C | E | D | L | A | S

Centro de Estudios
Distributivos, Laborales y Sociales

Maestría en Economía
Universidad Nacional de La Plata

Tests for Unbalanced Error Component Models Under Local Misspecication

Walter Sosa Escudero y Anil K. Bera

Documento de Trabajo Nro. 65
Marzo, 2008

www.depeco.econo.unlp.edu.ar/cedlas

Tests for Unbalanced Error Component Models Under Local Misspecification

Walter Sosa Escudero[*]

Department of Economics

Universidad de San Andrés

Anil K. Bera

Department of Economics

University of Illinois

December 20, 2007

Abstract: This paper derives *unbalanced* versions of the tests statistics for first order serial correlation and random individual effects summarized in Sosa Escudero and Bera (2001), and updates their `xttest1` routine. The derived tests statistics should be useful for applied researchers faced with the increasing availability of panel information where not every individual or country is observed for the full time span. Also, as it was the case of the previously available tests, the test statistics proposed here are based on the OLS residuals, and hence are computationally very simple.

JEL Classification: C12, C23, C52
Keywords: error components model, unbalanced panel data, testing, misspecification.

Introduction

A standard specification check that accompanies the output of almost every estimated error component model is a simple test for the presence of random individual effects. The well-known Breusch-Pagan statistic (Breusch and Pagan, 1980) based on the Rao-score (RS) principle, is a frequent choice. Bera, Sosa-Escudero and Yoon (2001, BSY(2001) hereafter) demonstrated that in the presence of first order serial correlation the test rejects the correct null hypothesis of no random effects too often. Consequently, they propose a modified version that is not affected by the presence of local

serial correlation. A similar concern affects the standard test for first order serial correlation derived by Baltagi and Li (1991), which overrejects the true null hypothesis when random effects are present. For this case, an adjusted RS test was also derived by BSY (2001). These test statistics, along with their `xttest1` routine in Stata and some empirical illustrations, are presented in Sosa-Escudero and Bera (2001, SB(2001) hereafter). For a textbook exposition, see Baltagi (2005, pp. 96-97).

These test procedures are originally derived for the *balanced* case, that is, in the panel data terminology, the case where all individuals are observed for the same number of periods, and in every period all individuals are observed. On the other hand, in applied work the availability of *unbalanced* panels is far from being an uncommon situation. Though in some cases statistical procedures designed for the balanced case can be straightforwardly extended to accomodate unbalanced panels, many estimation or test procedures require less trivial modifications.

Baltagi and Li (1990) derived an unbalanced version of the Breusch-Pagan statistic. The purpose of this paper is to derive unbalanced versions of the test for first order serial correlation originally proposed by Baltagi and Li (1991), and of the modified tests proposed by BSY(2001). As a simple extension we also derive an unbalanced version of the joint test of serial correlation and random effects proposed by Baltagi and Li (1991). Also, as it is the case of the previously available tests, the derived test statistics, being based on OLS residuals after pooled estimation, are computationally very simple. Finally, the SB(2001) `xttest1` routine is appropriately updated to handle unbalanced panels.

Tests for the unbalanced case

Consider a simple linear model for panel data allowing for the presence of random individual effects and first order serial correlation

$$
\begin{aligned}
y_{it} &= x_{it}'\beta + u_{it}, \\
u_{it} &= \mu_i + \nu_{it}, \\
\nu_{it} &= \lambda \nu_{i,t-1} + \epsilon_{it}, \quad |\lambda| < 1,
\end{aligned}
$$

where x_{it} is a $(k \times 1)$ vector of explanatory variables with 1 in its first position, β is a $(k \times 1)$ vector of parameters including an intercept, $\mu_i \sim IIDN(0, \sigma_\mu^2)$ and $\epsilon_{it} \sim IIDN(0, \sigma_\epsilon^2)$. We will assume $\nu_{i,0} \sim N(0, \sigma_\epsilon^2/(1 - \lambda^2))$.

We will be interested in testing for the absence of random effects ($H_0 : \sigma_\mu^2 = 0$) and/or first order serial correlation ($H_0 : \lambda = 0$). The panel will be *unbalanced* in the sense that for every individual

$i = 1, \ldots, N$ we will observe, possibly, a different number of time observations T_i. We will restrict the analysis to the case where missing observations occur either at the beginning or at the end of the sample period for each individual (that is, there are no 'gaps' in the series), and the starting and final periods are determined randomly. Hence, without loss of generality and in order to avoid complicating the notation too much, we can safely assume that series for each individual start at the same period ($t = 1$) and finish randomly at period $t = T_i$.

