Test of Random Effects with Incomplete Panel Data

K. Oya

Graduate School of Economics, Osaka University, Osaka 560-0043, Japan
(kosuke@econ.osaka-u.ac.jp)

Abstract : This paper examines properties of test statistics for random effects with incomplete panel data. We can divide incomplete panel data into two groups. One group arises from randomly missing or unbalanced data and the other arises from systematically missing data. We focus on the former case. Some statistical properties when there are missing independent (explanatory) variables in regression analysis are well-known. Many procedures such as imputation have been proposed to deal with missing independent variables, but few of them are practical to use in empirical analysis. For the dependent (explained) variable, we may use the procedures developed for limited dependent variable analysis. However, we often face the situation that we are not able to apply such procedures since we do not have any appropriate exogenous variables for the missing dependent variable. A simple approach to treat missing observations is to just discard the missing cases, but such approach may be highly inefficient. In this paper, instead of discarding the missing cases, we consider the missing data to be the outcome of a random variable. The test statistic for random effects with randomly missing panel data is derived. We examine the statistical properties of the derived test statistic and compare it with test statistics derived without randomness. We find that our test statistic is conservative in comparison with the test statistic derived without randomness.

Keywords : Incomplete panel; Random effect; Markov chain; Missing value

1. INTRODUCTION

We rarely have complete panel data although we have fancy statistical techniques. In the case of household surveys, some respondents do not answer all the items of a questionnaire. For stock and corporate bond analysis, several firms drop out from our data list for some reason such as bankruptcy or default. We frequently use a filling in or imputation method for missing values. However, this method can lead to biases in statistical inference when the imputed values are different from the true (unknown) missing value. For the dependent variable, we may use a model based on limited dependent variable method. Unfortunately, we do not always have the appropriate exogenous variables to construct a probit or tobit model for the missing dependent variable. Another standard method for missing value is to discard all the incomplete cases. Although this approach is quite simple and easy to implement, we should not neglect the possibility of selection biases.

Baltagi and Li [1990] proposed a statistic for testing for individual effects in the case of an incomplete panel. This test statistic is based on the conditional distribution given the observations. On the other hand, test statistic proposed in this paper is an unconditional one.

One object of this paper is to determine the statistical properties of the test statistics that is proposed for the ideal situation and also derive a more robust test when we do not have complete panel data. The outline of this paper is as follows. Section 2 details the model and notation. Section 3 develops the new test statistics and its asymptotic properties. In section 4, some Monte Carlo experiments are conducted to compare several test statistics. Section 5 contains some concluding remarks.

2. MODEL AND NOTATION

We consider an error components model

$$y_{i,t} = \alpha + x'_{i,t}\beta + u_{i,t}, \quad (1)$$

$$u_{i,t} = \mu_i + v_{i,t}, \quad (2)$$

$i = 1, \dots, N$ and $j = 1, \dots, T$, where $x_{i,t}$ is a $k \times 1$ non-stochastic vector of explanatory variables. The error term $u_{i,t}$ consists of two errors, μ_i and $v_{i,t}$, which indicate the individual effect of i-th component and the remainder effect, respectively. For simplicity, we do not treat the time effect in our analysis. We can construct a model with both individual and time effects as in Baltagi and Li [1990].

We assume that μ_i and $v_{i,t}$ are identically, independently normally distributed with zero mean and variances σ_μ^2 and σ_v^2. Obviously $u_{i,t}$ is i.i.d. normal with mean zero and variance $\sigma^2 = \sigma_\mu^2 + \sigma_v^2$. This normality assumption is not necessary for the results in the following sections. We will assume that the usual regularity conditions for the error terms instead of the normality assumption.

The hypothesis being tested is that the *variance component for the individual effect*, σ_μ^2, is zero, that is,

$$\mathcal{H}_0 : \sigma_\mu^2 = 0 \ \ and \ \ \mathcal{H}_a : \sigma_\mu^2 > 0. \tag{3}$$

Instead of discarding the incomplete cases, the standard method for dealing with missing values in empirical research, we treat the incomplete cases as randomly missing cases to avoid incorporating biases cased by discarding the incomplete cases in statistical inference.

To handle incomplete panel data, that is, randomly missing observations, we introduce a Markov chain $\{m_{i,t}\}$ that takes the value zero or unity. This approach is also used in Nishino and Yajima [1999] to consider the unit root process with missing observations. There is an alternative treatment of time series with missing values such as Toda and McKenzie [1999].

