应用概率统计 第 32 卷
第 3 期 2016 年 6 月

Chinese Journal of Applied Probability and Statistics
Jun., 2016, Vol. 32, No. 3, pp. 301-312
doi: 10.3969/j.issn.1001-4268.2016.03.006

Testing for Significance in Binary Choice Model with Stochastic Trend Process*

XU Peng
(*School of Economics, Nankai University, Tianjin, 300071, China*)

WANG Lujun
(*School of Public Finance and Taxation, Nanjing University of Finance and Economics, Nanjing, 210000, China*)

YAN Zichun
(*Economics and Management College, Civil Aviation University of China, Tianjin, 300300, China*)

Abstract: This paper investigates the test of significance for the binary choice model with stochastic trend process. The results show that when the true parameter vector is zero, the limiting distribution of the t statistic follows standard normal distribution. The joint significance test statistics Wald, LM and LR are asymptotically equivalent and have a Chi-square limiting distribution.

Keywords: binary choice model; stochastic trend process; significance test

2010 Mathematics Subject Classification: 62G10; 62G20

§1. Introduction

Nonstationary binary choice models have attracted lots of attention during the process of the financial and macroeconomic time series modeling. Park and Phillips [1] studies the asymptotic theory about logit and probit models with multiple explanatory variables following integrated processes. Their theory fundamentally based on the assumption that $\|\beta\| \neq 0$, where β is the true parameter vector and $\|\cdot\|$ is the Euclidean norm. In practice, however, we may not have prior information about this assumption. Guerre and Moon [2] and Mao [3] study the significance test for $\|\beta\| = 0$. It is worth noting that they assume the explanatory variables following unit root processes, while numerous types of time series in real life otherwise exhibit stochastic trend characteristics. Therefore, it should test $\|\beta\| = 0$ before applying binary choice model with multiple explanatory variables generated from stochastic trend process.

*The project was supported by China Scholarship Council and China's Civil Aviation Social Economic Contribution Evaluation Theory and Empirical Research (3122015Z008).

Received September 28, 2014. Revised June 23, 2015.

This paper is structured as follows. Section 2 outlines the model and assumptions. Section 3 gives the main results of the limiting distribution of the t, Wald, LM and LR statistics and also presents some Monte Carlo simulation results. For better understanding the theoretical research, Section 4 gives a simple empirical case. Some useful proofs of the main theorems are given in Appendix.

§2. Model and Assumptions

The classical binary choice model is defined as

$$y_t = 1\{y_t^* > 0\} \quad \text{and} \quad y_t^* = x_t'\beta - \varepsilon_t \qquad \text{for } t = 1, 2, \ldots, T, \tag{1}$$

where the error term ε_t is widely assumed to follow logistic or normal distribution. To derive the limiting distribution results, we make the following assumptions regarding the data generating process x_t.

Assumption 1 Let x_t follows a k-dimensional stochastic trend process,

$$x_t = \alpha + x_{t-1} + \mu_t, \tag{2}$$

where the true value of α is not zero and $\mu_t \sim$ i.i.d.$(0, \Omega_\mu)$. Then (2) can be written as

$$x_t = x_0 + \alpha t + (\mu_1 + \mu_2 + \cdots + \mu_t) = x_0 + \alpha t + \xi_t$$

for simplicity, set $\xi_0 = 0$. By functional central limit theorem (FCLT) we have

$$T^{-1/2}\xi_{[Tr]} \Rightarrow B_\mu(r), \tag{3}$$

where $[\cdot]$ denotes the floor function, $\Rightarrow$ signifies convergence in distribution, and $B_\mu(r)$ is the Brownian motion with positive definite variance matrix Ω_μ.

Since ML estimator involves nonlinear functions of the process x_t, we make some assumptions about the distribution function F and density $\dot{F}$ of ε_t.

Assumption 2 The functions $G(s)$ and $K(s)$ satisfy

$$G(s) = \dot{F}(s)/F(s)(1 - F(s)), \qquad K(s) = G(s)\dot{F}(s) = G(s)^2 F(s)(1 - F(s)),$$

where $F(s)$ is three times differentiable so that the first derivatives $\dot{F}(s)$, $\dot{G}(s)$ and the second derivatives $\ddot{F}(s)$, $\ddot{G}(s)$ all exist. Further: $\ddot{G}(s)$ and $\dot{K}(s)$ is bounded.