Let $m = \sum_{i=1}^{N} T_i$ be the total number of observations. Let u be an $m \times 1$ vector with typical element u_{it} where observations are sorted first by individuals and then by time, so the time index is the faster one. Then in our setup, $V(u) \equiv \Omega$ can be written as

$$V(u) = \sigma_\mu^2 \tilde{H} + \sigma_\epsilon^2 \tilde{V},$$

where $\tilde{H}$ is an $m \times m$ block diagonal matrix with blocks H_i equal to matrices of ones, each with dimension $T_i \times T_i$. Similarly, $\tilde{V}$ will be a block diagonal $m \times m$ matrix with blocks V_i equal to

$$V_i = \begin{bmatrix} 1 & \lambda & \lambda^2 & \cdots & \lambda^{T_i - 1} \\ \lambda & 1 & \lambda & \cdots & \lambda^{T_i - 2} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda^{T_i - 1} & \lambda^{T_i - 2} & \lambda^{T_i - 3} & \cdots & 1 \end{bmatrix}.$$

For the purpose of deriving the test statistics, the log-likelihood function will be

$$L(\beta, \lambda, \sigma_\epsilon^2, \sigma_\mu^2) = \text{constant} - \frac{1}{2} \log |\Omega| - \frac{1}{2} u' \Omega^{-1} u.$$

The information matrix for this problem is well-known to be block diagonal between β and the remaining parameters. Therefore, for the purposes of this paper, we will concentrate only on the parameters λ, σ_μ^2 and σ_ϵ^2. Under a more general setup, suppose the log-likelihood can be characterized by a three parameter vector $\theta = (\psi, \phi, \gamma)'$. Let $d(\theta)$ be the score vector and $J(\theta)$ the information matrix. If it can be assumed that $\phi = 0$, the standard Rao-score test statistic for the null hypothesis $H_o : \psi = 0$ is given by

$$RS_\psi = d_\psi(\hat{\theta}) J_{\psi \cdot \gamma}^{-1}(\hat{\theta}) d_\psi(\hat{\theta}), \tag{1}$$

where d_ψ is the element of the score corresponding to the parameter ψ, $J_{\psi \cdot \gamma}(\theta) = J_\psi - J_{\psi \gamma} J_\gamma^{-1} J_{\gamma \psi}$, and $\hat{\theta}$ is the maximum likelihood estimator of θ under the restriction implied by the null hypothesis and the assumption $\phi = 0$. Asymptotically, this test statistic under the null hypothesis $H_0 : \psi = 0$ is well known to have a central chi-squared distribution. In the context of our error components

model, if $\gamma = \sigma_\epsilon^2$ and if we set $\psi = \sigma_\mu^2$ and $\phi = \lambda$, (1) is a test for random effects assuming no serial correlation, and if instead we set $\psi = \lambda$ and $\phi = \sigma_\mu^2$, (1) gives a test for serial correlation assuming no random effects. The standard Breusch-Pagan (1980) test for random effects (assuming no serial correlation) and the Baltagi and Li (1991) test for first order serial correlation (assuming no random effects) are derived from this principle.

Bera and Yoon (1993) showed that the test statistic (1) is invalid when $\phi \neq 0$, in the sense that the test tends to reject the null hypothesis too frequently even when it is correct. More specifically, the RS_ψ statistic is found to have an asymptotic *non-central* chi-squared distribution under $H_o : \psi = 0$, when $\phi = \delta/\sqrt{n}$, that is, when the alternative is *locally misspecified*. In particular, this implies that when the null is correct, the Breusch-Pagan test tends to reject the true null of absence of random effects if the error term is serially correlated, even in a local sense. A similar situation arises for the test for serial correlation of Baltagi and Li (1991) in the local presence of random effects. In order to remedy this problem, Bera and Yoon (1993) proposed the following modified Rao-score statistic

