

# Alternative Tests for the Nested Error Component Regression Model <sup>†</sup>

Seuck Heun Song<sup>1</sup> and Byoung Cheol Jung<sup>2</sup>

## ABSTRACT

We consider the panel data regression model with nested error components. In this paper, the several Lagrange Multiplier tests for the nested error component model are derived. These tests extend the earlier work of Honda(1985), Moulton and Randolph(1989), Baltagi, et al.(1992) and King and Wu(1997) to the nested error component case. Monte Carlo experiments are conducted to study the performance of these LM tests.

*Keywords:* Panel data; Nested error component; LM Tests.

## 1. INTRODUCTION

In the error components regression model, most researcher have been provided the Lagrange Multiplier(LM) tests for tesing the existence of the various error components(Baltagi et al.(1992), Baltagi and Li(1991, 1995) and Jung, et al.(1999)). The LM (or Rao-Score) test is based on the estimation of the model under the null hypothesis and in most cases its computation requires only ordinary least squares residuals.

This paper considers the panel data regression model in which the economic data has a natural nested grouping. For example, data on firms may be grouped by industry, data on states by region and data on individuals by profession. In this case, one can control for unobserved group and nested subgroup effects using a nested error component model(see Baltagi(1993)).

Recently, Jung, et al.(1999) derived a LM test which jointly tests for the presence of random group and nested subgroup effects. In this paper, we derive a new one-sided version of these tests given that these variance components are non-negative. This extends the Honda(1985) test to the nested case. We also consider

---

<sup>†</sup>The present study as supported by KISTEP Fund of 1999

<sup>1</sup>Department of Statistics, Korea University, Seoul 136-701, Korea

<sup>2</sup>Institute of Statistics, Korea University, Seoul 136-701, Korea

the standardized version of these LM tests, denoted by SLM. This extends the results of Moulton and Randolph(1989) to the nested case. When one uses the one-directional LM test of the existence of nested subgroup effects, one implicitly assumes that the group effects do not exist. In this case, the resulting test may lead to incorrect decisions(see Baltagi et al.(1992)). To overcome this problem, we propose a conditional LM test for the existence of nested subgroup effects assuming the presence of random group effects. This guards against misleading inference caused by a LM test that ignores the presence of random industry effects, when in fact they are present.

Further, We investigate the performance of these proposed LM tests using Monte Carlo experiments and exact power comparisons. Other tests included in this comparison are the likelihood ratio(LR) tests, and the F tests used in an ANOVA framework.

Section 2 gives the nested error component model and the hypotheses to be tested. Section 3 derives the various LM tests. The proofs are relegated to the Appendices. Section 4 compares the performance of these LM tests using Monte Carlo experiments. Section 5 gives a conclusion.

## 2. THE MODEL

We consider the following panel data regression model

$$y_{ijt} = x'_{ijt}\beta + u_{ijt}, \quad i = 1, \dots, M, \quad j = 1, \dots, N \text{ and } t = 1, \dots, T, \quad (2.1)$$

where  $y_{ijt}$  be an observation on a dependent variable for the  $j$ th nested subgroup within the  $i$ th group(for example,  $j$ th firm in the  $i$ th industry) for the  $t$ th time period.  $x_{ijt}$  denotes of  $k$  nonstochastic regressor vector. The disturbances term  $u_{ijt}$  in (2.1) are assumed that

$$u_{ijt} = \mu_i + \nu_{ij} + \varepsilon_{ijt}, \quad i = 1, \dots, M, \quad j = 1, \dots, N \text{ and } t = 1, \dots, T, \quad (2.2)$$

where  $\mu_i$  denote  $i$ th group specific effects which are assumed to be *i.i.d.*  $(0, \sigma_\mu^2)$ ,  $\nu_{ij}$  denote the nested subgroup effects within the  $i$ th group which are *i.i.d.*  $(0, \sigma_\nu^2)$  and  $\varepsilon_{ijt}$  are the remainder disturbances which are also assumed to be *i.i.d.*  $(0, \sigma_\varepsilon^2)$ . The  $\mu_i$ 's,  $\nu_{ij}$ 's and  $\varepsilon_{ijt}$ 's are independent of each other and among themselves.

This is a panel data regression model with nested error components. The model (2.1) can be rewritten in a matrix notation as

$$y = X\beta + u, \quad (2.3)$$

where  $y$  is a  $MNT \times 1$  observation vector,  $X$  is a  $MNT \times k$  design matrix.  $\beta$  is a  $k \times 1$  parameter vector, and  $u$  is a  $MNT \times 1$  disturbance vector. Both  $N$  and  $T$  are assumed to be larger than  $k$ . The equation (2.2) is in vector form:

$$u = (I_M \otimes i_N \otimes i_T)\mu + (I_M \otimes I_N \otimes i_T)\nu + \varepsilon, \quad (2.4)$$

where  $\mu' = (\mu_1, \dots, \mu_M)$ ,  $\nu' = (\nu_{11}, \dots, \nu_{MN})$ ,  $\varepsilon' = (\varepsilon_{111}, \dots, \varepsilon_{11T}, \dots, \varepsilon_{MNT})$ ,  $i_N$  and  $i_T$  are vectors of ones of dimension  $N$  and  $T$ , respectively.  $I_M$  and  $I_N$  are identity matrices of dimension  $M$  and  $N$ , and  $\otimes$  denotes the Kronecker product.

The hypotheses under considerations are the following :

- a)  $H_0^a : \sigma_\mu^2 = \sigma_\nu^2 = 0$ , and the alternative  $H_1^a$  is that at least one component is greater than zero.
- b)  $H_0^b : \sigma_\mu^2 = 0$ , and the one-sided alternative is  $H_1^b : \sigma_\mu^2 > 0$  (assuming  $\sigma_\nu^2 = 0$ ).
- c)  $H_0^c : \sigma_\nu^2 = 0$ , and the one-sided alternative is  $H_1^c : \sigma_\nu^2 > 0$  (assuming  $\sigma_\mu^2 = 0$ ).
- d)  $H_0^d : \sigma_\nu^2 = 0$  (assuming  $\sigma_\mu^2 > 0$ ), and the one-sided alternative is  $H_1^d : \sigma_\nu^2 > 0$  (assuming  $\sigma_\mu^2 > 0$ ).

