A Note on the Nonstationary Binary Choice Logit Model

Emmanuel Guerre and Hyungsik Roger Moon

USC Center for Law, Economics & Organization
Research Paper No. C02-3
A Note on the Nonstationary Binary Choice Logit Model

Emmanuel Guerre
LSTA (Université Paris 6)
February 2002

Hyungsik Roger Moon*
University of Southern California

Abstract

This paper derives the rate and the asymptotic distribution of the
MLE of the parameter of a logit model with a nonstationary covariate
when the true parameter is zero. The limit distribution of the t-statistic
is also given.

Keywords: binary choice model; nonstationary covariates.

JEL Classification number: C22, C25.

1 Introduction

Suppose that the observed data \((y_t, x_t')\)' is generated by
\[
\begin{align*}
y_t &= 1\{y_t^* > 0\}, \\
y_t^* &= \beta_0 x_t - \varepsilon_t, \ t = 1, ..., n.
\end{align*}
\]

The model (1) is the standard binary choice model. When the regressor \(x_t\)
and the error term \(\varepsilon_t\) are stationary, it is well known that under regularity
conditions the maximum likelihood estimator (MLE) of \(\beta_0\) is \(\sqrt{n}\)–
consistent and asymptotically normal. However, if the regressors \(x_t\) are nonstationary,
the MLE of \(\beta_0\) do not have these standard asymptotic properties. Recently,
Park and Phillips (2000) consider the model (1) in which the regressors \(x_t\) are
nonstationary, i.e.,
\[
x_t = x_{t-1} + u_t
\]
and coefficient \(\beta_0\) is not zero, i.e.,
\[
\beta_0 \neq 0.
\]

*Corresponding Author. Address: KAP 300, Department of Economics, University of
Southern California, Los Angeles, CA 90089. Tel: 213-740-2108. Email: moonr@usc.edu
Under regularity conditions, Park and Phillips (2000) obtain dual convergence rates for the MLE. In an orthogonal direction to $\beta_0$, the MLE converges to a mixed normal random variable at a rate of $n^{3/4}$ while in all other directions, the convergence rate is $n^{1/4}$.

In this note, we derive the asymptotic properties of the MLE of $\beta_0$ when the true $\beta_0$ is zero. The model we are considering in this note is quite simple. We assume that (i) $x_t$ is univariate, (ii) $u_t \sim iid (0, 1)$, (iii) $\varepsilon_t$ are iid over $t$ with the logistic distribution $P\{\varepsilon_t < x\} = \Lambda (x) = \frac{e^x}{1 + e^x}$, (iv) $u_t$ and $\varepsilon_s$ are independent for all $t$ and $s$. These restrictive assumptions are made to make derivations of asymptotics simple and short. The assumptions in this note can be relaxed at a cost of lengthy calculations and derivations.

The main result we find in this paper is that when the true $\beta_0$ is zero and the regressor of the nonstationary binary choice is nonstationary, the MLE is $n$-consistent and its limit distribution is similar to so-called the unit root limit distribution. This main result will be derived in the Section 2.

2 Results

The MLE $\hat{\beta}$ maximizes the following log-likelihood function,

$$l_n (\beta) = \sum_{t=1}^{T} y_t \log \Lambda (\beta x_t) + \sum_{t=1}^{T} (1 - y_t) \log (1 - \Lambda (\beta x_t)).$$

As well known, the log-likelihood function $l_n (\beta)$ is smooth and strictly concave with respect to the parameter $\beta$. Thus, at the MLE $\hat{\beta}$, we have

$$\frac{\partial l_n}{\partial \beta} \hat{\beta} = 0.$$

Define

$$S_n (\beta) = \frac{\partial l_n (\beta)}{\partial \beta} = \sum_{t=1}^{T} x_t (y_t - \Lambda (\beta x_t)),$$

the score function,

and

$$J_n (\beta) = \frac{\partial^2 l_n (\beta)}{\partial \beta^2} = - \sum_{t=1}^{T} \Lambda (\beta x_t) x_t^2,$$