Following similar techniques as in [1] and [3], it can be shown that both logit and probit models satisfy Assumption 2.

To test $\|\beta\|=0$, let $e_t=y_t-F(0)$. For logit and probit models, $e_t=y_t-0.5$. It is easy to verify that $e_t\sim\text{i.i.d.}(0,0.25)$. By FCLT we have

$$T^{-1/2}\sum_{t=1}^{[Tr]}e_t\Rightarrow B_e(r), \tag{4}$$

where $B_e(r)$ is the Brownian motion with variance 0.25 and $B_e(r)$, $B_\mu(r)$ are independent.

§3. Main Results

The log likelihood of the model (1) has the form

$$\ln L_T(\beta)=\sum_{t=1}^{T}y_t\ln F(x_t'\beta)+\sum_{t=1}^{T}(1-y_t)\ln F(x_t'\beta),$$

and under Assumption 2 we can write the score $S_T(\beta)$ and hessian $H_T(\beta)$ as

$$S_T(\beta)=\sum_{t=1}^{T}G(x_t'\beta)x_t[y_t-F(x_t'\beta)],$$
$$H_T(\beta)=-\sum_{t=1}^{T}K(x_t'\beta)x_tx_t'+\sum_{t=1}^{T}\dot G(x_t'\beta)x_tx_t'[y_t-F(x_t'\beta)].$$

Let $\widehat\beta$ be the MLE of model (1). Then we have the following lemma:

Lemma 3 Let Assumptions 1 and 2 hold. If

$$\sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1}\Big\|\frac{H_T(\beta)-H_T(0)}{T^{3-3\delta}}\Big\|=o_p(1) \tag{5}$$

for some $\delta>0$, then we have

$$T^{3/2}\widehat\beta\Rightarrow\Big(\frac{K(0)}{3}\alpha\alpha'\Big)^{-1}\Big(G(0)\alpha\int_0^1 r\mathrm{d}B_e(r)\Big)\equiv\mathrm{MN}_k\Big(0,\frac{3G(0)^2}{4K(0)^2}(\alpha\alpha')^{-1}\Big), \tag{6}$$

where $\equiv$ denotes distributional equivalence and MN_k signifies a k-dimensional mixed normal distribution.

Note that in this case the convergence rate of MLE is $T^{3/2}$, which is faster than that of [2] and [3]. In the Appendix we prove that, for logit and probit models, there exists $\delta>0$ that makes the condition (5) satisfied. Therefore, (6) can be applied to the analysis of single significance t test and joint significance test such as Wald, LM and LR tests. Based on the results above, we have the following theorem.

Theorem 4 Let Assumptions 1 and 2 hold. For the logit or probit model, we have

$$t_{\beta_i} \Rightarrow N(0,1), \qquad i = 1, 2, \ldots, k,$$

$$\text{Wald} \approx \text{LM} \approx \text{LR} \Rightarrow \chi^2(k), \text{ as } T \to \infty \text{ and under the null hypothesis } \|\beta\| = 0.$$

The theorem suggests that the t, Wald, LM and LR statistics can be used in the usual way for testing the significance of the nonstationary logit and probit models in large samples. Besides, it is worth noting that the Wald, LM and LR statistics are asymptotically equivalent and they all have the $\chi^2(k)$ limiting distribution, where k denotes the number of restrictions. To further illustrate this theorem, we consider a simple Monte Carlo simulation performed with 10 000 repetitions. Firstly, we consider the property of the four statistics in finite samples. Under $\|\beta\| = 0$, the independent variables' DGPs are assumed to be $x_t = \alpha + x_{t-1} + \mu_t$, $k = 3$, $\mu_t \sim \text{i.i.d.}(0, I_3)$, $x_0 = 0$, $\alpha = \{0.02, 0.05, 0.08\}$ and $T = \{50, 100, 500, 1\,000\}$. Figure 1 gives the statistical distribution of the t statistic of the first independent variable and can be well approximated by $N(0,1)$ distribution in different samples. Figure 2 shows that the Wald, LM and LR statistics are asymptotically equivalent and can be well approximated by the $\chi^2(3)$ distribution. Therefore, the standard normal and Chi-Square distribution tables can be used to test for the significance in finite samples.

