2001

Asymptotic Inference in Censored Regression Models Revisited

Chihwa Kao
Syracuse University. Center for Policy Research, cdkao@maxwell.syr.edu

Follow this and additional works at: https://surface.syr.edu/cpr

Part of the Mathematics Commons

Recommended Citation
https://surface.syr.edu/cpr/119

This Working Paper is brought to you for free and open access by the Maxwell School of Citizenship and Public Affairs at SURFACE at Syracuse University. It has been accepted for inclusion in Center for Policy Research by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.
Center for Policy Research
Working Paper No. 36

ASYMPTOTIC INFERENCE IN CENSORED
REGRESSION MODELS REVISITED

Chihwa Kao

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

February 2001

$5.00

Up-to-date information about CPR’s research projects and other activities is available from our World Wide Web site at www-cpr.maxwell.syr.edu. All recent working papers and Policy Briefs can be read and/or printed from there as well.
CENTER FOR POLICY RESEARCH – Spring 2001

Timothy M. Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret M. Austin
Associate Director, Budget and Administration

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

Douglas Holtz-Eakin
Chair, Professor of Economics
Associate Director, Center for Policy Research

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Scott Allard .......................................................Public Administration
Dan Black ..................................................Economics
Stacy Dickert-Conlin.................................Economics
William Duncombe ........................................Public Administration
Gary Engelhardt ........................................Economics
Deborah Freund ........................................Public Administration
Vernon Greene ................................................Public Administration
Leah Gutierrez ........................................Public Administration
Madonna Harrington Meyer .......................Sociology
Christine Himes ........................................Sociology
Jacqueline Johnson .......................................Sociology
Bernard Jump ................................................Public Administration
Duke Kao .................................................Economics

Eric Kingson .......................................................Social Work
Thomas Kriesner ........................................Economics
Jeff Kubik ................................................Economics
Jerry Miner ................................................Economics
Jan Ondrich .................................................Economics
John Palmer ...............................................Public Administration
Lori Ploutz-Snyder .....................................Health and Physical Education
Grant Reeher .............................................Political Science
Stuart Rosenthal ..........................................Economics
Jodi Sandfort .............................................Public Administration
Michael Wasylenko .......................................Economics
Assata Zerai ..............................................Sociology

GRADUATE ASSOCIATES

Reagan Baughman ........................................Economics
Robert Bifulco ............................................Public Administration
Caroline Bourdeaux .................................Public Administration
Christine Caffrey ........................................Sociology
Christopher Cunningham .........................Economics
Tae Ho Eom .............................................Public Administration
Seth Gertz...................................................Economics
Andrzej Godnierz .......................................Economics
Rain Henderson ....................................Public Administration
Pam Herd ....................................................Sociology
Lisa Hotchkiss ........................................Public Administration
Peter Howe ...............................................Economics
Benjamin Johns ........................................Public Administration
Anil Kumar ...............................................Economics

Kwangho Jung ................................................Public Administration
James Ladlifka ............................................Public Administration
Xiali Liang .................................................Economics
Donald Marples .........................................Economics
Neddy Matshalaga ......................................Sociology
Suzanne Plourde ..........................................Economics
Nora Ranney .............................................Public Administration
Catherine Richards ....................................Sociology
Adriana Sandu ............................................Public Administration
Mehmet Serkan Tosun ................................Economics
Mark Trembley .........................................Public Administration
James Williamson .......................................Economics
Bo Zhao ...................................................Economics

STAFF

JoAnna Berger ..................................................Receptionist
Martha Bonney ........................................ Publications/Events Coordinator
Karen Cimilluca ...................................... Librarian/Office Coordinator
Kati Foley ..................................................Administrative Assistant, LIS
Esther Gray ................................................ Administrative Secretary
Kitty Nasto ................................................ Administrative Secretary
Denise Paul ................................................ Editorial Assistant, NTJ

Mary Santy ........................................ Administrative Secretary
Amy Storfer-Isser ................................ Computer Support,
Debbie Tafel ................................................ Secretary to the Director
Ann Wicks ............................................. Administrative Secretary
Lobrenzo Wingo ........................................... Computer Consultant
Abstract

This paper establishes that regressors in the models with censored dependent variables need not be bounded for the standard asymptotic results to apply. Thus regressors which grow monotonically with the observation index may be acceptable. It also purports to provide an upper bound on the rate at which regressors may grow.