Three classes of tests are considered :

- (1) LM tests : the LM tests are simple to apply and which were popularized in econometrics and statistics by Godfrey(1989).
- (2) ANOVA F tests : the ANOVA F tests used to test the significant of the fixed effects and which were recently considered by Moulton and Randolph(1989) and Baltagi et al.(1992).
- (3) LR tests : the likelihood ratio tests are computationally more expensive than the LM tests, but under the proper specification of the likelihood which are known to have desirable properties.

In the following section, we discuss each of these test statistics for the null hypothesis considered.

### 3. TEST STATISTICS

#### 3.1. LM Tests

(i) Testing  $H_0^a$

The joint LM test statistic for testing  $H_0^a$  vs  $H_1^a$  is given by

$$LM_1 = \frac{MN}{2(N-1)} \left( A^2 - 2AB + \frac{NT-1}{T-1} B^2 \right), \quad (3.1)$$

where  $A = \frac{\tilde{u}'(I_M \otimes J_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1$ ,  $B = \frac{\tilde{u}'(I_M \otimes I_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1$ ,  $\tilde{u}$  is OLS residuals, and  $J_N$  and  $J_T$  are matrices of ones of dimension  $N$  and  $T$ . See the derivation of (3.1) in Jung, et al.(1999). Under  $H_0^a$ ,  $LM_1$  is asymptotically distributed as  $\chi^2(2)$ . One weakness of the joint test is that, if  $H_0^a$  is rejected, one cannot infer without further testing whether  $\sigma_\mu^2$  or  $\sigma_\nu^2$  or both are different from zero. Also, this joint test will not be optimal if only one of the variance

components is actually zero. This is the problem of overtesting discussed in Bera and Jarque (1982).

The presence of the interaction term in joint LM test statistic (3.1) emphasizes the importance of joint test, but when  $N$  is large the interaction term becomes negligible (Jung, et al.(1999)). Therefore, for the joint test  $H_0^a$ , the 'handy' one-sided test suggested by Honda(1985) is extended as

$$HO = \left( \sqrt{\frac{MNT}{2(NT-1)}} A + \sqrt{\frac{MNT}{2(T-1)}} B \right) / \sqrt{2} \quad (3.2)$$

which is asymptotically distributed as  $N(0, 1)$ .

Furthermore, the locally mean most powerful(LMMP) one-sided test suggested by King and Wu(1997) and derived for the model (2.3) is given by

$$KW = \sqrt{\frac{MNT}{2(NT+3T-4)}} (A+B), \quad (3.3)$$

see the derivation in appendix 2.

Note that  $A$  or  $B$  can be negative when one or both variance components are small and close to zero. For the purpose of immune to the negative values of  $A$  and  $B$ , Gourieroux, Holly, and Monfort(1982) proposed the mixed  $\chi^2$  test, hereafter GHM, and Baltagi, et al.(1992) extended the GHM test to the one-sided joint test for the two-way error component. Applying to the GHM test in our model, we get the following test :

$$\chi_m^2 = \begin{cases} A^2 + B^2 & \text{if } A > 0, B > 0, \\ A^2 & \text{if } A > 0, B \leq 0, \\ B^2 & \text{if } A \leq 0, B > 0, \\ 0 & \text{if } A \leq 0, B \leq 0, \end{cases} \quad (3.4)$$

where  $\chi_m^2$  denote the mixed  $\chi^2$  distribution. Under the null hypothesis of  $H_0^a$ ,

$$\chi_m^2 \sim \left(\frac{1}{4}\right)\chi^2(0) + \left(\frac{1}{2}\right)\chi^2(1) + \left(\frac{1}{4}\right)\chi^2(2), \quad (3.5)$$

where  $\chi^2(0)$  equals zero with probability one, see Gourieroux, et al.(1982).

(ii) Testing  $H_0^b$  and Testing  $H_0^c$

From the appendix 1, it can be shown that the  $A^2$  term is the basis for the LM test statistic for testing  $H_0^b$  vs  $H_1^b$ . In fact,

$$LM_2 = \sqrt{\frac{MNT}{2(NT-1)}} A \quad (3.6)$$

is asymptotically distributed (for large  $M$ ) as  $N(0, 1)$  under  $H_0^b$ , see Jung, et al.(1999). Also the  $B^2$  is the basis for the LM test statistic for  $H_0^c$  vs  $H_1^c$  assuming there are no group effect. Similar derivation to equation (3.6), the one-sided LM test statistic for  $H_0^c$  vs  $H_1^c$  is

$$LM_3 = \sqrt{\frac{MNT}{2(T-1)}} B \quad (3.7)$$

which is asymptotically distributed (for large  $MN$ ) as  $N(0, 1)$ .