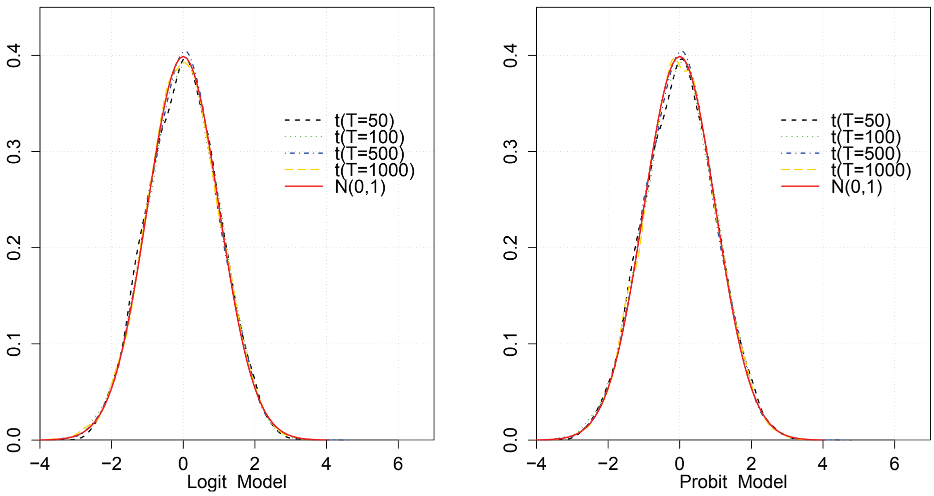

Figure 1 **Statistical distribution of the t statistics in finite samples**

Then we consider the actual size of the four statistics. The simulation setting are the same as the above. Table 1 reports the actual sizes of the four statistics under different nominal sizes. We can see that the four statistic has actual sizes reasonably close to the nominal sizes.

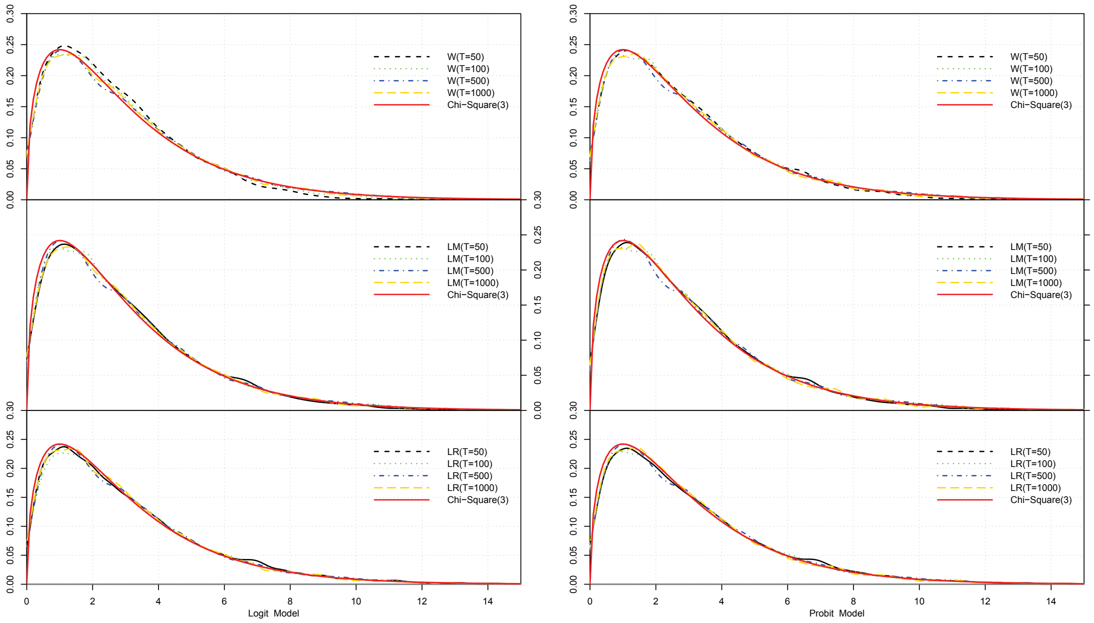

Figure 2 Statistical distribution of the Wald, LM and LR statistics in finite samples