If $\{m_{i,t} = 1\}$, then the process is in state 1. In state 1 the observation set for the i-th individual at time t is available. If $\{m_{i,t} = 0\}$, then the process is in state 0 the missing case. That is,

$$m_{i,t} = \begin{cases} 1 & y_{i,t} \ and \ x_{i,t} \ are \ observed, \\ 0 & otherwise. \end{cases} \tag{4}$$

Using the variable $\{m_{i,t}\}$, we rewrite (1) as a model with missing observations

$$y_{i,t}^* = z_{i,t}^* \delta + u_{i,t}^* \tag{5}$$

where $y_{i,t}^* = m_{i,t} y_{i,t}$, $z_{i,t}^* = m_{i,t}(\ 1 \ \ x'_{i,t})$, $\delta = (\ \alpha \ \ \beta')'$ and $u_{i,t}^* = m_{i,t} u_{i,t}$.

We suppose that the transition probabilities are $p_i = Pr(m_{i,t} = 1 \mid m_{i,t-1} = 1)$ and $q_i = Pr(m_{i,t} = 1 \mid m_{i,t-1} = 0)$ for each i and $t = 1, \cdots, T$ and assume that $0 < p_i < 1$ and $0 \le q_i < 1$. The case $q_i = 0$ for all i is excluded.

Let $\mathbf{P}_i$ be the matrix of one-step transition probabilities for the i-th individual given as

$$\mathbf{P}_i = \begin{pmatrix} 1 - q_i & q_i \\ 1 - p_i & p_i \end{pmatrix}. \tag{6}$$

Define the first moment of $m_{i,t}$ as ϕ_i, then we can easily obtain

$$\phi_i = E[\ m_{i,t} \ ] = \frac{q_i}{1 - p_i + q_i}. \tag{7}$$

The second moment of $m_{i,t}$ is

$$E[\ m_{i,t} m_{j,t-s} \ ] = \begin{cases} \phi_i = \sigma_{m,i}^2 + \phi_i^2 & (i = j, \ s = 0), \\ \phi_i \, r_i(s) & (i = j, \ 1 \le s \le T-1), \\ \phi_i \, \phi_j & (i \ne j, all \ s) \end{cases} \tag{8}$$

where $r_i(s)$ is $P[m_{i,t} = 1 \mid m_{i,t-s} = 1]$ which is the (2,2)-element of the matrix of s-step transition probabilities for the i-th individual $\mathbf{P}_i^s$. The matrix $\mathbf{P}_i^s$ can be obtained by multiplying the matrix $\mathbf{P}_i$ by itself s times. See Taylor and Karlin [1993] for details.

The model (5) is rewritten in vector form as

$$y^* = Z^* \delta + u^*, \quad u^* \sim N(\ 0, \ W), \tag{9}$$

where $y^* = (y_1^{*\prime}, y_2^{*\prime}, \cdots, y_N^{*\prime})'$, $Z^* = (z_1^{*\prime}, \ z_2^{*\prime}, \ \cdots, \ z_N^{*\prime} \ )'$, $u^* = (u_1^{*\prime}, \ u_2^{*\prime}, \ \cdots, \ u_N^{*\prime})'$. y_i^* and u_i^* are $T \times 1$ vectors whose t-th elements are $y_{i,t}^*$ and $u_{i,t}^*$, respectively. z_i^* is a $T \times (k+1)$ matrix whose t-th row is $z_{i,t}^*$.

The error term u^* is

$$u^* = \begin{pmatrix} \mu_1^* \\ \mu_2^* \\ \vdots \\ \mu_N^* \end{pmatrix} + \begin{pmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_N \end{pmatrix} = \mu^* + v^*,$$

where μ_i^* and v_i^* are $T \times 1$ vectors whose t-th elements are $m_{i,t} \mu_i$ and $m_{i,t} v_{it}$, respectively. The second moment matrices of μ^* and v^* are as follows.