Moulton and Randolph(1989) showed that the asymptotic  $N(0, 1)$  approximation for testing the individual effect in one-way error component model can be poor even in large samples. This occurs when the number of regressors is large or intra-class correlation is high, see Moulton and Randolph(1989). They suggest an alternative Standardized LM (SLM) test which centers and scales the one-sided LM test so that its mean is zero and its variance is one. The SLM procedure extends to our model for testing  $H_0^b$  and  $H_0^c$ . For the  $H_0^b : \sigma_\mu^2 = 0$ , it can be shown that the SLM statistic

$$SLM_1 = \frac{LM_2 - E(LM_2)}{\sqrt{var(LM_2)}} = \frac{d_2 - E(d_2)}{\sqrt{var(d_2)}}, \quad (3.8)$$

where  $d_2 = \tilde{u}'D_2\tilde{u}/\tilde{u}'\tilde{u}$  and  $D_2 = (I_M \otimes J_N \otimes J_T)$ . Using the results on moments of quadratic forms in regression residuals, we get

$$E(d_2) = tr(D_2(I_{MNT} - P_X))/s, \quad (3.9)$$

where  $s = MNT - k$ ,  $P_X = X(X'X)^{-1}X'$  and

$$Var(d_2) = 2 \frac{\left\{ tr[(D_2(I_{MNT} - P_X))^2]s - [tr(D_2(I_{MNT} - P_X))]^2 \right\}}{s^2(s+2)}, \quad (3.10)$$

see Evans and King(1985). Also, for the  $H_0^c : \sigma_\nu^2 = 0$ , the SLM statistic can be

$$SLM_2 = \frac{LM_3 - E(LM_3)}{\sqrt{var(LM_3)}} = \frac{d_3 - E(d_3)}{\sqrt{var(d_3)}}, \quad (3.11)$$

where  $d_3 = \tilde{u}'D_3\tilde{u}/\tilde{u}'\tilde{u}$  and  $D_3 = (I_M \otimes I_N \otimes J_T)$ . Under the null hypothesis,  $SLM_1$  and  $SLM_2$  are asymptotically distributed as  $N(0, 1)$ , see Moulton and Randolph(1989).

(iii) Testing  $H_0^d$

When one use  $B$  to test the zero nested subgroup effects, one assumes that the group effects do not exist. When the group effects exist, this may lead to the incorrect decisions. To overcome this problem, we propose the following test which tests the nested subgroup effects assuming that the group effects are present. The corresponding test (for testing  $H_0^d$  vs  $H_1^d$ ) is derived as follow, see appendix 3,

$$LM_\nu = \sqrt{\frac{M(N-1)(NT-1)}{2N(T-1)}} \left( \frac{NT-1}{N-1} \frac{\hat{u}'(I_M \otimes E_N \otimes \bar{J}_T)\hat{u}}{\hat{u}'(I_M \otimes E_{NT})\hat{u}} - 1 \right), \quad (3.12)$$

where  $E_{NT} = I_{NT} - \bar{J}_{NT}$ ,  $\bar{J}_{NT} = J_{NT}/NT$ , and  $J_{NT}$  is the matrix of ones of dimension  $NT$ , and  $\hat{u}$  denote the restricted GLS residuals using the maximum

likelihood estimate. Under the null hypothesis, the  $LM_\nu$  is also asymptotically distributed as  $N(0, 1)$ .

### 3.2. F tests

Moulton and Randolph(1989) found that the ANOVA F test which tests the significance of the fixed effects for the one-way error component model. Later, Baltagi, et al.(1992) extends their work to the two-way error component model. In these paper, the ANOVA F tests performs well comparing to the LR and LM tests. In our model, the idea of ANOVA F tests extendet to the nested error components model, and we will compare the performance of ANOVA F test with that of the LM and LR tests.

The ANOVA F test statistics have the following general form :

$$F = \frac{(RRSS - URSS)/(df2 - df1)}{URSS/df1}, \quad (3.13)$$

where  $RRSS$  and  $URSS$  are the restricted and unrestricted residual sums of squares, respectively, and  $df1$  and  $df2$  are the degree of freedoms of  $RRSS$  and  $URSS$ , respectively. Under the null hypothesis, this statistic has a central  $F$  distribution with  $(df2 - df1)$  and  $df1$  degrees of freedom. Table 1 defines the respective components of  $F$  for each hypothesis considered.

Table1. ANOVA F test statistics

$H_0$	URSS	RRSS	df1	df2
$H_0^a$	$\tilde{u}'_1 \tilde{u}_1$	$\tilde{u}'_{OLS} \tilde{u}_{OLS}$	MN(T-1)-1	MNT-2
$H_0^b$	$\tilde{u}'_2 \tilde{u}_2$	$\tilde{u}'_{OLS} \tilde{u}_{OLS}$	M(NT-1)-1	MNT-2
$H_0^c$	$\tilde{u}'_3 \tilde{u}_3$	$\tilde{u}'_{OLS} \tilde{u}_{OLS}$	MNT-M(N-1)-2	MNT-2
$H_0^d$	$\tilde{u}'_1 \tilde{u}_1$	$\tilde{u}'_2 \tilde{u}_2$	MN(T-1)-1	M(NT-1)-1

where  $\tilde{u}_{OLS}$  is the familiar OLS residual vector and  $\tilde{u}_1$  is the residual vector from the regression of  $y_{ijt} - \bar{y}_{ij}$ . on  $X_{ijt} - \bar{X}_{ij}$ , where  $\bar{y}_{ij} = \sum_t y_{ijt}/T$  and  $\bar{X}_{ij} = \sum_t X_{ijt}/T$ . Also,  $\tilde{u}_2$  is the residual vector from the regression of  $y_{ijt} - \bar{y}_{i..}$  on  $X_{ijt} - \bar{X}_{i..}$ , where  $\bar{y}_{i..} = \sum_j \sum_t y_{ijt}/NT$  and  $\bar{X}_{i..} = \sum_j \sum_t X_{ijt}/NT$ , and  $\tilde{u}_3$  is the residual vector from the regression of  $(y_{ijt} - \bar{y}_{ij} + \bar{y}_{i..})$  on  $(X_{ijt} - \bar{X}_{ij} + \bar{X}_{i..})$ .