Table 1 Actual size of the four statistics

Statistic	T	Logit Model			Probit Model		
		0.01	0.05	0.1	0.01	0.05	0.1
t	50	0.0072	0.0498	0.1022	0.0075	0.0517	0.1027
	100	0.0077	0.0455	0.1022	0.0095	0.0501	0.1043
	500	0.0113	0.0538	0.1013	0.0094	0.0485	0.0994
	1 000	0.0106	0.0515	0.0997	0.0090	0.0464	0.0970
Wald	50	0.0019	0.0255	0.0723	0.0034	0.0324	0.0844
	100	0.0043	0.0388	0.0833	0.0066	0.0440	0.0929
	500	0.0096	0.0494	0.0946	0.0090	0.0511	0.1021
	1 000	0.0092	0.0468	0.0958	0.0096	0.0466	0.0961
LR	50	0.0106	0.0559	0.1075	0.0106	0.0559	0.1075
	100	0.0110	0.0505	0.0993	0.0110	0.0505	0.0993
	500	0.0108	0.0526	0.0986	0.0108	0.0526	0.0986
	1 000	0.0102	0.0486	0.0972	0.0102	0.0486	0.0972
LM	50	0.0075	0.0445	0.0981	0.0075	0.0445	0.0981
	100	0.0074	0.0472	0.0966	0.0074	0.0472	0.0966
	500	0.0105	0.0484	0.1036	0.0105	0.0484	0.1036
	1 000	0.0110	0.0485	0.0983	0.0110	0.0485	0.0983

Lastly, we examine the power of the four statistic. For simplicity, we only consider four cases of DGPs for y^*. In Case 1 $\beta_0 = \{0.25, 0, -0.01\}$, In Case 2 $\beta_0 = \{0.25, 0.1, -0.01\}$, In Case 3 $\beta_0 = \{0.5, 0, -0.01\}$ and In Case 4 $\beta_0 = \{0.5, 0.1, -0.01\}$. Other simulation settings are the same as above. Table 2 shows that the four statistic has substantially higher power at nominal size of 5%.

Table 2 Power test of the four statistics

Statistic	T	Logit Model				Probit Model			
		Case1	Case2	Case3	Case4	Case1	Case2	Case3	Case4
t	50	0.4565	0.5077	0.6585	0.671	0.3867	0.4406	0.5969	0.6125
	100	0.6139	0.6387	0.7425	0.7481	0.526	0.562	0.6839	0.6951
	500	0.6228	0.636	0.7133	0.717	0.5411	0.563	0.6515	0.6576
	1 000	0.8233	0.8161	0.8592	0.8647	0.7891	0.7966	0.8339	0.8400
Wald	50	0.8392	0.8472	0.8869	0.8933	0.6789	0.703	0.77	0.7798
	100	0.9475	0.9498	0.9672	0.9666	0.9017	0.9081	0.9336	0.9342
	500	0.993	0.9912	0.9931	0.9931	0.9951	0.9936	0.9947	0.9931
	1 000	0.9985	0.9987	0.9994	0.9991	0.998	0.9984	0.9992	0.9992
LR	50	0.9239	0.9325	0.9696	0.9702	0.9236	0.9326	0.9696	0.9701
	100	0.9696	0.9744	0.9872	0.9883	0.9694	0.9738	0.9872	0.9883
	500	0.9992	0.999	0.9995	0.9996	0.9991	0.9989	0.9993	0.9996
	1 000	0.9991	0.9995	0.9996	0.9996	0.999	0.9995	0.9996	0.9996
LM	50	0.968	0.9724	0.9989	0.9987	0.968	0.9724	0.9989	0.9987
	100	0.9983	0.9991	1.0000	1.0000	0.9983	0.9991	1.0000	1.0000
	500	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1 000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

§4. Empirical Case

For better understanding the theoretical research, we consider an empirical case. This case mainly concentrates on China's monetary policy stance and dynamic adjustment. Following Taylor's model and the subsequent studies in this area, we specified the China's monetary reaction as:

$$\mathrm{MP}_t = \alpha * (\mathrm{GDP}_t - \mathrm{GDP}_t^*) + \beta * (\mathrm{CPI}_t - \mathrm{CPI}_t^*) + \varepsilon_t, \tag{7}$$

Where MP_t is a monetary policy stance indicator which is mainly derived from [4]'s study, except that we construct two monetary stance indices in this case. In binary choice model, we set 0 to indicate easy monetary policy and 1 tight monetary policy. GDP_t and CPI_t are actual GDP growth and CPI inflation, while GDP_t^* and CPI_t^* are growth and inflation targets respectively. The empirical data is from 2001Q1 to 2014Q4.