$$
\begin{aligned}
RS_\psi^* \;\; = \;\; & \frac{1}{n} \big[d_\psi(\hat{\theta}) - J_{\psi\phi\cdot\gamma}(\hat{\theta}) J_{\phi\cdot\gamma}^{-1}(\hat{\theta}) d_\phi(\hat{\theta}) \big]' \\
& \big[J_{\psi\cdot\gamma}(\hat{\theta}) - J_{\psi\phi\cdot\gamma}(\hat{\theta}) J_{\phi\cdot\gamma}^{-1}(\hat{\theta}) J_{\phi\psi\cdot\gamma}(\hat{\theta}) \big]^{-1} \\
& \big[d_\psi(\hat{\theta}) - J_{\psi\phi\cdot\gamma}(\hat{\theta}) J_{\phi\cdot\gamma}^{-1}(\hat{\theta}) d_\phi(\hat{\theta}) \big],
\end{aligned}
\tag{2}
$$

where $\hat{\theta}$ is the MLE of θ under the joint null $\psi = \phi = 0$. This modified test statistic has an asymptotic *central* χ_1^2 distribution under the null hypothesis $H_o : \psi = 0$ and when $\phi = \delta/\sqrt{n}$, that is, the modified test statistic has the correct size even when the underlying model is locally misspecified. Based on this principle, BSY (2001) derived modified tests for random effects (serial correlation) which are valid in the presence of local first order serial correlation (random effects), but assuming that the panel is balanced.

In order to derive tests for the unbalanced case, let $\theta = (\lambda, \sigma_\mu^2, \sigma_\epsilon^2)'$ and $\hat{\theta} = (0, 0, \hat{\sigma}_\epsilon^2)'$ be the maximum likelihood estimator of θ under the joint null hypothesis $H_0 : \lambda = \sigma_\mu^2 = 0$. The following formula by Hemmerle and Hartley (1973) will be useful to derive the score vector for the problem

$$
d_{\theta_r} \equiv \frac{\partial L}{\partial \theta_r} = -\frac{1}{2} \, tr \left[\Omega^{-1} \frac{\partial \Omega}{\partial \theta_r} \right] + \frac{1}{2} \left[u' \Omega^{-1} \frac{\partial \Omega}{\partial \theta_r} \Omega^{-1} u \right],
\tag{3}
$$

where θ_r denotes the r-th element of θ, $r = 1, 2, 3$. Note that $\partial\Omega/\partial\sigma_\mu^2 = \tilde{H}$, with $tr(\tilde{H}) = m$. Similarly, $\partial\Omega/\partial\sigma_\epsilon^2 = \tilde{V}$, which under the restricted MLE is an $m \times m$ identity matrix with trace equal to m. Also, $\partial\Omega/\partial\lambda = \sigma_\epsilon^2 \tilde{G}$, where $\tilde{G}$ is a block diagonal matrix with blocks equal to G_i, with $G_i = \partial V_i/\partial\lambda$ given by

4

$$
G_i = \begin{bmatrix}
0 & 1 & 2\lambda & \cdots & (T_i - 1)\lambda^{T_i-2} \\
1 & 0 & 1 & \cdots & (T_i - 2)\lambda^{T_i-3} \\
\vdots & & \vdots & \vdots & \vdots \\
\vdots & & \vdots & 1 & 0 & 1 \\
(T_i - 1)\lambda^{T_i-2} & \cdots & \cdots & 1 & 0
\end{bmatrix}.
$$

Under the restricted MLE, G_i is a bi-diagonal matrix as follows

$$
G_i(\hat{\theta}) = \begin{bmatrix}
0 & 1 & 0 & & \cdots & 0 \\
1 & 0 & 1 & & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & & 1 & 0 & 1 \\
0 & \cdots & & \cdots & 1 & 0
\end{bmatrix},
$$

hence $tr(G_i(\hat{\theta})) = 0$. Replacing these results in formula (3) and evaluating the expression under the restricted MLE we obtain

$$
\begin{aligned}
d_{\sigma_\mu^2}(\hat{\theta}) &= -\frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}I_m \tilde{H}\right] + \frac{1}{2}e'\frac{1}{\hat{\sigma}_\epsilon^2}I_m \tilde{H}\frac{1}{\hat{\sigma}_\epsilon^2}e \\
&= -\frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^2}m + \frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^4}e'\tilde{H}e = -\frac{m}{\hat{\sigma}_\epsilon^2}A,
\end{aligned}
$$

where e is an $m \times 1$ vector with typical element $e_{it} = x'_{it}\hat{\beta}$ and $\hat{\beta}$ is the restricted MLE of β. Similarly, $\hat{\sigma}_\epsilon^2 = e'e/m$ is the restricted MLE of σ_ϵ^2, and $A \equiv 1 - e'\tilde{H}e/(e'e)$. In a similar fashion,