We show that if $\|x_t\| \leq c$ for all $t$, then $\lambda_{\min} \sum_{t=1}^{T} x_t x_t^T \rightarrow \infty$ is a sufficient condition for the consistency and asymptotic normality of the MLE in censored regression models, which are different from those used by Amemiya (1973). For the case of growing regressors, we show that the sufficient conditions for the consistency and asymptotic normality of the MLE are $\|x_t\|^2 = o(\log t)$ and $\lambda_{\min} \sum_{t=1}^{T} x_t x_t^T \geq cT^\alpha$, for some $\alpha > 0$ and $c > 0$, but only for one-half of the parameter space. The admissible growth rate given above implies that the Fisher Information matrix diverges, which seems to be an indispensable requirement for asymptotic inference for the censored regression models. More importantly, it represents a critical upper bound in the Fisher information matrix if it is exceeded monotonically. It also implies that $x_t = (1, t)^T$ in Judge et al. (1985:791) for the Tobit is not admissible.
1. Introduction

Consider the linear regression model

\[ y_t = x'_t \beta + \varepsilon_t, \quad (t = 1, \ldots, T) \]  

(1)

where the \( \varepsilon_t \) are i.i.d. normal random variables having mean zero and finite variance \( \sigma^2 \), \( x_t \) is a k x 1 vector of regressors and \( \beta \) is a k x 1 vector of parameter of interest. However, for each \( t \), instead of observing \((y_t, x_t)\), one observes \((z_t, \delta_t, x_t)\), where \( z_t = \max\{y_t, 0\} \), and \( \delta_t \) is an indicator variable taking the value 1 if \( y_t > 0 \) and 0 if \( y_t \leq 0 \).

The consistency and asymptotic normality of the maximum likelihood estimator (MLE) of \((\beta^T, \sigma^2)\) have been studied by Amemiya (1973) under assumptions of \( \{x_t\} \), e.g., \( x_t \) is bounded, i.e.,

\[ \|x_t\| \leq c \quad \text{for all } t, \text{ and } \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x_t x'_t = V, \text{ a positive definite matrix. However, the conditions given by Amemiya (1973) above may be too strong to study the cases where growing regressors are of interest (e.g., time trend model, i.e, } x_t = (1, t)', \text{ see Judge et al. 1985:791).}

The likelihood function of this model is

\[ L = \prod_{t=1}^{T} [1 - \Phi(x'_t \beta, \sigma^2)]^{-\delta_t} \left[ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (y_t - x'_t \beta)^2} \right]^{\delta_t} \]  

(2)

with

\[ \Phi(x'_t \beta, \sigma^2) = \int_{-\infty}^{x'_t \beta} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (\lambda/\sigma)^2} d\lambda \]  

(3)

In section 2, we confirmed Fahrmeir’s (1987:103) result that if \( \|x_t\| \leq c \) for all \( t \), then

\[ \lambda_{\min} \sum_{t=1}^{T} x_t x'_t \to \infty \]  

(4)
is a sufficient condition for the consistency and asymptotic normality of the MLE for the censored model but only for one-half of the parameter space, where $\lambda_{\min} A$ is the smallest eigenvalues of a symmetric matrix $A$. In Section 3 the case of growing regressors is discussed. We show that the sufficient conditions for the consistency and asymptotic normality of the MLE are

$$\|x_i\|^2 = o(\log t)$$

and

$$\lambda_{\min} \sum_{i=1}^T x_i'x_i \geq cT^\alpha, \text{ for some } \alpha > 0 \text{ and } c > 0,$$

for one-half of the parameter space. The summary is given in the Section 4.