$$E[\ \mu^* \mu^{*\prime}] = \sigma_\mu^2 \, \phi_i \begin{pmatrix} \mathcal{M}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{M}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{M}_N \end{pmatrix}$$

where the sub-matrix $\mathcal{M}_i$ is

$$\mathcal{M}_i = \begin{pmatrix} r_i(0) & r_i(1) & \cdots & r_i(T-1) \\ r_i(1) & r_i(0) & & \vdots \\ \vdots & & \ddots & r_i(1) \\ r_i(T-1) & \cdots & r_i(1) & r_i(0) \end{pmatrix}$$

and

$$E[\,v^*v^{*\prime}] = \sigma_v^2 \begin{pmatrix} \phi_1 I_T & 0 & \cdots & 0 \\ 0 & \phi_2 I_T & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \phi_N I_T \end{pmatrix}.$$

Then, we have

$$E[\,u^*u^{*\prime}] = \begin{pmatrix} \mathcal{W}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{W}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{W}_N \end{pmatrix} = W.$$

The covariance matrix W is a block-diagonal whose i-th block is $\mathcal{W}_i = \phi_i\,(\sigma_\mu^2\,\mathcal{M}_i + \sigma_v^2\,I_T)$.

3. TEST STATISTIC

In this section, we derive a new test statistic for the individual random effect using the estimators of the variance components. The hypothesis being tested is $\mathcal{H}_0 : \sigma_\mu^2 = 0$.

3.1 Estimators of Variance Components

The structure of the variance-covariance matrix W allows us to estimate the variance components σ_μ^2 and σ^2 as shown in following lemma.

Lemma 1. *If $\hat{\phi}_i$ and $\hat{r}_i(s)$ are consistent estimators of ϕ_i and $r_i(s)$, respectively, then the variance components σ_μ^2 and σ^2 can be estimated consistently as*

$$\hat{\sigma}_\mu^2 = \frac{\hat{u}^{*\prime}(\,I_N \otimes 1_T 1_T^\prime)\hat{u}^* \;-\; \hat{u}^{*\prime}\hat{u}^*}{2\,\sum_{s=1}^{T-1}(T-s)\,\sum_{i=1}^N \hat{\phi}_i\,\hat{r}_i(s)} \tag{10}$$

and

$$\hat{\sigma}^2 = \frac{1}{T\sum_{i=1}^N \hat{\phi}_i}\,\{\,\hat{u}^{*\prime}\hat{u}^*\,\}, \tag{11}$$

where $\hat{u}^*$ is the OLS residuals and $\hat{\phi}_i$ is a consistent estimator of ϕ_i.

Proof. For the OLS residuals $\hat{u}^*$, we have

$$\frac{1}{NT}\hat{u}^{*\prime}\hat{u}^* \;-\; \frac{1}{NT}u^{*\prime}u^* \;=\; o_p(1).$$

Since the expectation of the random variable u_{it}^{*2} is $E[u_{it}^{*2}] = (\sigma_\mu^2 + \sigma_v^2)\,\phi_i \;=\; \sigma^2\,\phi_i$, we have

$$\begin{aligned} \frac{1}{NT}u^{*\prime}u^* &= \frac{1}{N}\sum_{i=1}^N \frac{1}{T}\sum_{t=1}^T u_{i,t}^{*2} \\ &= \frac{1}{N}\sum_{i=1}^N E[u_{i,t}^{*2}] \;+\; o_p(1). \end{aligned}$$

If a consistent estimator of ϕ_i is available, the consistency of the variance component estimator $\hat{\sigma}^2$ is readily shown as

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{(1/N)\sum_{i=1}^N \hat{\phi}_i} \times \frac{1}{NT}\hat{u}^{*\prime}\hat{u}^* \\ &= \frac{1}{(1/N)\sum_{i=1}^N \hat{\phi}_i} \times \frac{1}{NT}u^{*\prime}u^* \;+\; o_p(1) \\ &\xrightarrow{p} \sigma^2 \quad as \;\; T \to \infty, \end{aligned}$$

where $\hat{\phi}_i$ is the consistent estimator of ϕ_i. The consistent estimator of ϕ_i will be given later. It is noted that the consistency of $\hat{\sigma}^2$ holds as NT goes to infinity if $\phi_i = \phi$.