$$
\begin{aligned}
d_\lambda(\hat{\theta}) &= -\frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\hat{\sigma}_\epsilon^2\tilde{G}(\hat{\theta})\right] + \frac{1}{2}2\frac{1}{\hat{\sigma}_\epsilon^2}e'\tilde{G}(\hat{\theta})e \\
&= \frac{1}{\hat{\sigma}_\epsilon^2}e'\tilde{G}(\hat{\theta})\, e = m\, B, \quad \text{say,}
\end{aligned}
$$

with $B \equiv e'\tilde{G}e/e'e$.

To derive the elements of the information matrix we will use the following formula in Baltagi (2005, pp. 59-60)

$$
J_{r,s}(\theta) = E\left[-\frac{\partial^2 L}{\partial\theta_r \partial\theta_s}\right] = \frac{1}{2}tr\left[\Omega^{-1}\frac{\partial\Omega}{\partial\theta_r}\Omega^{-1}\frac{\partial\Omega}{\partial\theta_s}\right].
$$

Then,

$$
J_{\sigma_\epsilon^2,\sigma_\epsilon^2}(\hat{\theta}) = \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{V}(\hat{\theta})\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{V}(\hat{\theta})\right] = \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^4}I_m\right] = \frac{m}{2\hat{\sigma}_\epsilon^4}
$$

5

$$
\begin{aligned}
J_{\hat{\sigma}_\mu^2,\hat{\sigma}_\mu^2}(\hat{\theta}) &= \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{H}\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{H}\right] = \frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^4}tr\left[\tilde{H}\tilde{H}\right] = \frac{\sum_{i=1}^N T_i^2}{2\hat{\sigma}_\epsilon^4} \\
J_{\lambda,\lambda}(\hat{\theta}) &= \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\hat{\sigma}_\epsilon^2\tilde{G}(\hat{\theta})\frac{1}{\hat{\sigma}_\epsilon^2}\hat{\sigma}_\epsilon^2\tilde{G}(\hat{\theta})\right] = \frac{1}{2}tr\left[\tilde{G}(\hat{\theta})\tilde{G}(\hat{\theta})\right] = \frac{1}{2}\sum_{i=1}^N 2(T_i-1) = m-N \\
J_{\hat{\sigma}_\epsilon^2,\hat{\sigma}_\mu^2}(\hat{\theta}) &= \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{V}(\hat{\theta})\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{V}(\hat{\theta})\right] = \frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^4}tr\left[\tilde{H}\right] = \frac{m}{2\hat{\sigma}_\epsilon^4} \\
J_{\hat{\sigma}_\epsilon^2,\lambda}(\hat{\theta}) &= \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{V}(\hat{\theta})\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{G}(\hat{\theta})\right] = \frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^4}tr\left[\tilde{G}(\hat{\theta})\right] = 0 \\
J_{\lambda,\hat{\sigma}_\mu^2}(\hat{\theta}) &= \frac{1}{2}tr\left[\frac{1}{\hat{\sigma}_\epsilon^2}\hat{\sigma}_\epsilon^2\tilde{G}(\hat{\theta})\frac{1}{\hat{\sigma}_\epsilon^2}\tilde{H}\right] = \frac{1}{2}\frac{1}{\hat{\sigma}_\epsilon^2}tr\left[\tilde{G}(\hat{\theta})\tilde{H}\right] = \frac{1}{2}\frac{2}{\hat{\sigma}_\epsilon^2}\left(\sum_{i=1}^N T_i - N\right) = \frac{1}{\hat{\sigma}_\epsilon^2}(m-N),
\end{aligned}
$$

where we have used the facts that $tr\left(\tilde{G}_i(\hat{\theta})\tilde{G}_i(\hat{\theta})\right) = tr\left(\tilde{G}_i(\hat{\theta})\tilde{H}_i\right) = 2(T_i-1)$, and $tr(\tilde{H}_i\tilde{H}_i) = T_i^2$. Collecting all the elements, the information matrix evaluated at the restricted MLE under the joint null, can be expressed as

