Economics Letters 143 (2016) 94–98

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Significance test in nonstationary multinomial logit model

Chia-Shang J. Chu[a,b,*], Nan Liu[c], Lina Zhang[c]

[a] HSBC, Peking University, China
[b] Research Institute of Economics and Management, Southwest University of Finance and Economics, China
[c] National School of Development, Peking University, China

HIGHLIGHTS

- The asymptotic distribution of a LM test in a nonstationary multinomial logit model when the qualitative response is serially correlated is derived.
- Significance test ignoring the serial correlation in qualitative response in a nonstationary multinomial logit model can result in spurious logit link.

ARTICLE INFO

Article history:
Received 20 October 2015
Received in revised form
20 February 2016
Accepted 28 March 2016
Available online 31 March 2016

JEL classification:
C11
C12

Keywords:
Nonstationary multinomial logit
Discrete autoregression
Significance test

ABSTRACT

We derive the asymptotic distribution of the overall significance/LM test in a multinomial logit model with nonstationary covariates when the qualitative response is serially correlated and show that conventional significance tests can result in spurious logit link due to its wrong test size.

1. Introduction

Time series models for binary dependent variables are widely used in many empirical studies. The 0–1 event indicator, say financial crisis or middle income trap, is often modeled with some relevant macroeconomic fundamentals such as GDP or investment. Due to nonstationary covariates, Park and Phillips (2000) develop new limit theories for the maximum likelihood estimator (MLE) in nonstationary binary choice models. Their main finding is that the limiting distribution of the MLE is mixed normal in all directions which ensures the Wald statistics still obey the chi-square distribution asymptotically. Hence standard statistical inference can proceed in the usual manner. Two extensions of the binary response models with nonstationary covariates by Park and Phillips (2000) are available in the literature, Hu and Phillips (2004) study nonstationary ordered response models and Jin (2009) investigates discrete choice models in nonstationary panels.

Results in the aforementioned papers are derived under the assumption of correctly specified models $y_t = F(x_t\beta)$, with the true parameter $\beta \neq 0$. In empirical studies, however, the situation is that we are not sure if there is, say a logit link with a set of nonstationary covariates. It is useful to consider the distribution of MLE under $\beta = 0$ so that the overall significance test can be conducted. Guerre and Moon (2002) consider the case of single covariate and test $H_0 : \beta = 0$ under the assumption that the binary time series is independently identically distributed. They prove that the t-statistics obey standard normal distribution asymptotically. While this is convenient, a discrete time series is often autocorrelated. In fact the binary time series such as financial crisis indicator or monetary decisions are very likely to be serially correlated. We show that failure to take into account the serial correlation can result in spurious logit link in empirical applications. Consequently, some nonstationary covariates may be mistakenly regarded to have the event predictability. Riddel (2000) documented such a spurious problem but offers no theoretical justifications.

* Corresponding author at: HSBC, Peking University, China.
E-mail address: cchu@phbs.pku.edu.cn (C.-S.J. Chu).

http://dx.doi.org/10.1016/j.econlet.2016.03.022
0165-1765/

In this paper we consider the overall significance test in a nonstationary unordered logit model when the qualitative response is serially correlated. Such a serially correlated response is characterized by a discrete autoregression (DAR) process introduced by Jacobs and Lewis (1978). The advantage of this particular characterization is that it allows us to quantify the effect of serial correlation on the overall significance test and it enables us to derive the analytically tractable limiting null distribution.

2. Main results

We use the first order discrete autoregression to characterize the discrete time series.

$$y_t = a_t y_{t-1} + (1-a_t)u_t, \tag{2.1}$$

where a_t is Bernoulli distributed with $P(a_t = 1) = p$, u_t is a discrete distribution with $P(u_t = j) = q_j, j = 0, 1 \ldots J$ and a_t and u_t are independent.

Simple computation shows that $P(y_t = j) = q_j$ and (2.1) is also a Markov chain with transition probabilities

$$\begin{aligned} P(y_t = j|y_{t-1} = j) &= P(a_t = 1) + P(a_t = 0, u_t = j) \\ &= p + (1-p)q_j \end{aligned}$$

$$P(y_t = l|y_{t-1} = j) = P(a_t = 0, u_t = l) = (1-p)q_l, \quad l \neq j.$$

Define a binary random variable as

$$y_{tj} = \begin{cases} 1, & \text{if } y_t = j \\ 0, & \text{if } y_t \neq j. \end{cases} \tag{2.2}$$

The autocorrelation for $\{y_{tj}, t = 1, \ldots, T\}$ is

$$\begin{aligned} Corr(y_{tj}, y_{t-1,j}) &= \frac{Cov(y_{tj}, y_{t-1,j})}{\sqrt{Var(y_{tj})}\sqrt{Var(y_{t-1,j})}} \\ &= \frac{[p + (1-p)q_j]q_j - q_j^2}{q_j(1-q_j)} = p. \end{aligned}$$

Thus, each $\{y_{tj}\}$, $j = 1, \ldots, J$, is also equivalent to the first order DAR.

$$y_{tj} = a_t y_{t-1,j} + (1-a_t)u_{tj}, \tag{2.3}$$

where

$$u_{tj} = \begin{cases} 1, & \text{if } u_t = j \\ 0, & \text{if } u_t \neq j \end{cases} \quad P(u_{tj} = 1) = q_j.$$

Consider now an unordered multinomial logit model with nonstationary covariates.

$$y_{tj} = \Lambda(\alpha_j + x_t'\beta_j) + \varepsilon_{tj}, \quad \Lambda(x) = \frac{e^x}{1+e^x}. \tag{2.4}$$

$$x_t = x_{t-1} + v_t, \tag{2.5}$$

where v_t is any weakly dependent stationary process that admits the functional central limit theorem.