In classical time series analysis, before building a model, we should test the stationarity of the variables. If a variable is not stationary, we should take difference or detrend it. By using ADF unit root test, we find that the term $(\mathrm{GDP}_t - \mathrm{GDP}_t^*)$ follows stochastic trend process. However, in binary choice model, we do not need to concern about it and still can use original data to do the regression, as proved in theoretical part above. Table 3 gives the estimation results. The values of LR, LM and Wald are larger, showing that the whole model fits better. Both α and β are positive, implying that the People's bank of China will tighten monetary stance if economic growth and inflation are above their targets.

Table 3 Empirical results of binary choice models

	Logit Model	Probit Model
GDP GDP*	0.273**	0.155*
	-2.103	-1.69
CPI CPI*	1.076***	0.629***
	-3.389	-2.763
LR	23.949	23.533
LM	17.96	17.96
Wald	11.966	7.776

Note: ***, ** and * denote significance at the 1%, 5% and 10% significance level, respectively. The value in parentheses is the standard error. All data is available upon request.

Appendix: Mathematical Proofs

Proof of Lemma 3 Theorem 10.1 in [5] is suitable to solve the present problem. To better use the theorem, we verify the key conditions (iii) and (iv) of this theorem.

By mean value theorem we have

$$0 = S_T(\widehat{\beta}) = S_T(0) + H_T(\beta^+)(\widehat{\beta} - 0), \tag{8}$$

where β^+ is the mean values. To arrive at nondegenerate limiting distribution, we premultiply the (8) by an inverse of scaling matrix $D_T^{1/2}$ and write the result as

$$\begin{aligned} 0 = {} & D_T^{-1/2} S_T(0) + \big[D_T^{-1/2} H_T(0) D_T^{-1/2}\big] D_T^{1/2}\widehat{\beta} \\ & + T^{-3\delta}\big(C_T^{-1/2}[H_T(\beta) - H_T(0)]C_T^{-1/2}\big) D_T^{1/2}\widehat{\beta}, \end{aligned} \tag{9}$$

where $C_T = D_T T^{-3\delta}$ for some $\delta > 0$. Next find out the scaling matrix $D_T^{1/2}$. Under $\|\beta\| = 0$ and the Assumptions 1 and 2, we have

$$S_T(0) = G(0)\sum_{t=1}^{T} x_t e_t = G(0)\Big(\sum_{t=1}^{T} x_0 e_t + \sum_{t=1}^{T} \alpha t e_t + \sum_{t=1}^{T} \xi_t e_t\Big), \tag{10}$$

$$H_T(0) = -K(0)\sum_{t=1}^{T} x_t x_t' = -K(0)\sum_{t=1}^{T} (x_0 + \alpha t + \xi_t)(x_0 + \alpha t + \xi_t)'. \tag{11}$$

We rewrite the (10) by order in probability

$$S_T(0) = O_p(T^{1/2}) + O_p(T^{3/2}) + O_p(T),$$

the second term in above equation asymptotically dominates the other two components, so we have

$$T^{-3/2} S_T(0) \Rightarrow G(0)\alpha \int_0^1 r \mathrm{d}B_e(r). \tag{12}$$

Similarly, we expand the (11) totally

$$H_T(0) = -K(0)\sum_{t=1}^{T} (x_0 x_0' + x_0\alpha' t + x_0\xi_t' + \alpha x_0' t + \alpha\alpha' t^2 + \alpha\xi_t' t + \xi_t x_0' + \xi_t \alpha' t + \xi_t \xi_t').$$

