computationally less intensive and easier to implement. In this section, we combine this moment estimator with the minimum distance estimation into a two-step procedure.

In the first step, the conditional means $\mu_y = E(y)$, $\mu_y^+ = E(y \mid y > 0)$ and $\mu_{zy}^+ = E(zy \mid y > 0)$ are first estimated using all y_i's and the positive y_i's, respectively. Then, by the moment equations (3) of Wang (1998), the other first-step parameters are consistently estimated through

$$\sigma_\eta = \frac{\mu_y\mu_y^+}{\mu_y\Phi^{-1} + \mu_y^+\phi(\Phi^{-1})},$$

$\mu_\eta = \sigma_\eta \Phi^{-1}$ and $\Sigma_{z\eta} = \mu_{zy}^+ - \mu_z \mu_y^+$, where $\Phi^{-1} = \Phi^{-1}(\mu_y/\mu_y^+)$. Then γ and σ_u^2 are estimated by the "least squares estimates" $\gamma_2 = \Sigma_z^{-1}\Sigma_{z\eta}$, $\gamma_1 = \mu_\eta - \gamma_2'\mu_z$ and $\sigma_u^2 = \sigma_\eta^2 - \Sigma_{z\eta}'\Sigma_z^{-1}\,\Sigma_{z\eta}$.

In the second step, the MDE for α is calculated using (13) and σ_ε^2 is estimated using (15) which implies $\sigma_\varepsilon^2 = \sigma_\eta^2 - \alpha_2'\Sigma_{x\eta}$, where $\Sigma_{x\eta} = \mu_{xy}^+ - \mu_x\mu_y^+$. Because now $\hat{\Sigma}_{\hat{\gamma}} = \hat{\sigma}_u^2(Z'Z)^{-1}$ and a multiplicative scalar in A_n does not affect the corresponding MDE, the efficient MDE can be calculated using $A_n = Z'Z$, which coincides with the generalized 2SLS estimator.

Analogue to the maximum likelihood based estimator, the asymptotic normality for this two-step estimator can be derived by using theorem 1 of Wang (1998) and the delta method. However, because the notations are quite complicated we omit the details here.

6. SIMULATION STUDIES

In this section, we carry out some simulation studies to investigate the finite sample behaviour of the two-stage estimators proposed in the previous sections and to compare their performances. For simplicity, we consider models with $k = 1$ predictor and $\ell = 2$ instrumental variables. In particular, we consider the following three models:

Normal model: $\eta_i = \alpha_1 + \alpha_2'\xi_i + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.
Non-normal model: $\eta_i = \alpha_1 + \alpha_2'\xi_i + \varepsilon_i$, where $\varepsilon_i \sim \sqrt{8}t(4)$ and $t(4)$ is the Student's t distribution with 4 degree of freedom.
Heteroskedastic model: $\eta_i = \alpha_1 + \alpha_2\xi_i + \varepsilon_i$, where $\varepsilon_i \mid \xi_i \sim (1 + 0.05\,|\xi_i|)t(4)$.

The true parameter values are $\alpha_1 = -4$, $\alpha_2 = 0.6$ in all the three and $\sigma_\varepsilon^2 = 16$ in the first two models. Other variables and parameters are as follows: $z_i \sim N_2[(5, 0)', \text{diag}(25, 25)]$; $\xi_i = \beta_1 + \beta_2'z_i + \tau_i$, where $\beta_1 = 5$, $\beta_2 = (2, -1)'$ and $\tau_i \sim N(0, 25)$; and $x_i = \xi_i + \delta_i$, where $\delta_i \sim N(0, 16)$.

All variables are generated independently and the observed responses are $y_i = \max\{\eta_i, 0\}$. In all the three models, the average amount of censoring in η_i is 27%. In the heteroskedastic model, $V(\varepsilon_i)$ varies between 2 and 26 approximately. The following estimators are calculated, where the optimal weighting matrices $A_n = \hat{\Sigma}_{\hat{\gamma}}^{-1}$ are used in the MDE step:

MLE: the maximum likelihood based estimator under normality;
TME: the two-step moment estimator under normality;
LAD: the MDE based on the censored least absolute deviations estimator;
NML: the naive maximum likelihood estimator ignoring the measurement errors.