### 3.3. Likelihood Ratio Tests

The one-sided LR tests have the following form :

$$LR = -2\log \frac{l(res)}{l(unres)}, \quad (3.14)$$

where  $l(res)$  and  $l(unres)$  denote the restricted and unrestricted maximum likelihood value, respectively. The LR tests require ML estimators of the one-way and nested effect models, therefore, they are comparatively more expensive than their LM counterparts. Under the null hypothesis, the LR test statistics have the same asymptotic properties as their LM counterparts, see Gourieroux, et al.(1982). More specifically, for  $H_0^a$ ,  $LR \sim \frac{1}{4}\chi^2(0) + \frac{1}{2}\chi^2(1) + \frac{1}{4}\chi^2(2)$ , and for  $H_0^b$ ,  $H_0^c$ , and  $H_0^d$ ,  $LR \sim \frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1)$ .

## 4. MONTE CARLO RESULTS

The model is set as follows :

$$y_{ijt} = \alpha + x_{ijt}\beta + u_{ijt}, \quad i = 1, \dots, M, \quad j = 1, \dots, N, \quad t = 1, \dots, T, \quad (4.1)$$

where  $\alpha = 5$  and  $\beta = 0.5$ .  $x_{ijt}$  was generated by a similar method to that of Nerlove(1971). In fact,  $x_{ijt} = 0.1t + 0.5x_{ij,t-1} + w_{ijt}$ , where  $w_{ijt}$  is uniformly distributed on the interval  $[-0.5, 0.5]$ . The initial values  $x_{ij0}$  was chosen as  $(100 + 250w_{ij0})$ . For the disturbances,  $u_{ijt} = \mu_i + \nu_{ij} + \varepsilon_{ijt}$  with  $\mu_i \sim i.i.d.N(0, \sigma_\mu^2)$ ,  $\nu_{ij} \sim i.i.d.N(0, \sigma_\nu^2)$  and  $\varepsilon_{ijt} \sim i.i.d.N(0, \sigma_\varepsilon^2)$ . We fix  $\sigma^2 = \sigma_\mu^2 + \sigma_\nu^2 + \sigma_\varepsilon^2 = 20$  and let  $\gamma_1 = \sigma_\mu^2/\sigma^2$  and  $\gamma_2 = \sigma_\nu^2/\sigma^2$  were varied over the set  $(0, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8)$  such that  $(1 - \gamma_1 - \gamma_2)$  is always positive. For the sample size combinations,  $M = 5$  and  $10$  were used, and for each  $M$ , the following  $(N, T)$  are used :  $(N, T) = (5, 5)(10, 5)(20, 5)(5, 10)(10, 10)(20, 10)(5, 20)(10, 20)$ . For all combination of  $(M, N, T)$ , and  $\gamma_1$  and  $\gamma_2$ , 1000 replications are performs. For each replications, we calculate the following test statistics for  $H_0^a$  : LM<sub>1</sub> test, HO test, KW test, GHM test, LR test, and F test, and for  $H_0^b$  and  $H_0^c$  : LM<sub>2</sub> test, LM<sub>3</sub> test, SLM<sub>1</sub> test, SLM<sub>2</sub> test, LR test, and F test, and for  $H_0^d$  : LM <sub>$\nu$</sub>  test, LR test, and F test.

#### 4.1. Testing $H_0^a : \sigma_\mu^2 = 0$ and $\sigma_\nu^2 = 0$

Table 2 gives the number of rejections in 1000 replications for the various tests employed for  $H_0^a : \sigma_\mu^2 = 0$  and  $\sigma_\nu^2 = 0$  when  $(M, N, T) = (5, 5, 5)$ ,  $(10, 5, 5)$ ,  $(5, 10, 10)$ . (Similar tables for other combinations of  $(M, N, T)$  are not produced here to save space. These results are available upon request from the

authors.) All the tests considered except for the LR test have estimated size that are not significantly different from the nominal size. The LR test significantly underestimate the nominal size for all considered  $(M, N, T)$  combinations, and it is not recommended.

The power of all the tests increase as  $M$ ,  $N$  and  $T$  increase, or  $\gamma_1$  and  $\gamma_2$  increase. Compared to the HO test and the GHM test, the two-sided  $LM_1$  test has lower rejection number. The power of the ANOVA F test is higher than the other tests considered when  $\gamma_1 = 0$ , also the KW test has high power rejecting the null hypothesis when  $\gamma_2 = 0$ . But when  $\gamma_1 \geq 0.05$  or  $\gamma_2 \geq 0.05$ , the HO and the GHM test has the higher power.

Therefore, the HO and GHM test are recommended in joint test for testing  $H_0^a : \sigma_\mu^2 = \sigma_\nu^2 = 0$  for the nested error components model.

#### 4.2. Testing $H_0^b : \sigma_\mu^2 = 0$ and Testing $H_0^c : \sigma_\nu^2 = 0$

Table 3 gives the number of rejections for the various tests employed for  $H_0^b : \sigma_\mu^2 = 0$  when  $(M, N, T) = (5, 5, 5), (10, 5, 5), (5, 10, 10)$ . When  $H_0^b$  is true ( $\gamma_1 = 0$ ) and  $\gamma_2$  is large, all the tests badly overrejects the null hypothesis since they ignore the fact that  $\sigma_\nu^2 > 0$ . This is because all the tests considered do not take into account the fact  $\sigma_\nu^2 > 0$  and implicitly assume that  $\sigma_\nu^2 = 0$ .

When  $\sigma_\mu^2$  is existent ( $\gamma_1 \geq 0.05$ ), all the tests perform well in rejecting the null hypothesis, and the power of all the tests increase as  $\gamma_2$  increases.

Similarly, Table 4 gives the results of testing  $H_0^c : \sigma_\nu^2 = 0$  when  $(M, N, T) = (5, 5, 5), (10, 5, 5), (5, 10, 10)$ . When  $H_0^c$  is true ( $\gamma_2 = 0$ ) and  $\gamma_1$  is large, all the tests perform badly since they ignore the fact that  $\sigma_\mu^2 > 0$ . In fact, *HO*, *SLM* and *LR* test badly overrejects the null while the ANOVA F underestimate the nominal size. The overrejection of the *HO*, *SLM* and *LR* test may caused by similar reason for testing  $H_0^b : \sigma_\mu^2 = 0$ . However, the underestimation of the ANOVA F test is caused by primary group effect is transferred to the nested subgroup.