**Remark 1:** In the classical linear regression model with i.i.d errors and the sequence $\{x_t\}$ of regressors, $\lambda_{\min} \sum_{i=1}^T x_i'x_i \rightarrow \infty$ is necessary and sufficient for weak (Drygas 1976) and strong (Lai, Robbins, and Wei 1979) consistency (also see Amemiya 1985:95).

**Remark 2:** Note that $\Phi(x_i'\beta, \sigma^2)$ will tend to one or zero if some regressors are growing monotonically to $+\infty$ or $-\infty$. Thus, for large $T$ nearly all response $y_t$ will fall into one category and there will be too less information to draw inference about the relevant parameters. The admissible growth rate in (5) and (6) assure that enough information is available and the asymptotic theory works.

**Remark 3:** The admissible growth rate, $\|x_i\|^2 = o(\log t)$, given in (5) implies that the Fisher information matrix diverges, which seems to be an indispensable requirement for asymptotic inference for the censored regression models. More importantly, it represents a critical upper bound in the Fisher information matrix if it is exceeded monotonically.
Remark 4: Fahrmeir and Kaufmann (1986) and Gourieroux and Monfort (1981) have discussed the sharp upper bounds on the admissible growth of regressors for logit, probit, cumulative logit, and loglinear, and linear Poisson model.

2. Statistical Inference

Olsen (1978) proved the global concavity of $\log L$ in the Tobit model in terms of the transformed parameter $\alpha = \beta / \sigma$ and $h = 1 / \sigma$. Without loss of generality (see Remark 7), we assume $\sigma = 1$. The $\log L$ in terms of the new parameters can written as (see Olsen 1978)

$$
\log L = \sum_0 \log[1 - \Phi(x_i \beta)] - \frac{1}{2} \sum_1 (y_i - x_i \beta)^2
$$

where $\sum_0$ is the summation over all observations where $y_t = 0$, $\sum_1$ is the summation over all observations where $y_t > 0$, and $T_1$ is the number of observations where $y_t > 0$. The score function $s_i(\beta)$ and the information matrix $F_i(\beta)$ are

$$
s_i(\beta) = -\sum_0 x_i \frac{\phi(x_i \beta)}{1 - \Phi(x_i \beta)} + \sum_1 x_i (y_i - x_i \beta)
$$

$$
= \sum_i x_i \left[ \delta_i \left( y_i - x_i \beta \right) - (1 - \delta_i) \left( \frac{\phi(x_i \beta)}{1 - \Phi(x_i \beta)} \right) \right]
$$

$$
= \sum_i x_i \left[ \left( \delta_i y_i + (1 - \delta_i) \left( x_i \beta - \frac{\phi(x_i \beta)}{1 - \Phi(x_i \beta)} \right) \right) - x_i \beta \right]
$$

$$
= \sum_i x_i \left( \hat{y}_i - x_i \beta \right), \text{ say,}
$$
\( F_i(\beta) = \text{cov}_{\beta_i, \beta_j}(\beta) = E_{\beta} H_i(\beta) = E_{\beta} \left[ \sum_{j=0}^{\phi(x_j \beta)} \left( \frac{\phi(x_i \beta)}{1 - \Phi(x_j \beta)} - x_j \beta \right) x_i x_j + \sum_j x_i x_j \right] \)

\( = E_{\beta} \left[ \sum_j \left(1 - \delta_j \right) \left[ \frac{\phi(x_j \beta)}{1 - \Phi(x_j \beta)} \left( \frac{\phi(x_i \beta)}{1 - \Phi(x_j \beta)} - x_j \beta \right) \right] \right] \)

\( = \sum_j \left( \frac{\phi(x_i \beta)}{1 - \Phi(x_j \beta)} - x_j \beta \right) \frac{\phi(x_i \beta)}{1 - \Phi(x_j \beta)} + \Phi(x_j \beta) \right) x_i x_j \]

(9)

where \( H_i(\beta) = -\frac{\partial^2 \log L}{\partial \beta \partial \beta} \) and \( \frac{\phi(x_i \beta)}{1 - \Phi(x_i \beta)} - x_i \beta > 0 \). Therefore \( H_i(\beta) \) and \( F_i(\beta) \) are positive definite.