Let MLE $\hat{\alpha}$ and $\hat{\beta}$ be the MLE for $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_J]'$ and $\beta = [\beta_1', \beta_2', \ldots, \beta_J']'$ obtained from maximizing the likelihood

$$l_T(\alpha, \beta) = \sum_{t=1}^{T}\sum_{j=1}^{J} y_{tj} \log \Lambda(\alpha_j + x_t'\beta_j).$$

The score and Hessian matrix are given in Box I:

Under $H_0 : \beta_1 = \beta_2 = \cdots = \beta_J = 0$,

$$S_T\begin{pmatrix}\tilde{\alpha}\\0\end{pmatrix} = \begin{bmatrix} \sum_{t=1}^{T}(y_{t1} - \tilde{\Lambda}_1) \\ \vdots \\ \sum_{t=1}^{T}(y_{tJ} - \tilde{\Lambda}_J) \\ \sum_{t=1}^{T}(y_{t1} - \tilde{\Lambda}_1)x_t \\ \vdots \\ \sum_{t=1}^{T}(y_{tJ} - \tilde{\Lambda}_J)x_t \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T}\tilde{e}_{t1} \\ \vdots \\ \sum_{t=1}^{T}\tilde{e}_{tJ} \\ \sum_{t=1}^{T}\tilde{e}_{t1}x_t \\ \vdots \\ \sum_{t=1}^{T}\tilde{e}_{tJ}x_t \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sum_{t=1}^{T}\tilde{e}_{t1}x_t \\ \vdots \\ \sum_{t=1}^{T}\tilde{e}_{tJ}x_t \end{bmatrix}$$

$J_T\begin{pmatrix}\tilde{\alpha}\\0\end{pmatrix} == \begin{bmatrix} \tilde{J}_T^{\alpha\alpha} & \tilde{J}_T^{\beta\alpha'} \\ \tilde{J}_T^{\beta\alpha} & \tilde{J}_T^{\beta\beta} \end{bmatrix}$, $\tilde{J}_T^{\alpha\alpha}$ is $J_T^{\alpha\alpha}$ matrix evaluated at $\begin{pmatrix}\tilde{\alpha}\\0\end{pmatrix}$ and similarly for other *J-tilde* matrix. Furthermore,

$$J_T\begin{pmatrix}\tilde{\alpha}\\0\end{pmatrix} = -\begin{bmatrix} \tilde{A}\otimes T & \tilde{A}\otimes\sum_{t=1}^{T}x_t' \\ \tilde{A}\otimes\sum_{t=1}^{T}x_t & \tilde{A}\otimes\sum_{t=1}^{T}x_tx_t' \end{bmatrix},$$

where $\tilde{A} = -\begin{bmatrix} -\tilde{\Lambda}_1(1-\tilde{\Lambda}_1) & \tilde{\Lambda}_1\tilde{\Lambda}_2 & \cdots & \tilde{\Lambda}_1\tilde{\Lambda}_J \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\Lambda}_J\tilde{\Lambda}_1 & \tilde{\Lambda}_J\tilde{\Lambda}_2 & \cdots & -\tilde{\Lambda}_J(1-\tilde{\Lambda}_J) \end{bmatrix}$,

$\tilde{\alpha} = (\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_J)'$, $\tilde{\Lambda}_j = \Lambda(\tilde{\alpha}_j)$, $j = 1, \ldots, J$, and $\tilde{\alpha}_j$ is the estimator of α_j under H_0.

Since $E(y_{tj}|I_{t-1,j}) = py_{t-1,j} + (1-p)q_j$,

$$y_{tj} = c_j + py_{t-1,j} + \eta_{tj}, \tag{2.6}$$

where $c_j = (1-p)q_j$, $E(\eta_{tj}) = 0$ and $var(\eta_{tj}) \equiv \sigma_j^2 = (1-p^2)q_j(1-q_j)$, $j = 1, 2, \ldots, J$.

Apply the lag operator "L" to rewrite (2.6) as $(1-pL)y_{tj} = c_j + \eta_{tj}$, it follows that

$y_{tj} = (1-pL)^{-1}c_j + (1-pL)^{-1}\eta_{tj} = q_j + (1-pL)^{-1}\eta_{tj}$ and $e_{tj} = y_{tj} - q_j = (1-pL)^{-1}\eta_{tj}$.

Thus $e_{tj} = (1-pL)^{-1}\eta_{tj}$ is an autoregressive process. Standard application of the functional central limit theorem yields

$$\begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]}e_{tj} \\ \frac{x_{[Tr]}}{\sqrt{T}} \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{\sigma_j}{1-p}B_j(r) \\ V^{1/2}B_u(r) \end{pmatrix},$$

$$V = \lim_{T\to\infty}\frac{1}{T}E\left[\left(\sum_{t=1}^{T}v_t\right)\left(\sum_{t=1}^{T}v_t\right)'\right], \tag{2.7}$$

where $B_u(r)$ and $B_j(r)$, $j = 1, 2, \ldots, J$ are mutually independent standard Brownian motion.

We shall apply some results on weak convergence that are well documented in the literature.

$$\frac{1}{T}\sum_{t=1}^{T}x_te_{tj} \Rightarrow \frac{\sigma_j}{1-p}V^{1/2}\int_0^1 B_u(r)dB_j(r). \tag{2.8}$$

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T}x_t \Rightarrow V^{1/2}\int_0^1 B_u(r)dr. \tag{2.9}$$

$$\frac{1}{T^2}\sum_{t=1}^{T}x_tx_t' \Rightarrow V^{1/2}\int_0^1 B_u(r)B_u'(r)drV^{1/2}. \tag{2.10}$$