It is easy to find that the time trend term t^2 dominates other components and the order in probability is $O_p(T^3)$, so we have the convergence of $H_T(0)$

$$T^{-3} H_T(0) \xrightarrow{p} -K(0)\alpha\alpha'/3. \tag{13}$$

Thus we find out the scaling matrix and let $D_T^{1/2} = T^{3/2} I_k$, where I_k is the k by k identity matrix. Consequently, $C_T D_T^{-1} = o(1)$ as $T \to \infty$, which shows that condition (iii) (a) holds. Besides, condition (iv) is ensured by (12) and (13). To obtain (6), it is sufficient to show that

$$T^{3/2}\widehat{\beta} = \big[-T^{-3} H_T(0)\big]^{-1}\big[T^{-3/2} S_T(0)\big] + o_p(1),$$

which follows from (9) if the last term of (9) is $o_p(1)$. This will be so, if the condition (iii) (b) holds. To show this is so, we need to proof (5). Note that

$$\begin{aligned}
&\sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \Big\| \frac{H_T(\beta) - H_T(0)}{T^{3-3\delta}} \Big\| \\
&= \sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \left\| \frac{\sum_{t=1}^{T} \big\{\dot{G}(x_t'\beta)[y_t - F(x_t'\beta)] - [K(x_t'\beta) - K(0)]\big\} x_t x_t'}{T^{3-3\delta}} \right\|.
\end{aligned} \tag{14}$$

Since for any square matrix $\|A\| = \big(\sum\limits_i \sum\limits_j |A_{ij}|^2\big)^{1/2} \leqslant k \max\limits_{i,j} |A_{ij}|$ and by absolute value inequality, (14) can be written as

$$\begin{aligned}
&\sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \left\| \frac{H_T(\beta) - H_T(0)}{T^{3-3\delta}} \right\| \\
&\leqslant k \max_{i,j} \sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \left| \frac{\sum\limits_{t=1}^{T} \big\{\dot{G}(x_t'\beta)[y_t - F(x_t'\beta)] - [K(x_t'\beta) - K(0)]\big\} x_{it}x_{jt}}{T^{3-3\delta}} \right| \\
&\leqslant k \max_{i,j} \sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \left(\left| \frac{\sum\limits_{t=1}^{T} \dot{G}(x_t'\beta)[y_t - F(x_t'\beta)] x_{it}x_{jt}}{T^{3-3\delta}} \right| + \left| \frac{\sum\limits_{t=1}^{T} [K(x_t'\beta) - K(0)] x_{it}x_{jt}}{T^{3-3\delta}} \right| \right).
\end{aligned}$$

By mean value theorem and Cauchy-Schwarz inequality $\|x_t'\beta\| \leqslant \|x_t\|\,\|\beta\|$, the right hand side of the inequality above can be bounded by

$$\begin{aligned}
k \max_{i,j} \sup_{\|T^{1.5-1.5\delta}\beta\|\leqslant 1} \Bigg(&\frac{\sum\limits_{t=1}^{T} |\ddot{G}(x_t'\beta^+)|\,|y_t - F(x_t'\beta)|\,|x_{it}x_{jt}|\,\|x_t\|\,\|\beta\|}{T^{3-3\delta}} \\
&+ \frac{\sum\limits_{t=1}^{T} |\dot{K}(x_t'\beta^+)|\,|x_{it}x_{jt}|\,\|x_t\|\,\|\beta\|}{T^{3-3\delta}} \Bigg).
\end{aligned}$$

Since $\|x_t\| \leqslant \sqrt{k} \max\limits_l |x_{lt}|$, $\|\beta\| \leqslant T^{-1.5+1.5\delta}$ and $|y_t - F(x_t'\beta)| \leqslant 2$, meanwhile $\ddot{G}(s)$ and $\dot{K}(s)$ is bounded under Assumption 2, so the above inequality can be further bounded by

$$\leqslant \frac{ck^{3/2} \sum\limits_{t=1}^{T} |x_{it}x_{jt}x_{lt}|}{T^{9/2-9\delta/2}},$$

where the c is a generic positive constant. Since $\sum\limits_{t=1}^{T} 3|x_{it}x_{jt}x_{lt}| \leqslant \sum\limits_{t=1}^{T} (|x_{it}|^3 + |x_{jt}|^3 + |x_{lt}|^3)$, to proof (5), it suffices to verify $\sum\limits_{t=1}^{T} |x_{it}|^3 / T^{9/2-9\delta/2} = o_p(1)$. Because $|x_{it}|^3$ is regular, we can use the Theorem 3.2 of [6], who proofed that $T^{-1-k/2} \sum\limits_{t=1}^{T} |x_{it}|^k = O_p(1)$ for $k \geqslant 0$. Choose $0 < \delta < 4/9$, we can easily deduce that $\sum\limits_{t=1}^{T} |x_{it}|^3 / T^{9/2-9\delta/2} = o_p(1)$. Hence (5) holds. we complete the proof of Lemma 3. □