Both MLE and NML have been computed using Newton–Raphson procedure (Wang 1998) which converged after four or five iterations with a convergence criterion of 10^{-8}. For the last two models, the LAD of Powell (1984,1986b) is calculated using the R implementation of Fitzenberger (1996) which is included in the R package *quantreg* by Koenker (2006). All models are simulated for each of the sample sizes $n = 50, 70, 100, 150, 300, 500$ and in each case $R = 1000$ Monte Carlo replications are carried out. The mean estimates, their Monte Carlo simulation standard errors and the root mean squared errors for all estimators are computed. The computation is done using the statistical computing language R on an IBM Workstation with a 2.2 MHz CPU and 4 GB RAM. For each given sample size, a simulation with 1000 runs is done within several seconds.

6.1. Normal model

The simulation results for the normal model are shown in Table 1. These results show that, even with an average censoring of 27% in η_i, both the MLE-based estimators and the TME perform quite satisfactorily except for the TME of σ_ε^2 for small sample sizes. This confirms the general perception that variance parameters are usually more difficult to estimate than regression parameters, especially when the sample sizes are relatively small. Overall, the MLE have smaller simulation standard errors as well as root mean squared errors than the TME, because they are statistically more efficient. However, the TME are computationally simpler and cheaper than the MLE. As expected, the NML are clearly biased and inconsistent for all the three parameters. It is worthwhile to note that we have also calculated the MLE and TME using the weight $A_n = Z'Z$. They perform very similar to their counterparts using the optimal weight $A_n = \hat{\Sigma}_{\hat{\gamma}}^{-1}$ except for having a slightly higher root mean squared errors.

6.2. Non-normal model

The simulation results of the LAD-based estimators for the non-normal model are given in Table 2. For comparison, we also included the MLE based estimator of Section 4, and the naive MLE ignoring measurement errors. The later two estimators are developed under the normality assumption. While the performance of the LAD-based estimators is satisfactory as expected, that of the MLE based estimators is a little surprisingly good, especially for the regression parameters. Overall, they have even smaller root mean squared errors than the LAD. This seems to indicate that the MLE-based estimators are not very sensitive to the deviation of the error distribution from normality if it remains symmetric. Again, the naive MLE are clearly biased and inconsistent for all parameters.

6.3. Heteroskedastic model

The simulation results for the heteroskedastic model are shown in Table 3. Again, we included the MLE and the NML for comparisons. Because in this model the random error ε_i is heteroskedastic, only parameters (α_1, α_2) are estimated. As has been mentioned earlier, although for the most part of the paper ξ_i and ε_i are assumed to be uncorrelated, this assumption is not necessary for the LAD or more general regression quantile based estimators (Powell 1984,1986b). The simulation results show that the LAD-based estimates are very good for all sample sizes. The MLE-based estimates are clearly biased in this model, but they have smaller root mean squared errors than the LAD based estimators except for sample size $n = 500$. This is in line with the previous simulation finding that the MLE based estimator shows certain degree of robustness against symmetric deviation from normality in the error terms. Finally, the performance of the naive MLE ignoring measurement errors is similar to that in previous simulations.

7. CONCLUSIONS AND DISCUSSION

Regression analysis with censored or truncated data and errors-in-variables are two important issues in many applied fields including econometrics. We have studied a combined model for censored outcome data with additive measurement errors in predictor variables. We

Table 1. Simulated mean estimates, simulation standard errors (in parentheses) and the root mean squared errors for normal model.