Remark 5: \( E \hat{y} = E \left[ \delta y_i + (1 - \delta) \left( x_i \beta - \frac{\phi(x_i \beta)}{1 - \Phi(x_i \beta)} \right) \right] = x_i \beta \) and

\( \text{Var}(\hat{y}) = \sigma_i^2 = \sigma^2 - \text{Var}(y_i | y_i < -x_i \beta) \). In particular, \( \sup_{i \in \mathbb{R}} \sigma_i^2 \leq \sigma^2 < \infty \) (see James and Smith 1984).

Let \( A^{1/2} \) be a unique symmetric positive definite matrix associated with a symmetric positive definite matrix A such that \((A^{1/2})^2 = A\). We approach the problem of the distribution of \( F_i^{-1/2}s \) first, and move to the large sample distribution of \( \hat{\beta} \), MLE of \( \beta \).

Assumption 1: \( |x_t| \leq c \) for all \( t \),

Assumption 2: \( \lambda_{\min} \sum_{t=1}^{T} x_t x_t \to \infty \).

Although Fahrmeir (1987:103) did not fully present a separate and detailed proof regarding the sufficient condition for the Tobit model with bounded regressors, he indicated that the stated sufficient conditions in Assumption 1 and Assumption 2 can be directly proved by his general results presented in his paper. Here, we merely confirm his conjecture.
Lemma 1: Under Assumption 1 and Assumption 2, the normed score function is asymptotically normal: \( F_i^{-1/2} s_i \xrightarrow{d} N(0, I) \).

Proof: We use the Lindeberg-Feller Theorem for triangular arrays. Fix \( \lambda \) with \( \lambda' \lambda = 1 \).

\[
z_{t,j} = \lambda F_i^{-1/2} x_i \left( \hat{y}_i - x_i \beta \right)
\]

we have \( E z_{t,i} = 0 \), \( \text{var} \sum_{i=1}^t z_{i,t} = \text{var} \lambda F_i^{-1/2} s_i = 1 \), i.e., \( z_{t,j} \) are independent and \( \sum_{i=1}^t z_{t,i} \) has mean 0 and variance 1. In order that \( z_{t,j} \) obey the central limit theorem, it is sufficient that the Lindeberg condition (see Billingsley 1986:369) is satisfied, i.e., for any \( \varepsilon > 0 \),

\[
\lim_{t \to \infty} \sum_{i=1}^t \int_{\{z_{i,t} > \varepsilon\}} z_{i,t}^2 \text{d}P = 0 .
\]

where \( P \) is the distribution of \( z_{t,i} \). Let \( \alpha^i_{t,i} = \lambda F_i^{-1/2} x_i \). By the Cauchy-Schwarz inequality, we have

\[
z_{t,i}^2 \leq \alpha^i_{t,i} \alpha_{t,i} \left( \hat{y}_i - x_i \beta \right)^2 = \alpha^i_{t,i} \alpha_{t,i} e_i^2 ,
\]

where \( e_i = \hat{y}_i - x_i \beta \). This gives

\[
\sum_{i=1}^t \int_{\{z_{i,t} > \varepsilon\}} z_{i,t}^2 \text{d}P \leq \sum_{i=1}^t \alpha^i_{t,i} \alpha_{t,i} \int_{B(t,i)} e^2 \text{d}G_x
\]

where \( G_x \) is the distribution of \( e \) for a given \( x \), and \( B(t,i) \) is the set \( \left\{ e^2 > \frac{e^2}{\alpha^i_{t,i} \alpha_{t,i}} \right\} \). From Assumption 1 and 2, we have \( \lambda_{\min} F_i \to \infty \). Define

\[
h_c(x) = \sup \int_{\{e^2 > \varepsilon^2\}} e^2 \text{d}G_x
\]