the hessian function,

where

$$\Lambda (x) = \frac{e^x}{(1 + e^x)^2}.$$

Now, we find the limit distributions of the score function $S_n (\beta)$ and the hessian function $J_n (\beta)$ at the true parameter $\beta_0 = 0$. For this, let $e_t = y_t - \Lambda (0) = e_t$.
\[ 1 \{ \varepsilon_t < 0 \} - \frac{1}{2}. \] Then, \( e_t \sim iid \ 0, \frac{1}{\sqrt{n}} \). By the conventional functional central limit theorem, we have
\[
\frac{1}{\sqrt{n}} \mathbb{P} \left[ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \varepsilon_t \right] \Rightarrow \frac{1}{2} B_c (r) - B_a (r),
\]
where \( B_c (r) \) and \( B_a (r) \) are two independent Brownian motions. Using this, it is easy to find that
\[
\frac{1}{n} S_n (0) = \frac{1}{n} \sum_{t=1}^{n} x_t e_t \Rightarrow \frac{\theta}{2} B_a (r) dB_c (r),
\]
and
\[
\frac{1}{n \theta^2} J_n (0) = - \frac{1}{n \theta^2} \sum_{t=1}^{n} x_t^2 \Rightarrow - \frac{\theta}{4} B_a (r) \, dr.
\] (2)

By the mean value theorem,
\[
0 = \frac{S_n (0) - J_n (0)}{n} = \frac{S_n (0)}{n} \frac{J_n (0)}{n \theta} + \frac{\hat{A} \beta + \hat{A} \beta^+ - J_n (0)}{n \theta} \frac{1}{n \theta},
\]
where \( \beta^+ \) locates between zero and \( \hat{\beta} \). Suppose that for some \( \delta > 0 \) we have
\[
\sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} (J_n (\beta) - J_n (0))^2 = o_p (1).
\] (3)

Then, by Theorem 10.1 in Wooldridge (1994), \( n \hat{\beta} = O_p (1) \) and
\[
\frac{1}{n} S_n (0) + o_p (1)
\]
\[
\Rightarrow 2 \frac{B_a (r)^2}{0} - \frac{B_a (r) dB_c (r)}{0}.
\] (4)

To prove (3), notice by the mean value theorem that
\[
\sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} (J_n (\beta) - J_n (0)) = \sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} (\hat{A} \beta + \hat{A} \beta^+ - J_n (0)) = \sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} \left( x_t \beta \Lambda (0) x_t \right) \right),
\]
where \( \Lambda (x) = \frac{e^x}{(1 + e^x)} - \frac{2 e^x}{(1 + e^x)^2} \). Since \( |\beta| < \frac{1}{n^{1-\delta}} \) and \( \Lambda (x) \leq 3 \),
\[
\sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} x_t^3 \beta \Lambda (0) x_t \leq 3 \sup_{|n^{1-\delta} \beta| \leq 1} \frac{1}{n^{2-2\delta}} x_t^3.
\]
Choose \( 0 < \delta < \frac{1}{6} \). Then, \( \sup_{|n^{1-\delta}\beta| \leq 1} \frac{1}{n^{1-\delta}} \left| \sum_{i=1}^{n} x_i \right|^3 = o_p(1) \), and in consequence,

\[
\sup_{|n^{1-\delta}\beta| \leq 1} \frac{1}{n^{1-\delta}} (J_n(\beta) - J_n(0)) = o_p(1),
\]
as required for (3). As a summary, we have the following theorem.

**Theorem 1** Under the assumptions in the previous section, if \( \beta_0 = 0 \), then

\[
\frac{\hat{\beta}}{\text{m}n} \Rightarrow N(0, \text{var})
\]

where notation \( \equiv \) signifies equivalence in distribution and \( MN \) denotes a mixed normal distribution.

Notice that in a univariate nonstationary binary choice model studied by Park and Phillips (2000), when the true parameter \( \beta_0 \neq 0 \), the MLE is \( n^{1/4} \) consistent and its limit distribution is mixed normal. When the true parameter \( \beta_0 = 0 \), Theorem 1 shows that the MLE has a faster convergence rate \( n \) and its limit distribution is similar to so-called “the unit root distribution”. In our note, since we assume that \( u_t \) and \( \varepsilon_s \) are independent, the limit distribution is also mixed normal.

The conventional \( t \)-statistic in this case is

\[
t = \frac{\hat{\beta}}{\text{m}n \hat{\beta}}
\]

By (3) and (2), we have

\[
- \frac{J_n}{n^2} \hat{\beta} = - \frac{J_n(0)}{n^2} + o_p(1) = \frac{1}{4} \int_{0}^{1} B_u(r)^2 dr.
\]

Thus,

\[
t = \frac{\hat{\beta}}{\text{m}n \hat{\beta}} \Rightarrow N(0, 1).
\]

Summarizing this, we have the following corollary.

**Corollary 2** Under the assumptions in the previous section, if \( \beta_0 = 0 \), then

\[
t = \frac{\hat{\beta}}{\text{m}n \hat{\beta}} \Rightarrow N(0, 1)
\]

as required for (3).
References