	$n = 50$	$n = 70$	$n = 100$	$n = 150$	$n = 300$	$n = 500$
$\alpha_1 = -4$						
MLE	−3.9816	−4.0368	−3.9957	−3.9945	−3.9791	−3.9887
	(0.0487)	(0.0404)	(0.0356)	(0.0270)	(0.0195)	(0.0152)
	1.5384	1.2778	1.1261	0.5835	0.6165	0.4816
TME	−3.9880	−4.0287	−3.9584	−4.0008	−4.0047	−3.9887
	(0.0554)	(0.0472)	(0.0403)	(0.0318)	(0.0226)	(0.0177)
	1.7514	1.4919	1.2759	1.0056	0.7154	0.5597
NML	−3.1596	−3.2176	−3.1452	−3.1450	−3.1484	−3.1275
	(0.0433)	(0.0370)	(0.0316)	(0.0240)	(0.0177)	(0.0137)
	1.6069	1.4067	1.3144	1.1401	1.0191	0.9743
$\alpha_2 = 0.6$						
MLE	0.6002	0.6011	0.5989	0.6000	0.5981	0.5994
	(0.0025)	(0.0021)	(0.0017)	(0.0014)	(0.0010)	(0.0008)
	0.0784	0.0650	0.0544	0.0430	0.0306	0.0244
TME	0.5960	0.5970	0.5948	0.5989	0.5986	0.5990
	(0.0028)	(0.0024)	(0.0020)	(0.0016)	(0.0011)	(0.0009)
	0.0897	0.0749	0.0627	0.0514	0.0357	0.0279
NML	0.5438	0.5453	0.5419	0.5432	0.5421	0.5418
	(0.0021)	(0.0018)	(0.0014)	(0.0011)	(0.0008)	(0.0007)
	0.0870	0.0785	0.0735	0.0673	0.0634	0.0618
$\sigma_\varepsilon^2 = 16$						
MLE	16.3518	16.0422	16.1673	15.9695	15.9553	15.9637
	(0.2060)	(0.1657)	(0.1393)	(0.1105)	(0.0781)	(0.0594)
	6.5190	5.2363	4.4051	3.4928	2.4687	1.8791
TME	17.3321	17.2531	16.9631	16.3493	16.2727	16.0905
	(0.3824)	(0.3198)	(0.2685)	(0.2135)	(0.1513)	(0.1158)
	12.1596	10.1864	8.5418	6.7569	4.7902	3.6617
NML	20.6708	20.5661	20.7125	20.8772	21.0296	21.0897
	(0.1582)	(0.1389)	(0.1161)	(0.0919)	(0.0679)	(0.0509)
	6.8426	6.3336	5.9723	5.6761	5.4689	5.3382

Table 2. Simulated mean estimates, simulation standard errors (in parentheses) and the root mean squared errors for non-normal model.

	$n = 50$	$n = 70$	$n = 100$	$n = 150$	$n = 300$	$n = 500$
$\alpha_1 = -4$						
LAD	−3.9143	−4.0107	−4.0451	−3.9290	−3.9964	−3.9996
	(0.0628)	(0.0543)	(0.0488)	(0.0391)	(0.0273)	(0.0219)
	1.9883	1.7167	1.5425	1.2370	0.8614	0.6927
MLE	−3.9892	−4.0734	−4.1203	−4.0497	−4.1200	−4.0854
	(0.0469)	(0.0399)	(0.0330)	(0.0263)	(0.0205)	(0.0150)
	1.4818	1.2622	1.0514	0.8329	0.6586	0.4805
NML	−3.1710	−3.2524	−3.2421	−3.2017	−3.2498	−3.2228
	(0.0430)	(0.0358)	(0.0306)	(0.0241)	(0.0185)	(0.0137)
	1.5919	1.3571	1.2282	1.1031	0.9503	0.8902
$\alpha_2 = 0.6$						
LAD	0.5935	0.6006	0.6022	0.5964	0.6002	0.5998
	(0.0030)	(0.0026)	(0.0023)	(0.0019)	(0.0013)	(0.0010)
	0.0960	0.0831	0.0738	0.0596	0.0410	0.0329
MLE	0.6001	0.6052	0.6073	0.6031	0.6069	0.6047
	(0.0024)	(0.0020)	(0.0017)	(0.0013)	(0.0010)	(0.0008)
	0.0749	0.0648	0.0532	0.0424	0.0328	0.0245
NML	0.5437	0.5485	0.5485	0.5458	0.5484	0.5470
	(0.0021)	(0.0017)	(0.0015)	(0.0012)	(0.0009)	(0.0007)
	0.0864	0.0756	0.0691	0.0654	0.0586	0.0570
$\sigma_\varepsilon^2 = 16$						
LAD	19.1620	18.3492	17.9263	16.7925	16.6038	16.0455
	(0.4184)	(0.3124)	(0.3795)	(0.2060)	(0.1813)	(0.1075)
	13.5962	10.1500	12.1474	6.5598	5.7614	3.3973
MLE	15.9278	15.6933	15.2337	15.3918	15.6280	15.2705
	(0.4450)	(0.2400)	(0.2076)	(0.1774)	(0.1565)	(0.0957)
	14.0644	7.5917	6.6056	5.6393	4.9590	3.1130
NML	19.6550	19.7973	19.7681	19.9394	20.4617	20.0625
	(0.3675)	(0.1986)	(0.1757)	(0.1459)	(0.1371)	(0.0814)
	12.1776	7.3361	6.7107	6.0660	6.2186	4.8085