Proof of Theorem 4 The t statistic is characterized by

$$t_{\widehat{\beta}_i} = \frac{\widehat{\beta}_i}{\sqrt{(-H_T(\widehat{\beta}))^{-1}_{ii}}} = \frac{T^{3/2}\widehat{\beta}_i}{\sqrt{(-T^{-3}H_T(\widehat{\beta}))^{-1}_{ii}}}.$$

By (5), (6) and (13), we have

$$t_{\widehat{\beta}_i} \Rightarrow \frac{\Big\{(K(0)\alpha\alpha'/3)^{-1}\Big(G(0)\alpha\displaystyle\int_0^1 r\mathrm{d}B_e(r)\Big)\Big\}_i}{\sqrt{(K(0)\alpha\alpha'/3)^{-1}_{ii}}}.$$

For logit and probit models, $G^2(0) = 4K(0)$. Consequently, we have

$$t_{\widehat{\beta}_i} \Rightarrow N(0,1), \qquad i = 1, 2, \ldots, k.$$

Next we consider testing $\|\beta\| = 0$ by the Wald, LM and LR statistics. Under $\|\beta\| = 0$, the Wald statistic can be written as

$$\text{Wald} = -\widehat{\beta}' H_T(\widehat{\beta})\widehat{\beta} = (T^{3/2}\widehat{\beta})'(-T^{-3}H_T(\widehat{\beta}))(T^{3/2}\widehat{\beta}).$$

By (5), (6) and (13), we have

$$\text{Wald} = \frac{G(0)^2}{4K(0)}(T^{3/2}\widehat{\beta})'\Big[\frac{3G(0)^2}{4K(0)^2}(\alpha\alpha')^{-1}\Big]^{-1}(T^{3/2}\widehat{\beta}) + o_p(1).$$

Since $G^2(0) = 4K(0)$, thus we can obtain the limiting distribution of Wald statistic

$$\text{Wald} \Rightarrow \Big(G(0)\alpha\int_0^1 r\mathrm{d}B_e(r)\Big)'\Big(\frac{K(0)}{3}\alpha\alpha'\Big)^{-1}\Big(G(0)\alpha\int_0^1 r\mathrm{d}B_e(r)\Big) \equiv \chi^2(k). \tag{15}$$

The LM statistic is defined by

$$\text{LM} = S_T(\widetilde{\beta})'[-H_T(\widetilde{\beta})]^{-1}S_T(\widetilde{\beta}),$$

where $\widetilde{\beta}$ is the restricted estimator vector. Under $\|\beta\| = 0$, it is easy to verify that $\widetilde{\beta} = 0$, so we can rewrite the LM as

$$\text{LM} = S_T(0)'[-H_T(0)]^{-1}S_T(0) = (T^{-3/2}S_T(0))'[-T^{-3}H_T(0)]^{-1}(T^{-3/2}S_T(0)).$$

By (12), we have

$$T^{-3/2}S_T(0) \Rightarrow G(0)\alpha\int_0^1 r\mathrm{d}B_e(r) \equiv \text{MN}_k(0, G(0)^2\alpha\alpha'/12),$$

Then, by (5), (6) and (13), we obtain

$$\text{LM} = \frac{G(0)^2}{4K(0)}(T^{-3/2}S_T(0))'[G(0)^2\alpha\alpha'/12]^{-1}(T^{-3/2}S_T(0)) + o_p(1).$$

Since $G^2(0) = 4K(0)$, we can obtain the limiting distribution of LM statistic by simple calculation

$$\text{LM} \Rightarrow \Big(G(0)\alpha\int_0^1 r\mathrm{d}B_e(r)\Big)'\Big(\frac{K(0)}{3}\alpha\alpha'\Big)^{-1}\Big(G(0)\alpha\int_0^1 r\mathrm{d}B_e(r)\Big) \equiv \chi^2(k). \tag{16}$$