$$
J(\hat{\theta}) = \frac{1}{2\hat{\sigma}_\epsilon^4}\begin{bmatrix} m & m & 0 \\ m & a & 2\hat{\sigma}_\epsilon^2(m-N) \\ 0 & 2\hat{\sigma}_\epsilon^2(m-N) & 2\hat{\sigma}_\epsilon^4(m-N) \end{bmatrix},
$$

where $a \equiv \sum_{i=1}^N T_i^2$. Note that for the balanced case $T_i = T$ we get exactly the same expression for $J(\hat{\theta})$ as in Baltagi and Li (1991, p. 279). From the above expression of $J(\hat{\theta})$, we can show that

$$
\begin{aligned}
J_{\mu\lambda\cdot\sigma_\epsilon^2} &= \frac{m-N}{\hat{\sigma}_\epsilon^2} \\
J_{\mu\cdot\sigma_\epsilon^2} &= \frac{a-m}{2\hat{\sigma}_\epsilon^4} \\
J_{\lambda\cdot\sigma_\epsilon^2} &= m-N.
\end{aligned}
$$

Substituting these results in (2), we obtain the unbalanced version of the modified test for random effects as

$$
RS_\mu^* = \frac{m^2\left[A+2B\right]^2}{2\left(a-3m+2N\right)}.
$$

When $T_i = T$ (the balanced case), the above expression boils down to

$$
RS_\mu^* = \frac{NT\left[A+2B\right]^2}{2(T-1)(1-(2/T))},
$$

as in BSY(2001) for the balanced case.

Similarly, the modified test statistic for serial correlation is

$$
RS_\lambda^* = \left[B + \frac{m-N}{a-m}A\right]^2 \frac{(a-m)m^2}{(m-N)(a-3m+2N)},
$$

and when $T_i = T$, we get

$$RS_\lambda^* = \left[B + \frac{A}{T}\right]^2 \frac{NT^2}{(T-1)(1-2/T)},$$

the expression in BSY(2001) for the balanced case.

For computational purposes, it is interesting to see that

$$A = 1 - \frac{\sum_{i=1}^{N}\left(\sum_{t=1}^{T_i} e_{it}^2\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T_i} e_{it}^2}$$

and

$$B = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T_i} e_{i,t}e_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T_i} e_{it}^2},$$

and therefore, there is no need to construct the $\tilde{G}$ or $\tilde{H}$ matrices, and hence the tests statistics can be easily computed right after OLS estimation without constructing any matrices.

The previous derivations allow us to obtain the unbalanced version of the test for serial correlation assuming no random effects

$$RS_\lambda = \frac{m^2 B^2}{m - N},$$

which again reduces to $NT^2 B^2/(T-1)$, originally derived by Baltagi and Li (1991) for balanced panels. Also, for completeness, the unbalanced version of the test for random effects assuming no serial correlation is given by

$$RS_\mu = \frac{\frac{1}{2}m^2 A^2}{a - m}.$$

This test statistic is a particular case of the Baltagi and Li (1991) test for the two-way error component model.

Suppose we are interested in the *joint* null hypothesis of no random effects and no first order serial correlation. Let $RS_{\phi,\psi}$ be the Rao-Score test statistic for the joint null hypothesis $H_o : \phi = \psi = 0$. Bera and Yoon (2001) show that the following identities hold

$$RS_{\phi\psi} = RS_\psi^* + RS_\phi = RS_\phi^* + RS_\psi,$$

which simplifies computations, as illustrated in Sosa Escudero and Bera (2001). Then, as a simple by-product of the previous derivations, we can obtain a statistic for jointly testing for serial correlation and random effects, as

$$RS_{\lambda\mu} = m^2 \left[\frac{A^2 + 4AB + 4B^2}{2(a - 3m + 2N)} + \frac{B^2}{m - N} \right].$$

When $T_i = T$, $RS_{\lambda\mu}$ simplifies to

$$RS_{\lambda\mu} = \frac{NT^2}{2(T - 1)(T - 2)} \left[A^2 + 4AB + 2TB^2 \right],$$

the original joint test statistic of Baltagi and Li (1991).