Table 3. Simulated mean estimates, simulation standard errors (in parentheses) and the root mean squared errors for heteroskedastic model.

	$n = 50$	$n = 70$	$n = 100$	$n = 150$	$n = 300$	$n = 500$
			$\alpha_1 = -4$			
LAD	−4.0389	−4.0770	−3.9991	−4.0075	−4.0707	−4.0756
	(0.0536)	(0.0460)	(0.0403)	(0.0315)	(0.0249)	(0.0183)
	1.6942	1.4548	1.2732	0.9942	0.7889	0.5836
MLE	−4.4858	−4.5274	−4.4438	−4.4886	−4.5031	−4.4814
	(0.0408)	(0.0357)	(0.0294)	(0.0239)	(0.0181)	(0.0127)
	1.3769	1.2465	1.0301	0.9008	0.7613	0.6269
NML	−3.5959	−3.6644	−3.6021	−3.6310	−3.6303	−3.6135
	(0.0368)	(0.0316)	(0.0266)	(0.0210)	(0.0161)	(0.0113)
	1.2307	1.0531	0.9306	0.7596	0.6301	0.5263
			$\alpha_2 = 0.6$			
LAD	0.5956	0.5977	0.5935	0.5945	0.5970	0.5971
	(0.0027)	(0.0024)	(0.0021)	(0.0016)	(0.0013)	(0.0009)
	0.0870	0.0746	0.0738	0.0505	0.0402	0.0290
MLE	0.6212	0.6225	0.6184	0.6205	0.6207	0.6198
	(0.0022)	(0.0020)	(0.0016)	(0.0013)	(0.0010)	(0.0007)
	0.0727	0.0658	0.0545	0.0464	0.0370	0.0295
NML	0.5605	0.5638	0.5606	0.5623	0.5619	0.5609
	(0.0019)	(0.0016)	(0.0014)	(0.0011)	(0.0008)	(0.0006)
	0.0720	0.0632	0.0598	0.0518	0.0463	0.0434

proposed a two-stage procedure based on the availability of the instrumental variables which produces consistent and asymptotically normally distributed estimators. We also examined a maximum likelihood based estimator and a two-step moment estimator under the joint normality assumption. The numerical computation of the proposed estimators is straightforward using widely available statistical or econometric computer packages. Simulation studies show that they behave satisfactorily for small samples and relatively high degree of censoring. In all simulations, the two-stage MDE using the weighting matrix $A_n = Z'Z$ perform very similarly to the efficient MDE using optimal weight, except that they have slightly higher root mean squared errors. Another interesting finding is that the normal maximum likelihood based estimators appear to be fairly robust against certain non-normal but symmetric random error distributions. How general this robustness property is remains to be examined through more extensive simulation studies in the future. The proposed method can be extended to other limited dependent variable models. As pointed out by a referee, equation (5) may be generalized to a non-linear function of the instruments z. Such a non-linear function can still be consistently estimated with non-linear regression method and hence the results of this paper continue to hold. Moreover, z may contain as its elements some non-linear transformations of exactly measured predictor variables.

ACKNOWLEDGMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would like to thank two referees and the co-editor for their constructive comments and suggestions.

REFERENCES

Aigner, D. J., C. Hsiao, A. Kapteyn and T. Wansbeek (1984). Latent Variable Models in Econometrics. In Z. Griliches and M. D. Intriligator eds., *Handbook of Econometrics,* Vol. II, North-Holland, Amsterdam.

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica 41*, 997–1016.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics 24*, 3–61.

Amemiya, T. (1985a). *Advanced Econometrics*. Basil Blackwell, Oxford, UK.

Amemiya, Y. (1985b). Instrumental variable estimator for the non-linear errors-in-variables model. *Journal of Econometrics 28*, 273–89.

Amemiya, Y. (1990). Two-stage instrumental variable estimators for the non-linear errors-in-variables model. *Journal of Econometrics*, *44*, 311–32.

Carroll, R.J., D. Ruppert and L.A. Stefanski (1995). *Measurement Error in Nonlinear Models*. Chapman & Hall, London.

Dood, L.E., E.L. Korn, L.M. McShane, G.V.R. Chandramouli and E.Y. Chuang (2004). Correcting log ratios for signal saturation in cDNA microarrays. *Bioinformatics 20*, 2685–93.

Fang, K.T. and Y.T. Zhang (1990). *Generalized Multivariate Analysis*. Springer-Verlag, New York.

Fitzenberger, B. (1996). A guide to censored quantile regression. In C. R. Rao and G. Maddala (eds), *Handbook of Statistics*. North-Holland, New York.

Fuller, W.A. (1987). *Measurement Error Models*. Wiley, New York.

Gupta, A.K. and T. Varga (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, Dordrecht.

Gustafson, P. (2004). *Measurement Error and Misclassification in Statistics and Epidemiology - Impacts and Bayesian adjustments*. Chapman & Hall /CRC, Boca Raton.

Heckman, J.J. and T.E. MaCurdy (1986). Labor Econometrics. In Z. Griliches and M.D. Intriligator (eds), *Handbook of Econometrics III*, North-Holland, Amsterdam.

Hsiao, C. (1983). Identification. In Z. Griliches and M.D. Intriligator (eds), *Handbook of Econometrics*, Vol. I. North-Holland, Amsterdam.

Hsiao, C. (1989). Consistent estimation for some non-linear errors-in-variables models. *Journal of Econometrics 41*, 159–85.

Hsiao, C. (1991). Identification and estimation of dichotomous latent variables models using panel data. *Review of Economic Studies 58*, 717–31.

Khan, S. and J.L. Powell (2001). Two-step estimation of semiparametric censored regression models. *Journal of Econometrics 103*, 73–110.

Killingsworth, M.R. and J.J. Heckman (1986). Female labor supply: A survey. In O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics I*, North-Holland, Amsterdam.

Koenker, R. (2006). quantreg: Quantile Regression. *R package version 3.85*. http://www.r-project.org

Maddala, G. S. (1985). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.

Powell, J.L. (1984). Least absolute deviations estimation of the censored regression models. *Journal of Econometrics 25*, 303–25.

Powell, J.L. (1986a). Symmetrically trimmed least squares estimation for Tobit models. *Econometrica*, *54*, 1435–60.

Powell, J.L. (1986b). Censored regression quantiles. *Journal of Econometrics 32*, 143–55.

Schennach, S.M. (2004). Instrumental variable estimation of non-linear errors-in-variables models. *Working Paper*, Department of Economics, University of Chicago.

Wang, L. (1998). Estimation of censored linear errors-in-variables models. *Journal of Econometrics 84*, 383–400.

Wang, L. (2002). A simple adjustment for measurement errors in some limited dependent variable models. *Statistics and Probability Letters 58*, 427–33.

Wang, L. and C. Hsiao (1995). A simulated semiparametric estimation of non-linear errors-in-variables models. *Working Paper*, Department of Economics, University of Southern California.

Weiss, A.A. (1993). Some aspects of measurement error in a censored regression model. *Journal of Econometrics 56*, 169–88.