The likelihood ratio test is defined by

$$\mathrm{LR} = -2(\ln L_T(\widetilde{\beta}) - \ln L_T(\widehat{\beta})).$$

To get the limiting distribution of LR, take a second-order Taylor's approximation of $\ln L_T(\widetilde{\beta})$ around $\ln L_T(\widehat{\beta})$

$$\ln L_T(\widetilde{\beta}) = \ln L_T(\widehat{\beta}) + S_T(\widehat{\beta})(\widehat{\beta} - \widetilde{\beta}) + 1/2(\widehat{\beta} - \widetilde{\beta})' H_T(\overline{\beta})(\widehat{\beta} - \widetilde{\beta}),$$

where $S_T(\widehat{\beta}) = 0$, $\overline{\beta}$ is the mean values. Then the above equation can be written as

$$\ln L_T(\widetilde{\beta}) - \ln L_T(\widehat{\beta}) = 1/2(\widehat{\beta} - \widetilde{\beta})' H_T(\overline{\beta})(\widehat{\beta} - \widetilde{\beta}),$$

so we have

$$\mathrm{LR} = (\widehat{\beta} - \widetilde{\beta})' H_T(\overline{\beta})(\widehat{\beta} - \widetilde{\beta})$$

under $\|\beta\| = 0$, we know that $\widetilde{\beta} = 0$. By (5), we obtain

$$\mathrm{LR} = (T^{3/2}\widehat{\beta})'(-T^{-3}H_T(0))(T^{3/2}\widehat{\beta}) + o_p(1).$$

Similar to Wald statistic, we can obtain the limiting distribution of LR statistic

$$\mathrm{LR} \Rightarrow \Big(G(0)\alpha \int_0^1 r\mathrm{d}B_e(r)\Big)' \Big(\frac{K(0)}{3}\alpha\alpha'\Big)^{-1} \Big(G(0)\alpha \int_0^1 r\mathrm{d}B_e(r)\Big) \equiv \chi^2(k). \tag{17}$$

From (15) to (17), we observe that the Wald, LM and LR statistics are asymptotically equivalent and they all have $\chi^2(k)$ limiting distribution, which completes the proof of Theorem 4. □

Acknowledgments Helpful comments and suggestions from the editor and the anonymous referee are greatly appreciated. Thanks to Zhuang J. and Zhao H. M. in writing help.

References

[1] Park J Y, Phillips P C B. Nonstationary binary choice [J]. *Econometrica*, 2000, **68(5)**: 1249–1280.

[2] Guerre E, Moon H R. A note on the nonstationary binary choice logit model [J]. *Econom. Lett.*, 2002, **76(2)**: 267–271.

[3] Mao G Y. Testing for joint significance in nonstationary binary choice model [J]. *Econom. Lett.*, 2014, **122(2)**: 311–313.

[4] Xu P, Xu W G. Monetary policy stance and dynamic adjustment: an empirical study of China monetary policy reports [J]. *J. Finance Economics*, 2015, **30(4)**: 3–16. (Chinese)

[5] Wooldridge J M. Estimation and inference for dependent processes [M] // Engle R F, McFadden D L. *Handbook of Econometrics, Volume 4*. North-Holland: Elsevier Science Publishers, 1994: 2639–2738.

[6] Park J Y, Phillips P C B. Asymptotics for nonlinear transformations of integrated time series [J]. *Econometric Theory*, 1999, **15(3)**: 269–298.

带有随机趋势项的二元选择模型显著性检验研究

徐 鹏

(南开大学经济学院数量经济研究所, 天津, 300071)

汪卢俊

(南京财经大学财政与税务学院, 南京, 210000)

严子淳

(中国民航大学经济与管理学院, 天津, 300300)

摘 要: 本文主要对非平稳二元选择模型的显著性检验进行研究. 研究结果显示, 当真实参数为零时, t统计量依分布收敛于标准正态分布. 同时联合显著性检验统计量Wald、LM和LR渐近相等且依分布收敛于卡方分布.

关键词: 二元离散选择模型; 随机趋势过程; 显著性检验

中图分类号: F224.0