Next, we will show the consistency of $\hat{\sigma}_\mu^2$. Since $u^{*\prime}(\,I_N \otimes 1_T 1_T^\prime)u^*$ can be written as

$$\sum_{i=1}^N \left(\sum_{t=1}^T u_{i,t}^{*2} \;+\; 2\sum_{j=1}^{T-1}\;\sum_{t=j+1}^T u_{i,t}^* u_{i,t-j}^* \right),$$

it is easy to see that $u^{*\prime}\{I_N \otimes 1_T 1_T^\prime \;-\; I_{NT}\}\,u^*/\{NT(T-1)\}$ takes the form

$$\frac{2\,\sigma_\mu^2\,\sum_{s=1}^{T-1}(T-s)\sum_{i=1}^N \phi_i r_i(s)}{NT(T-1)} \;+\; o_p(1). \tag{12}$$

Multiplying both sides of (12) by $NT(T-1)/\{2\sum_{s=1}^{T-1}(T-s)\,\sum_{i=1}^N \phi_i r_i(s)\}$, we obtain

$$\frac{u^{*\prime}(I_N \otimes 1_T 1_T^\prime \;-\; I_{NT})u^*}{2\,\sum_{s=1}^{T-1}(T-s)\sum_{i=1}^N \phi_i r_i(s)} \;-\; \sigma_\mu^2 \;=\; o_p(1).$$

We have to estimate the unknown parameters ϕ_i and $r_i(s)$ consistently in order to obtain the variance component estimators of σ_μ^2 and σ^2. The moments (7) and (8), that is, $\phi_i = E[\,m_{i,t}\,]$ and $\phi_i r_i(s) = E[\,m_{i,t}m_{i,t-s}\,]$, $s = 1, \cdots, T-1$, lead to simple consistent estimators for the unknown parameter ϕ_i and $r_i(s)$ as

$$\hat{\phi}_i \;=\; \frac{1}{T}\sum_{t=1}^T m_{i,t} \;=\; \frac{T_i}{T}, \tag{13}$$

$$\widehat{\phi_i r_i(s)} \;=\; \frac{1}{T}\sum_{t=s+1}^T m_{i,t}m_{i,t-s}, \tag{14}$$

where T_i is the number of observations for each i. It is natural to define T_i as $\sum_{t=1}^T m_{i,t}$.

The consistency of these estimators of ϕ_i can be easily shown using the Strong Law of Large Numbers. Unfortunately, the consistency of (14) does not hold when s is not fixed.

We should provide other estimation method to avoid such difficulties. It is noted that $\phi_i =$

$q_i/(1-p_i+q_i)$ and $r_i(s)$ = (2,2)-th element of $\mathbf{P}_i^s$. p_i and q_i defined in (6) compose the matrix of s-step transition probabilities $\mathbf{P}_i^s$.

Estimation consists of two steps. First we estimate the transition probabilities p_i and q_i. Then, we obtain estimators of ϕ_i and $r_i(s)$ by substituting consistent estimators of p_i and q_i into ϕ_i and $r_i(s)$. We obtain $\hat{p}_i$ and $\hat{q}_i$ to solve following two equation.

$$\begin{aligned}\hat{r}_i(1) &= \hat{p}_i = \frac{\sum_{t=2}^T m_{i,t}\, m_{i,t-1}}{\sum_{t=1}^T m_{i,t}} \\ &= \frac{\sum_{t=2}^T m_{i,t}\, m_{i,t-1}}{T_i}, \qquad (15)\\ \hat{r}_i(2) &= \hat{q}_i(1-\hat{p}_i)+\hat{p}_i^2. \qquad (16)\end{aligned}$$

Explicitly

$$\hat{p}_i = \frac{\sum_{t=2}^T m_{i,t}\, m_{i,t-1}}{T_i} \qquad (17)$$

$$\hat{q}_i = \frac{T_i-\sum_{t=2}^T m_{i,t}\, m_{i,t-1}}{T-T_i}. \qquad (18)$$

It should be noted that $\hat{\phi}_i = 1$ and $\hat{r}_i(s) = 1$ for all s where $T_i = T$, that is, there is no missing observation.

3.2 Construction of Test Statistic

We provide a lemma for the asymptotic normality for $u^{*\prime}(I_N \otimes 1_T 1_T' - I_{NT})u^*$ for the construction of the test statistic.