Finally, since $\sigma_\mu^2 \geq 0$, it is natural to consider one-sided versions of the tests for the null $H_0 : \sigma_\mu^2 = 0$. As in BSY(2001), appropriate tests statistics can be readily constructed by taking the signed square roots of the original two-sided tests RS_μ and RS_μ^*. Denoting their one-sided versions, respectively, as RSO_μ and RSO_μ^*, we have

$$RSO_\mu = -\sqrt{\frac{\frac{1}{2} m^2}{a - m}} \, A$$

and

$$RSO_\mu^* = -\sqrt{\frac{m^2}{2(a - 3m + 2N)}} \, [A + 2B].$$

Empirical illustration

As an illustration of these procedures we provide an empirical exercise based on Gasparini, Marchionni and Sosa Escudero (2001). It consists in a simple linear panel data model where the dependent variable is the Gini coefficient for 17 regions of Argentina. The vector of explanatory variables includes mean income and its square (*ie* and *ie2*), proportion of the population employed in the manufacturing industry (*indust*) and in public adminstration, health or education (*public*), unemployment rate (*unemp*), activity rate (*tactiv*), public investment as percentege of GDP (*invgdp*), degree of openess (*open*), social assistance (*pyas4*), proportion of population older than 64 (*e64*), proportion of population with complete high school (*supc*), and average family size (*tamfam*); for details see Gasparini et al. (2001). Models of this type have been used extensively in the literature exploring the links between inequality and development, usually to study the so-called 'Kuznets hypothesis' which postulates an inverted U-shaped relationship between these two variables [for example, see Anand and Kanbur (1993) and Gustafsson and Johansson (1999)].

Income related variables, including the Gini coefficients, are constructed using Argentina's *Permanent Household Survey* (Encuesta Permamente de Hogares) which surveys several socioeconomic

variables at the household level for several regions of the country. Due to certain administrative deficiencies the panel is largely unbalanced, so the number of available temporal observations ranges from five to eight years in the period 1992 to 2000.

First we `tsset` the data and estimate the parameters of a one-way error components model with region-specific random effects, using `xtreg`:

```
. tsset naglo ano
      panel variable:  naglo, 2 to 33
       time variable:  ano, 1992 to 2000, but with a gap

. xtreg gini ie ie2 indus adpubedsal desempleo tactiv invipib apertura pyas4 e64 supc tamfam, re i(naglo)

Random-effects GLS regression             Number of obs      =         128
Group variable (i): naglo                 Number of groups   =          17

R-sq:  within  = 0.5096                    Obs per group: min =           6
       between = 0.6153                                   avg =         7.5
       overall = 0.5344                                   max =           8

Random effects u_i ~ Gaussian              Wald chi2(12)      =      121.30
corr(u_i, X)       = 0 (assumed)           Prob > chi2        =      0.0000

------------------------------------------------------------------------------
        gini |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ie |  -.0000995   .0001823    -0.55   0.585    -.0004568    .0002578
         ie2 |   1.64e-08   2.19e-07     0.08   0.940    -4.12e-07    4.45e-07
       indus |   -.041974   .0704982    -0.60   0.552    -.1801478    .0961999
   adpubedsal |  -.0635789   .0531777    -1.20   0.232    -.1678053    .0406475
   desempleo |  -.1177452   .0638999    -1.84   0.065    -.2429868    .0074963
      tactiv |   .0999584   .0737997     1.35   0.176    -.0446864    .2446031
     invipib |  -.3307239   .1912258    -1.73   0.084    -.7055197    .0440718
    apertura |   .4289793   .0768693     5.58   0.000     .2783183    .5796404
       pyas4 |   2.884162   1.626136     1.77   0.076    -.3030061    6.071331
         e64 |  -.1339182   .1505384    -0.89   0.374    -.4289681    .1611316
        supc |   .2417907   .0946423     2.55   0.011     .0562952    .4272861
      tamfam |   .0169905   .0174328     0.97   0.330    -.0171771    .0511581
       _cons |   .3084864   .1031351     2.99   0.003     .1063453    .5106274
-------------+----------------------------------------------------------------
     sigma_u |  .01370805
     sigma_e |  .01377936
       lambda |  .49740589   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Next, the command `xttest1` with the 'unadjusted' option presents the following output:

```
. xttest1, unadjusted

Tests for the error component model:

        gini[naglo,t] = Xb + u[naglo] + v[naglo,t]
           v[naglo,t] = lambda v[naglo,(t-1)] + e[naglo,t]

        Estimated results:
```

```
                    Var      sd = sqrt(Var)
      ---------+----------------------------
         gini |   .0006167         .0248335
            e |   .0001899         .01377936
            u |   .0001879         .01370805

Tests:
   Random Effects, Two Sided:
   LM(Var(u)=0)          =   13.50 Pr>chi2(1) =  0.0002
   ALM(Var(u)=0)         =    6.03 Pr>chi2(1) =  0.0141