#### 4.3. Testing $H_0^d : \sigma_\nu^2 = 0$ (given $\sigma_\mu^2 > 0$ )

Table 5 gives the results of testing  $H_0^d : \sigma_\nu^2 = 0$  (assuming  $\sigma_\mu^2 > 0$ ). The estimated size is not significantly different from the nominal size for all tests considered. For  $\gamma_2 \geq 0.2$ , all the tests have high power rejecting the null hypothesis in 70.8 % to 100 % of the cases when  $(M, N, T) = (5, 5, 5)$  and the power of all the test increase as  $M, N$  and  $T$  increase. For  $0 < \gamma_2 < 0.2$ , we first con-



sider the case that  $\gamma_1 > 0$ . In table 5, as moved from the top block( $\gamma_1 = 0$ ) to the bottom block( $\gamma_1 = 0.8$ ), the power of all the tests improves as  $\gamma_1$  increases. Next, we consider the case that  $\gamma_1 = 0$ . The first block corresponds to the case  $\sigma_\mu^2 = 0$  ( $\gamma_1 = 0$ ). Comparing this block with the first block of Table 4, we see that the power of tests in  $H_0^d$  does not largely deteriorates to compare with the power of tests for testing  $H_0^c$  even though  $\sigma_\mu^2 = 0$ . Hence overspecifying the model, i.e., assuming the model is nested( $\sigma_\mu^2 > 0$ ) when in fact it is one-way( $\sigma_\mu^2 = 0$ ) does not seem to hurt the power of these tests. Also, the conditional LM tests shows the higher power rejection number compared to LR and ANOVA F tests.

The results in this subsection 4.2 and 4.3 strongly support the fact that one should not ignore the possibility of  $\sigma_\mu^2 > 0$  when testing  $\sigma_\nu^2 = 0$ . In fact, our results suggest that it may be better to overspecify the model rather than underspecify it in testing the variance components. Also, the conditional LM test is recommended for testing  $H_0^d : \sigma_\nu^2 = 0$  (assuming  $\sigma_\mu^2 > 0$ ).

## 5. CONCLUSION

This paper derives several LM tests for the nested error component model. The Monte Carlo experiments show that: (i) the HO and the GHM test are recommended for  $H_0^a$ . (ii) The one directional LM tests, LR test and ANOVA F test that assume the other variance component is zero overreject the null hypothesis( $H_0^b$  and  $H_0^c$ ). (iii) The conditional LM test that explicitly assume the other variance component is positive perform well in the Monte Carlo experiments and is recommended.

## APPENDICES

### Appendix 1

This appendix derives the LM test statistic for testing  $H_0^b : \sigma_\mu^2 = 0$  and  $H_0^c : \sigma_\nu^2 = 0$ . From the Jung, et al.(1999), the partial derivatives and the information matrix in order to test  $H_0^a : \sigma_\mu^2 = \sigma_\nu^2 = 0$  is given by

$$\tilde{D} = \begin{bmatrix} \frac{MNT}{2\tilde{\sigma}_\epsilon^2} \left( \frac{\tilde{u}'(I_M \otimes J_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1 \right) \\ \frac{MNT}{2\tilde{\sigma}_\epsilon^2} \left( \frac{\tilde{u}'(I_M \otimes I_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1 \right) \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{MNT}{2\tilde{\sigma}_\epsilon^2} A \\ \frac{MNT}{2\tilde{\sigma}_\epsilon^2} B \\ 0 \end{bmatrix}, \quad (A.1)$$

where  $\tilde{u}$  is the OLS residuals, and

$$\tilde{J} = \frac{MNT}{2\tilde{\sigma}_\varepsilon^4} \begin{bmatrix} NT & T & 1 \\ T & T & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad (\text{A.2})$$

where  $A = \frac{\tilde{u}'(I_M \otimes J_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1$  and  $B = \frac{\tilde{u}'(I_M \otimes I_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}} - 1$ .

It is note that, if  $\sigma_\nu^2 = 0$  and one is testing  $H_0 : \sigma_\mu^2 = 0$ , then we ignore the second element of  $\tilde{D}$  in (A.1) and the second row and column of  $\tilde{J}$  in (A.2). In this case, the LM statistic becomes

$$LM^b = \frac{MNT}{2(NT-1)} A^2. \quad (\text{A.3})$$

Under the null hypothesis, the LM statistic is asymptotically distributed as  $\chi^2(1)$ . Similarly, if  $\sigma_\mu^2 = 0$  and one is testing  $H_0 : \sigma_\nu^2 = 0$ , then we ignore the first element of  $\tilde{D}$  in (A.1) and the first row and column of  $\tilde{J}$  in (A.2). In this case, the LM statistic becomes

$$LM^c = \frac{MNT}{2(T-1)} B^2. \quad (\text{A.4})$$

Under the null hypothesis, the LM statistic is asymptotically distributed as  $\chi^2(1)$ .