Lemma 2. *Assume that D_i is a positive constant where*

$$D_i = \lim_{T\to\infty} \frac{1}{T(T-1)} \sum_{s=1}^{T-1}(T-s)\phi_i r_i(s),$$

and define $\bar{D} = \lim_{N\to\infty} N^{-1}\sum_{i=1}^N D_i$. Under the null hypothesis, that is $\sigma_\mu^2 = 0$ and $\sigma^2 = \sigma_v^2$,

$$\frac{u^{*\prime}(\, I_N \otimes 1_T 1_T')u^* - u^{*\prime}u^*}{2\sigma_v^2\sqrt{NT(T-1)}} \xrightarrow{d} N\,(\,0, \bar{D})$$

as N and T go to ∞.

Proof. It is noted that $\phi_i r_i(s)$ is positive and less than unity. Thus, the assumption about D_i is natural. Define a $T\times T$ matrix as

$$\Lambda = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & & \vdots \\ \vdots & & \ddots & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Then, $u^{*\prime}(\, I_N \otimes 1_T 1_T')u^* - u^{*\prime}u^*$ can be written as follows

$$u^{*\prime}\begin{pmatrix} \Lambda & 0 & \cdots & 0 \\ 0 & \Lambda & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Lambda \end{pmatrix} u^* = \sum_{i=1}^N a_i^* \qquad (19)$$

where $a_i^* = u_i^{*\prime}\Lambda u_i^*$. Under the null hypothesis, $\sigma_\mu^2 = 0$, we have

$$E[\, a_i^*\,] = 0, \qquad (20)$$

and

$$Var[\, a_i^*\,] = 4\,\sigma_v^4 \sum_{s=1}^{T-1}(T-s)\phi_i r_i(s). \qquad (21)$$

It is obvious that the expectation of a_i^* is

$$E[\, a_i^*\,] = 2\,\sigma_\mu^2 \left\{\sum_{s=1}^{T-1}(T-s)\ \phi_i r_i(s)\right\} \qquad (22)$$

under the alternative hypothesis. Define

$$b_i^* = \frac{1}{2\,\sigma_v^2\sqrt{T(T-1)}}\, a_i^*. \qquad (23)$$

Since b_i^* is independent random variable whose expectation is $E[\, b_i^*\,] = 0$ and variance is

$$\begin{aligned} Var[\, b_i^*\,] &= \frac{1}{T(T-1)}\sum_{s=1}^{T-1}(T-s)\phi_i r_i(s) \\ &= D_i + o(1), \end{aligned}$$

and

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N D_i = \bar{D} < \infty,$$

we can apply a central limit theorem for

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N b_i^* = \frac{u^{*\prime}(I_N \otimes 1_T 1_T')u^* - u^{*\prime}u^*}{2\sigma_v^2\sqrt{NT(T-1)}}.$$

We are now going to construct a new test statistic for individual effects for incomplete panel data using the ratio of $\hat{\sigma}_\mu^2$ and $\hat{\sigma}^2$.

Theorem 1. *Define $\hat{\rho} = \hat{\sigma}_\mu^2/\hat{\sigma}^2$. Under the null hypothesis $\mathcal{H}_0 : \sigma_\mu = 0$, the new test statistic τ is asymptotically distributed as standard normal.*

$$\begin{aligned}\tau &= \sqrt{\sum_{i=1}^N\sum_{s=1}^{T-1}(T-s)\ \hat{\phi}_i\hat{r}_i(s)} \times \hat{\rho} \\ &\xrightarrow{d} N\,(\,0,\ 1)\end{aligned}$$

where $\hat{\sigma}_\mu^2$ and $\hat{\sigma}^2$ are defined earlier.

Proof. The test statistic τ can be rewritten as

$$\tau = \frac{\sum_{i=1}^{N}\sum_{s=1}^{T-1}(T-s)\,\hat{\phi}_i\hat{r}_i(s)}{\sqrt{\sum_{i=1}^{N}T(T-1)\hat{D}_i}}\,\frac{\hat{\sigma}_\mu^2}{\hat{\sigma}^2}$$

$$= \frac{1}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{D}_i}} \times \frac{\{\hat{u}^{*\prime}(\,I_N \otimes 1_T 1_T'\,)\hat{u}^* \;-\; \hat{u}^{*\prime}\hat{u}^*\,\}}{2\hat{\sigma}^2\sqrt{NT(T-1)}}.$$

It is easy to see that $\hat{\sigma}^2$ is converges to σ^2 under $\mathcal{H}_0$, $\hat{D}_i$ is a consistent estimator of D_i and $\sum_{i=1}^{N}\hat{D}_i/N$ converges to $\bar{D}$. Thus, *Lemma* 2 and its proof complete the proof of theorem.