Under Assumption 1, \( \sum_{i=1}^t \alpha^i_{t,i} \alpha_{t,i} \leq K < \infty \) with a constant \( K \). Note that \( \sum_{i=1}^t \alpha^i_{t,i} \alpha_{t,i} \int_{B(t,i)} e^2 \text{d}G_x \) is bounded above by \( Kh_c(x) \).

\[
\sum_{i=1}^t \int_{\{z_{i,t} > \varepsilon\}} z_{i,t}^2 \text{d}P \leq K h_c(x) \to 0
\]
because \( h_i(x) \to 0 \) as \( c \to \infty \).

**Theorem 1:** Under Assumption 1 and Assumption 2

(i) \( \hat{\beta} \xrightarrow{p} \beta \)

(ii) \( F_t^{-1/2}(\hat{\beta} - \beta) \xrightarrow{d} N(0, I) \)

**Proof:** The proof is similar to the Theorem 4 in Fahrmeir and Kaufmann (1985:364).

**Remark 6:** A main step is the verification of

\[
F_t^{-1/2}(\beta)H_t(\beta)F_t^{-1/2}(\beta) \xrightarrow{p} I. \tag{14}
\]

This assertion is equivalent to \( \frac{\lambda H_t(\beta)\lambda}{\lambda F_t(\beta)\lambda} \to 1 \), uniformly all \( \lambda \neq 0 \), i.e., Condition (13) is a continuity condition on \( H_t(\beta) \) as well as a convergence condition on the asymptotic relation between \( F_t(\beta) \) and \( H_t(\beta) \), requiring that the ratio between observed information \( H_t(\beta) \) and expected information \( F_t(\beta) \) converges to one.

**Remark 7:** The above results hold even if \( \sigma^2 \) is unknown. Note that

\[
\lambda_{\min} F_t \geq \lambda_{\min} \begin{bmatrix}
\sum_0 \left( \frac{\phi(x_i, \alpha)}{1 - \Phi(x_i, \alpha)} - x_i \alpha \right) \phi(x_i, \alpha) + \Phi(x_i, \alpha) \right) x_i x_i & 0 \\
0 & T_i \frac{h}{h^2}
\end{bmatrix}.
\]

Therefore \( \lambda_{\min} F_t \to \infty \) if and only if \( \lambda_{\min} \sum_{i=1}^T x_i x_i \to \infty \) because

\[
\lambda_{\min} \sum_{i=1}^T x_i x_i \to \infty \text{ iff } \lambda_{\min} \sum_{i=0}^T x_i x_i \to \infty.
\]
3. Growing Regressors

However, there are situations where growing regressors are of interest, e.g., time trend models (e.g., Judge et al. 1985:791). Fahrmeir and Kaufmann (1986:187) have given gives a sharp upper bound for admissible growth of regressors for the Probit. Note that Tobit like-likelihood in (2) is expressible as the sum of the probit log-likelihood and the truncated likelihood and the Probit MLE is asymptotically normal and $\sqrt{T}$—consistent for the original Tobit parameter vector divided by the standard deviation of the Tobit error term, one may think that the sufficient conditions on the growing regressors in the Tobit model and in the Probit model coincide.

It is true that $\Phi(x_i' \beta, \sigma^2)$ in (2) will tend to one or zero if some regressors are growing monotonically to $+\infty$ or $-\infty$. However, the statement by Fahrmeir and Kaufmann (1986:189) about the Probit: “Thus for large $T$ nearly all response $y_t$ will fall into one category and there will be too less information to draw inference about the relevant parameters” is only one-half right for the Tobit. Specifically, if the variable $x_{it}$ is growing monotonically with $t$, and if its associated coefficient $\beta_i$ is negative, then for large $T$ nearly all responses $y_t$ will be zero, so that additional observations will be indeed add too little information for asymptotic theory to work. That is the case that looks like Probit. But if the coefficient $\beta_i$ is positive, then nearly all $y_t > 0$ for large $T$, meaning new observations will add as much information as they would in the classical linear model. Thus, the upper bound on the growth rates for regressors in the censored normal model, is true as stated in (5) and (6), but only for half of the parameter space to assure that enough information is available and the asymptotic theory works.