## Appendix 2

This appendix derives the KW LM test for testing  $H_0^a : \sigma_\mu^2 = \sigma_\nu^2 = 0$ . The parameter  $\theta$  can be divided as  $\theta' = (\theta'_1, \theta'_2)$ , where  $\theta'_1 = (\sigma_\mu^2, \sigma_\nu^2)$ , and  $\theta_2 = (\sigma_\varepsilon^2)$ . The null hypothesis is  $H_0^a : \theta_1 = 0$  and the one-sided alternative is  $H_1^a : \theta_1 > 0$ . Then, using the King and Wu(1997), the KW test statistic is given by

$$KW = \tilde{S}^+ / (i_2'(\tilde{J}^{11})^{-1}i_2)^{1/2}, \quad (\text{A.5})$$

where  $\tilde{S}^+ = \sum_{i=1}^2 \frac{\partial \ln L}{\partial \theta_i} |_{\tilde{\theta}=(\theta', \tilde{\theta}_2)'}$ ,  $i_2$  is a vector of ones of dimension 2 and  $\tilde{J}^{11}$  denote the upper  $2 \times 2$  block of the inverse of the information matrix evaluated at  $\tilde{\theta} = (\theta', \tilde{\theta}_2)'$ . For the nested effect model, the score vector is given by

For the nested effect model,  $\theta'_1 = (\sigma_\mu^2, \sigma_\nu^2)$  and  $\theta_2 = \sigma_\varepsilon^2$ , and the score vector is given by (A.1) and the information matrix is given by (A.2). Therefore,  $\tilde{J}^{11}$  is given by

$$\tilde{J}^{11} = \frac{2\tilde{\sigma}_\varepsilon^2}{MNT(N-1)T(T-1)} \begin{bmatrix} T-1 & -(T-1) \\ -(T-1) & NT-1 \end{bmatrix} \quad (\text{A.6})$$

$$\text{and } (\tilde{J}^{11})^{-1} = \frac{MNT}{2\tilde{\sigma}_\varepsilon^2} \begin{bmatrix} NT-1 & T-1 \\ T-1 & T-1 \end{bmatrix}. \quad (\text{A.7})$$

Substituting the score vector in (A.1) and (A.7) into (A.5), we obtain the KW test

$$\begin{aligned}
 KW &= \frac{MNT}{2\tilde{\sigma}_\epsilon^2} (A + B) / \sqrt{\frac{MNT(NT + 3T - 4)}{2\tilde{\sigma}_\epsilon^4}} \\
 &= \sqrt{\frac{MNT}{2(NT + 3T - 4)}} (A + B), \tag{A.8}
 \end{aligned}$$

Under the null hypothesis, the KW statistic of (A.8) is asymptotically distributed as  $N(0, 1)$ .

### Appendix 3

Let us consider the LM test for  $\sigma_\nu^2 = 0$  given the existence of random group effects. The null hypothesis for this model is  $H_0^d : \sigma_\nu^2 = 0$  (given  $\sigma_\mu^2 > 0$ ) vs  $H_1^d : \sigma_\nu^2 \neq 0$  (given  $\sigma_\mu^2 > 0$ ). Under the null hypothesis, the variance-covariance matrix is reduced as  $\Omega_0 = \sigma_\mu^2(I_M \otimes J_N \otimes J_T) + \sigma_\epsilon^2(I_M \otimes I_N \otimes I_T)$ . Following Wansbeek and Kapteyn(1982, 1983), we get  $\Omega_0^{-1} = \frac{1}{\sigma_1^2}(I_M \otimes \bar{J}_{NT}) + \frac{1}{\sigma_\epsilon^2}(I_M \otimes E_{NT})$ , where  $\bar{J}_{NT} = J_{NT}/NT$ , and  $E_{NT} = I_{NT} - \bar{J}_{NT}$ , and  $\sigma_1^2 = NT\sigma_\mu^2 + \sigma_\epsilon^2$ . Using the formula of Hemmerle and Hartly(1973), we obtain

$$\begin{aligned}
 \frac{\partial L}{\partial \sigma_\epsilon^2} &= D(\sigma_\epsilon^2) = 0, \quad \frac{\partial L}{\partial \sigma_\mu^2} = D(\sigma_\mu^2) = 0, \\
 \frac{\partial L}{\partial \sigma_\nu^2} &= \hat{D}(\sigma_\nu^2) = \frac{M(N-1)T}{2\hat{\sigma}_\epsilon^2} \left\{ \frac{M(NT-1)}{M(N-1)T} \frac{\hat{u}'[I_M \otimes E_N \otimes J_T]\hat{u}}{\hat{u}'[I_M \otimes E_{NT}]\hat{u}} - 1 \right\}, \tag{A.9}
 \end{aligned}$$

where  $\hat{u} = y - X\hat{\beta}_{GLS}$  is the GLS residuals under the null hypothesis,  $\hat{\sigma}_\epsilon^2 = \hat{u}'[I_M \otimes E_{NT}]\hat{u}/M(NT-1)$  are the solution of  $D(\sigma_\epsilon^2) = 0$ . Also, using the the formula of Harville(1977), we obtain the information matrix, when evaluated under the null hypothesis ( $\sigma_\nu^2 = 0$ ) is

$$\hat{J}_0 = \frac{1}{2} \begin{bmatrix} \frac{MN^2T^2}{\hat{\sigma}_1^4} & \frac{MNT^2}{\hat{\sigma}_1^4} & \frac{MNT}{\hat{\sigma}_1^4} \\ \frac{MNT^2}{\hat{\sigma}_1^4} & \frac{MT^2}{\hat{\sigma}_1^4} + \frac{M(N-1)T^2}{\hat{\sigma}_\epsilon^4} & \frac{MT}{\hat{\sigma}_1^4} + \frac{M(N-1)T}{\hat{\sigma}_\epsilon^4} \\ \frac{MNT}{\hat{\sigma}_1^4} & \frac{MT}{\hat{\sigma}_1^4} + \frac{M(N-1)T}{\hat{\sigma}_\epsilon^4} & \frac{M}{\hat{\sigma}_1^4} + \frac{M(NT-1)}{\hat{\sigma}_\epsilon^4} \end{bmatrix}, \tag{A.10}$$

where  $\hat{\sigma}_1^2 = \hat{u}'(I_M \otimes \bar{J}_{NT})\hat{u}/M$  is the solution of  $D(\sigma_\mu^2) = 0$ .