In particular, the use of $\hat{\phi}_i = T_i/T$ allows us to have a relationship between our τ and the test statistic derived in Baltagi and Li [1990] denoted as τ^{BL}. The test statistic τ^{BL} is for the two-way error component model, however it can be used for the one-way error component model. In such a case, the test statistic is asymptotically distributed as χ_1^2. Since the alternative hypothesis should be one sided, we now compare the square root of the test statistic τ^{BL} for the one-way error component model and our new test statistic τ.

Corollary 1. *In the case of $\hat{\phi}_i = T_i/T$, the ratio of τ^{BL} and τ does not depend on σ_μ^2. Thus, the power of the two test statistics is same. Further if we assume that p_i and q_i are known, then we have an inequality between τ and τ^{BL} such that*

$$\tau \;\leq\; \tau^{BL}. \qquad (24)$$

The equality holds when $T_i = T$ for all i, that is, when there is no missing data for all series i.

Proof. Although different notations are used in Baltagi and Li [1990], it is easy to see that the ratio of the squared statistics can be rewritten as

$$\left(\frac{\tau^{BL}}{\tau}\right)^2 = \frac{2\sum_{i=1}^{N}\sum_{s=1}^{T-1}(1-\frac{s}{T})\frac{1}{T}\,\hat{\phi}_i\hat{r}_i(s)}{\sum_{i=1}^{N}\hat{\phi}_i(\hat{\phi}_i - \frac{1}{T})}.$$

It is clear that this ratio depends only $m_{i,t}$ which is introduced to handle missing observations. Next we show that this ratio is greater than 1 when p_i and q_i are known which leads to $\hat{r}_i(s) = r_i(s)$. The variance of the random variable T_i/T is

$$Var[\,\frac{T_i}{T}\,] = E[\,(\frac{T_i}{T})^2\,] - \{E[\,\frac{T_i}{T}\,]\}^2$$
$$= \left\{2\sum_{s=1}^{T-1}(1-\frac{s}{T})\frac{1}{T}\,\phi_i r_i(s) \;-\; \phi_i(\phi_i - \frac{1}{T})\right\}.$$

Using the form of $Var[\,T_i/T\,]$,

$$\sum_{i=1}^{N}Var[\,\frac{T_i}{T}\,] = \sum_{i=1}^{N}\phi_i(\phi_i - \frac{1}{T}) \times\left\{\frac{2\sum_{i=1}^{N}\sum_{s=1}^{T-1}(1-\frac{s}{T})\frac{1}{T}\,\phi_i r_i(s)}{\sum_{i=1}^{N}\phi_i(\phi_i - \frac{1}{T})} - 1\right\} \geq 0. \qquad (25)$$

Since (25) holds for any value of ϕ_i between zero and unity, the ratio $(\tau^{BL}/\tau)^2$ is greater than unity.

Generally, the ratio τ^{BL}/τ would be greater than unity even when we use consistent estimators instead of the true value of $r_i(s)$, though the difference between the two test statistics is asymptotically negligible.

Corollary 1 makes the point that there is a possibility that the statistic τ^{BL} rejects and τ does not reject the null hypothesis when the null hypothesis is true in the case of the analysis of panel data with missing values.

4. NUMERICAL EXAMPLE

We conduct some Monte Carlo experiments to compare the tests τ^{BL} and τ for individual effects in the error component model. As shown in corollary 1, the powers of these test statistics are the same. Thus, we will focus on the properties of the statistics under the null hypothesis. Though these test statistics are asymptotically the same, their finite sample behaviors are unknown. To determine their properties, we conduct some Monte Carlo experiments.

The model is set up as in Baltagi and Chang [1994].

$$y_{it} = 5.0 + 0.5x_{it} + u_{it},\quad i = 1,\dots,N,\ t = 1,\dots,T.$$

x_{it} was generated as

$$x_{it} = 0.1t + 0.5x_{it-1} + \omega_{it}$$

and ω is distributed uniformly over the interval $[-0.5,\ 0.5]$. The initial values of x_{i0} were chosen as $(5 + 10\omega_{i0})$. The error term was specified as $u_{it} = \mu_i + v_{it}$ with $\mu_i \sim iidn(0, \sigma_\mu^2)$, $v_{it} \sim iidn(0, \sigma_v^2)$ and we set $\sigma^2 = \sigma_\mu^2 + \sigma_v^2 = 20$. The null hypothesis is H_0: $\sigma_\mu^2 = 0$ and the alternative is H_a: $\sigma_\mu^2 > 0$. To determine the finite sample behavior under the null hypothesis, we set $\sigma_\mu^2 = 0$.