The following theorem gives a sharp upper bound for admissible growth of regressors for half of the parameter space.
Theorem 2: If \( \| x \|^2 = o \left( \log t \right) \) and \( \lambda_{\min} \sum_{r=1}^{T} x_r x_r' \geq c T^\alpha \), for some \( \alpha > 0 \) and \( c > 0 \). Then
\[ \lambda_{\min} F_t \to \infty. \]

Proof: The information matrix \( F_t(\beta) \) can also be written as \( F_t(\beta) = \sum_{r=1}^{T} x_r x_r' \), with
\[ a_r^2 = \left( \frac{\phi(x_r \beta)}{(1 - \Phi(x_r \beta))} - x_r \beta \right) \phi(x_r \beta) + \Phi(x_r \beta) \]  \hspace{1cm} (15)

From Magnus and Neudecker (1988:204), we have
\[ \lambda_{\min} F_t(\beta) = \lambda_{\min} \sum_{r=1}^{T} x_r x_r' a_r^2 \geq \lambda_{\min} \sum_{r=1}^{T} \left[ \Phi(x_r \beta) \right] x_r x_r'. \] \hspace{1cm} (16)

We also note that
\[ \Phi(x_r \beta) \geq c \exp(-\{x_r \beta\}^2) \] \hspace{1cm} (17)
for some \( c > 0 \). The admissible growth rate of \( x_t, \| x \|^2 = o \left( \log t \right) \), in Theorem 2 is equivalent to
\[ \exp\left( -\| x \|^2 \| \beta \|^2 \right) \geq t^{-\delta} \text{ for all } \beta, \delta > 0, t \geq t_1(\beta, \delta). \] \hspace{1cm} (18)

In combination with (16) and (17),
\[ \lambda_{\min} F_t(\beta) \geq \lambda_{\min} \sum_{r=1}^{T} x_r x_r' t^{-\delta}. \]

Since \( t^{-\delta} \geq T^{-\delta} \) for \( t \leq T, \delta > 0 \). Under \( \lambda_{\min} \sum_{r=1}^{T} x_r x_r' \geq c T^\alpha \), therefore, we have
\[ \lambda_{\min} F_t(\beta) \geq c T^{\alpha-\delta}, \quad T \geq T_1 \]
With \( \delta = \alpha / 2 \), this implies that \( \lambda_{\min} F_t \to \infty. \)

Remark 8: If \( \| x \|^2 \geq c \log t, c > 1 \), then \( \lambda_{\min} F_t \) converges.
Example: Let \( x_t = (1, w_t)' \) with \( w_t^2 = (\log t)^\alpha \) for the Tobit Model in (1). \( \|x_t\|^2 = o(\log t) \) holds if 

\[ 0 < \alpha < 1. \quad \text{If } \alpha \geq 1, \text{ then } \lambda_{\min} F_t \text{ converges for some } \beta. \]

Remark 9: It implies that \( x_t = (1, t)' \) in Judge et al. (1985:791) for the Tobit is not admissible.

Remark 10: It also implies that when \( x_t = (1, w_t) \) has a unit root with the drift, i.e., \( w_t = \mu + w_{t-1} + \omega_t \)

for the Tobit is not admissible. Since \( w_t \) can be written as \( w_t = w_0 + ut + \sum_{i=1}^{t} \omega_i \).

4. Summary

Regression models for censored data have found numerous applications. Statistical analysis of these models relies heavily on large sample theory, i.e., asymptotic properties of the MLE. However, previously published conditions assuring these properties may be too strong. Consistency and

Asymptotic normality of the MLE are shown under weak and easily verifiable requirements. This paper gives a sharp upper bound on the admissible growth of regressors.
References