Thus, the  $\det(\hat{J}_0)$  is given by

$$\det(\hat{J}_0) = \frac{M^3N^3T^4(N-1)(T-1)}{8\hat{\sigma}_1^4(\hat{\sigma}_\epsilon^4)^2}, \tag{A.11}$$

where, the cofactor of (2,2)th element is

$$\begin{aligned} C(\widehat{J}_0)_{22} &= \frac{1}{4} \left[ \frac{MN^2T^2}{\widehat{\sigma}_1^4} \left( \frac{M}{\widehat{\sigma}_1^4} + \frac{M(NT-1)}{\widehat{\sigma}_\varepsilon^4} \right) - \frac{M^2N^2T^2}{(\widehat{\sigma}_1^4)^2} \right] \\ &= \frac{M^2N^2T^2(NT-1)}{4\widehat{\sigma}_1^4\widehat{\sigma}_\varepsilon^4}. \end{aligned} \quad (\text{A.12})$$

Therefore,

$$(\widehat{J}_0^{-1})_{22} = \frac{C(\widehat{J}_0)_{22}}{\det(\widehat{J}_0)} = \frac{2(NT-1)}{MNT^2(N-1)(T-1)} \widehat{\sigma}_\varepsilon^4. \quad (\text{A.13})$$

Therefore, the resulting LM test statistic is

$$\begin{aligned} LM &= \widehat{D}'\widehat{J}_0\widehat{D} \\ &= (\widehat{D}(\sigma_\nu^2))^2 (\widehat{J}_0^{-1})_{22} \\ &= \frac{M(N-1)(NT-1)}{2N(T-1)} \left( \frac{NT-1}{N-1} \frac{\widehat{u}'(I_M \otimes E_N \otimes \bar{J}_T)\widehat{u}}{\widehat{u}'(I_M \otimes E_{NT})\widehat{u}} - 1 \right)^2, \end{aligned} \quad (\text{A.14})$$

where  $E_N = I_N - \bar{J}_N$ , and  $\bar{J}_N = J_N/N$ . Under the null hypothesis, LM is asymptotically distributed as  $\chi^2(1)$ .

And the one-sided LM test is

$$LM_\nu = \sqrt{\frac{M(N-1)(NT-1)}{2N(T-1)}} \left( \frac{NT-1}{N-1} \frac{\widehat{u}'(I_M \otimes E_N \otimes \bar{J}_T)\widehat{u}}{\widehat{u}'(I_M \otimes E_{NT})\widehat{u}} - 1 \right), \quad (\text{A.15})$$

which is asymptotically distributed as  $N(0, 1)$ , see Baltagi, et al.(1992).





Table 4. Rejection number of  $H_0^c : \sigma_\nu^2 = 0$

$\gamma_1$	$\gamma_2$	$(M, N, T) = (5, 5, 5)$					$(M, N, T) = (10, 5, 5)$					$(M, N, T) = (5, 10, 10)$				
		$LM^c$	$LM_3$	$SLM_2$	LR	F	$LM^c$	$LM_3$	$SLM_2$	LR	F	$LM^c$	$LM_3$	$SLM_2$	LR	F
0.00	0.00	62	49	50	36	54	82	58	56	31	46	53	59	61	30	46
0.00	0.05	157	199	208	157	178	226	301	292	256	258	594	691	703	593	605
0.00	0.10	352	424	439	332	345	569	658	648	585	536	935	958	958	942	938
0.00	0.20	746	799	808	769	736	962	978	977	967	935	1000	1000	1000	1000	1000
0.00	0.40	991	996	996	992	981	1000	1000	1000	1000	999	1000	1000	1000	1000	1000
0.00	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.05	0.00	145	180	183	136	37	215	298	292	242	35	447	510	521	447	30
0.05	0.05	326	399	407	331	154	573	652	639	567	188	889	925	927	893	578
0.05	0.10	560	644	653	521	315	837	876	872	825	487	988	994	994	992	936
0.05	0.20	868	911	915	869	700	986	994	994	989	917	1000	1000	1000	1000	999
0.05	0.40	998	998	999	998	977	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.05	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.10	0.00	278	331	341	302	28	530	596	592	522	22	714	755	760	722	25
0.10	0.05	500	566	577	495	121	764	824	819	802	141	958	973	974	964	515
0.10	0.10	711	775	782	709	300	936	957	954	937	456	999	999	999	1000	927
0.10	0.20	933	951	952	932	670	1000	1000	1000	998	879	1000	1000	1000	1000	1000
0.10	0.40	998	999	999	1000	967	1000	1000	1000	1000	999	1000	1000	1000	1000	1000
0.10	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.20	0.00	554	605	610	570	14	834	869	865	868	13	897	915	918	909	17
0.20	0.05	744	796	799	749	75	968	978	977	962	79	995	999	999	991	434
0.20	0.10	865	895	900	865	248	997	998	998	995	318	999	999	999	1000	875
0.20	0.20	979	986	987	980	600	1000	1000	1000	1000	814	1000	1000	1000	1000	999
0.20	0.40	1000	1000	1000	1000	949	1000	1000	1000	1000	997	1000	1000	1000	1000	1000
0.20	0.60	1000	1000	1000	1000	995	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.40	0.00	868	887	892	860	4	985	988	988	993	0	975	978	979	981	4
0.40	0.05	939	957	958	930	55	1000	1000	1000	1000	23	999	1000	1000	1000	324
0.40	0.10	972	986	986	980	185	1000	1000	1000	1000	153	1000	1000	1000	1000	749
0.40	0.20	1000	1000	1000	1000	552	1000	1000	1000	1000	613	1000	1000	1000	1000	981
0.40	0.40	1000	1000	1000	1000	918	1000	1000	1000	1000	970	1000	1000	1000	1000	1000
0.60	0.00	966	973	975	965	2	1000	1000	1000	1000	0	998	998	998	996	1
0.60	0.05	993	995	995	989	43	1000	1000	1000	1000	10	1000	1000	1000	1000	281
0.60	0.10	998	1000	1000	999	154	1000	1000	1000	1000	85	1000	1000	1000	1000	674
0.60	0.20	1000	1000	1000	1000	494	1000	1000	1000	1000	466	1000	1000	1000	1000	941
0.80	0.00	996	996	996	993	0	1000	1000	1000	1000	0	999	999	999	1000	0
0.80	0.05	1000	1000	1000	1000	40	1000	1000	1000	1000	7	1000	1000	1000	1000	237
0.80	0.10	1000	1000	1000	1000	162	1000	1000	1000	1000	59	1000	1000	1000	1000	552