The test statistics being examined are τ^{BL} which is the square-root of the LM test for the one-way error component model proposed

Table 1. Rejection probabilities: Bernoulli.

N	Probability of Missing	Ratio of Complete Series	τ^{BL}	τ
		3/4	0.058	0.058
50	0.8	2/4	0.065	0.064
		1/4	0.064	0.058
		3/4	0.056	0.055
50	0.4	2/4	0.053	0.052
		1/4	0.053	0.051
		3/4	0.049	0.049
100	0.8	2/4	0.043	0.042
		1/4	0.056	0.055
		3/4	0.051	0.050
100	0.4	2/4	0.051	0.050
		1/4	0.043	0.042

Table 2. Rejection probabilities: AB sampling.

N	A,B	Ratio of Complete Series	τ^{BL}	τ
		3/4	0.050	0.050
50	2,1	2/4	0.065	0.064
		1/4	0.049	0.049
		3/4	0.053	0.053
50	3,1	2/4	0.053	0.053
		1/4	0.046	0.046
		3/4	0.047	0.046
100	2,1	2/4	0.045	0.045
		1/4	0.048	0.047
		3/4	0.050	0.050
100	3,1	2/4	0.057	0.056
		1/4	0.056	0.056

in Baltagi and Li [1990] and τ which is the test statistic proposed in the previous section. To estimate the rejection probabilities, we conduct the experiments 1000 times. The nominal size is set 5%, and T is set to 20. We consider two kinds of missing structures, one is a Bernoulli trial, and the other is A-B sampling. A-B sampling is considered in Nishino and Yajima [1999] and the references therein. The sampling scheme is that there are A observations followed by B missing values, repeated m times. The total number of observations T is $m \times (A+B)$. The number of available observations is $m \times A$.

The results of the simulations summarized in Table 1 and Table 2. In Table 1, we see that the size distortion becomes large when the probability of observation being missing increases. As the sample size increases, the properties of two test statistics become similar. In the case of A-B sampling, the difference between τ^{BL} and τ is small. This is because the sequence of A-B sampling is deterministic and periodic, and is not random. In such cases, both tests have reasonably good sizes. All rejection probabilities using τ are smaller than those of τ^{BL}.

5. CONCLUDING REMARK

We examine the statistical properties of a test statistic for random effects with incomplete panel data and compare it with the test statistics derived without randomness such as Baltagi and Li [1990]. Though the two test statistics are equivalent asymptotically, there is some difference in their behavior in finite samples. We find that the test statistic derived in this paper is conservative when compared with the test statistic derived in Baltagi and Li [1990]. It is natural to consider the non-response for query in the questionnaire to be the outcome of a random variable when we do not have any information about non-responses since many of household surveys are based on the random sampling. It is possible that the test statistic derived without randomness leads upward bias in finite sample.

6. ACKNOWLEDGEMENT

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology: Grant-in-Aid for Scientific Research (C)(2), 13630031, 2001. The author would like to thank Colin Ross McKenzie and Michael McAleer for helpful their comments.

7. REFERENCES

Baltagi, B.H. and Y. Chang, Incomplete Panels A Comparative Study of Alternative Estimators for the Unbalanced One-way Error Component Regression Model, *Journal of Econometrics*, 62, 67-89, 1994.

Baltagi, B.H. and Q. Li, A Lagrange Multiplier Test for the Error Component Model with Incomplete Panels, *Econometric Reviews*, 9, 103-107, 1990.

Nishino, H and Y. Yajima, Parameter Estimation of Unit Root Processes with Missing Observations, *Journal of the Japan Statistical Society*, 29, 181-200, 1999.

Taylor, H.M. and S. Karlin, *An Introduction to Stochastic Modeling*, Revised Edition, Academic Press,1993.

Toda, Y.H. and C.R. McKenzie, LM Tests for Unit Roots in the Presence of Missing Observations: Small Sample Evidence, *Mathematics and Computers in Simulation*, 48, 457-468, 1999.