Table 5. Rejection number of  $H_0^d : \sigma_\nu^2 = 0$  (given  $\sigma_\mu^2 > 0$ )

$\gamma_1$	$\gamma_2$	$(M, N, T) = (5, 5, 5)$			$(M, N, T) = (10, 5, 5)$			$(M, N, T) = (5, 10, 10)$		
		$LM_\nu$	LR	F	$LM_\nu$	LR	F	$LM_\nu$	LR	F
0.00	0.00	62	30	56	48	20	42	55	25	44
0.00	0.05	205	109	188	295	188	281	638	545	615
0.00	0.10	394	271	370	594	495	573	946	925	942
0.00	0.20	788	708	769	963	945	956	1000	999	1000
0.00	0.40	991	987	991	1000	1000	1000	1000	1000	1000
0.00	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.05	0.00	61	38	54	60	45	55	60	33	48
0.05	0.05	221	162	196	282	235	265	679	635	655
0.05	0.10	418	324	390	631	579	610	972	965	969
0.05	0.20	801	747	780	969	957	964	1000	1000	1000
0.05	0.40	995	991	994	1000	1000	1000	1000	1000	1000
0.05	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.10	0.00	61	44	54	58	49	53	60	49	54
0.10	0.05	201	167	191	318	274	292	712	666	686
0.10	0.10	446	371	415	697	656	677	983	978	982
0.10	0.20	831	796	812	980	970	976	1000	1000	1000
0.10	0.40	998	997	997	1000	1000	1000	1000	1000	1000
0.10	0.80	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.20	0.00	58	43	51	63	52	57	65	50	56
0.20	0.05	228	185	207	362	328	346	789	744	761
0.20	0.10	505	447	474	763	731	739	994	988	989
0.20	0.20	873	842	864	986	983	985	1000	1000	1000
0.20	0.40	999	999	1000	1000	1000	1000	1000	1000	1000
0.20	0.60	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.40	0.00	68	48	59	61	53	58	51	36	43
0.40	0.05	327	284	304	497	459	479	910	887	898
0.40	0.10	700	650	672	887	867	874	998	998	998
0.40	0.20	974	964	970	1000	1000	1000	1000	1000	1000
0.40	0.40	1000	1000	1000	1000	1000	1000	1000	1000	1000
0.60	0.00	65	53	59	61	49	53	71	51	57
0.60	0.05	505	453	480	701	675	684	988	984	987
0.60	0.10	906	879	891	989	986	988	1000	1000	1000
0.60	0.20	1000	999	999	1000	1000	1000	1000	1000	1000
0.80	0.00	60	43	49	57	47	52	69	55	61
0.80	0.05	904	879	891	991	987	990	1000	1000	1000
0.80	0.10	998	998	999	1000	1000	1000	1000	1000	1000



## REFERENCES

- Batagi, B. (1993). "Nested effects," *Econometric Theory*, **9**, 687-688.
- Baltagi, B.H., Chang, Y.J. and Li, Q. (1992). "Monte Carlo results on several new and existing tests for the error component model," *Journal of Econometrics*, **54**, 95-120.
- Baltagi, B.H. and Li, Q. (1995). "Testing AR(1) against MA(1) disturbances in an error component model," *Journal of Econometrics*, **68**, 133-151.
- Baltagi, B.H. and Li, Q. (1991). "A joint test for serial correlation and random individual effects," *Statistics and Probability letters*, **11**, 277-280.
- Bera, A.K. and Jarque, C.M. (1982). "Model specification tests: a simultaneous approach," *Journal of Econometrics*, **20**, 59-82.
- Evans, M.A. and King, M.L. (1989). "Critical value approximations for tests of linear regression disturbances," *Australian Journal of Statistics*, **27**, 68-83.
- Godfrey, L.G. (1989). *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.
- Gourieroux, C., Holly, A. and A. Monfort, A. (1982). "Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters," *Econometrica*, **50**, 63-80.
- Harville, D.A. (1977). "Maximum likelihood approaches to variance component estimation and to related problems," *Journal of the American Statistical Association*, **72**, 320-338.
- Hemmerle, W.J. and Hartley, H.O. (1973). "Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation." *Technometrics*, **15**, 819-831.
- Honda, Y. (1985). "Testing the error components model with non-normal disturbances," *Review of Economic Studies*, **52**, 681-690.
- Jung, B.C., Jhun, M. and Song, S.H. (1999). "The asymptotic property of LM test statistic for the nested error component regression model," *Working papers*, Department of Statistics, Korea University.
- King, M.L., and Wu, P.X. (1997), "Locally optimal on sided tests for multiparameter hypotheses," *Econometric Reviews*, **16**, 131-156.

- Moulton, B.R. and Randolph, W.C. (1989). "Alternative Tests of the error components model," *Econometrica*, **57**, 685-693.
- Nerlove, M. (1971). "Further evidence on the estimation of dynamic economic relations from a time-series of cross-sections," *Econometrica* **39**, 359-382.
- Wansbeek, T. and Kapteyn, A. (1982). "A simple way to obtain the spectral decomposition of variance components models for balanced data," *Communications in Statistics - Theory and Method*, **11**, 2105-2111.
- Wansbeek, T. and Kapteyn, A. (1983). "A note on spectral decomposition and maximum likelihood estimation in ANOVA models with balanced data," *Statistics and Probability Letters*, **1**, 213-215.