   Random Effects, One Sided:
   LM(Var(u)=0)          =    3.67 Pr>N(0,1)  =  0.0001
   ALM(Var(u)=0)         =    2.46 Pr>N(0,1)  =  0.0070

   Serial Correlation:
   LM(lambda=0)          =    9.32 Pr>chi2(1) =  0.0023
   ALM(lambda=0)         =    1.86 Pr>chi2(1) =  0.1732

   Joint Test:
   LM(Var(u)=0,lambda=0) =   15.35 Pr>chi2(2) =  0.0005
```

The first part of the table is identical to the one produced by `xttest0`, presenting estimates of the corresponding variance components. Next, the table presents seven test statistics and their corresponding p-values. `LM(Var(u)=0)` and `ALM(Var(u)=0)` are, respectively, the test statistics for random effects and their adjusted versions, that is, RS_μ and RS_μ^* in our earlier notation. `LMO(Var(u)=0)` and `LMO(Var(u)=0)` are their one-sided versions. `LM(lambda=0)` and `ALM(lambda=0)` are the test statistics for serial correlation RS_λ and RS_λ^*, and `LM(Var(u)=0,lambda=0)` is the joint test statistic $RS_{\lambda\mu}$ for testing no random effects and no serial correlation.

The unadjasted version of the tests for random effects (`LM(Var(u)=0)`) and serial correlation (`LM(lambda=0)`), and the test for the joint null (`LM(Var(u)=0,lambda=0)`) suggest rejecting their nulls at the 5% significance level. Care must be taken in deriving conclusions about the direction of the misspecification since in light of the results in BSY(2001) rejections may arise due to the presence either of random effects, serial correlation, or both. In order to explore the possible nature of the misspecification, we restore to the modified versions of the test. The adjusted version of the test for serial correlation `ALM(lambda=0)` now fails to reject the null hypothesis while the adjusted version of the test for random effect `ALM(Var(u)=0)` still does. This suggests that the possible misspecification is more likely due to the presence of random effects rather than the serial correlation. Consequently, and to stress the main usefulness of these procedures, in this example the presence of the random effect seems to confound the unadjusted test for serial correlation making it spuriously reject its null.

References

Anand, S. and Kanbur, R., 1993, Inequality and Development: A Critique, *Journal of Development Economics*, 41, 19-43.

Baltagi, B., 2005, *Econometric Analysis of Panel Data, 3rd. edition*, New York: Wiley & Sons.

Baltagi, B. and Q. Li, 1990, A Lagrange Multiplier Test for the Error Components Model with Incomplete Panels, *Econometric Reviews*, 9(1), 103-107.

Baltagi, B. and Q. Li, 1991, A joint test for serial correlation and random individual effects, *Statistics & Probability Letters*, 11, 277–280.

Bera, A. and Yoon, M., 1993, Specification teting with misspecified alternatives, *Econometric Theory*, 9, 649-658.

Bera, A. and Yoon, M., 2001, Adjustments of Rao's score test for distributional and local parametric misspecifications, mimeo, University of Illinois at Urbana-Champaign.

Bera, A., Sosa-Escudero, W. and Yoon, M., 2001, Tests for the error component model in the presence of local misspecification, *Journal of Econometrics*, 101, 1-23.

Breusch, T.S. and A.R. Pagan, 1980, The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, 47, 239–253.

Gasparini, L., Marchionni, M. and Sosa-Escudero, W., 2001, *Distribucion del Ingreso en la Argentina: Perspectivas y Efectos sobre el Bienestar*, Fundacion Arcor-Triunfar, Cordoba.

Gustafsson, B. and Johansson, M., 1999, In search of smoking guns: what makes income inequality vary over time in different countries?, *American Sociological Review*, 64, 585-605.

Hemmerle, W.J. and Hartley, H., 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 component model with non-normal disturbances, *Review of Economic Studies*, 52, 681-690.

Sosa-Escudero, W. and Bera, A., 2001, Specification tests for linear panel data models, *Stata Technical Bulletin*, STB-61, 18-21.

SERIE DOCUMENTOS DE TRABAJO DEL CEDLAS

Todos los Documentos de Trabajo del CEDLAS están disponibles en formato electrónico en <*www.depeco.econo.unlp.edu.ar/cedlas*>.

- Nro. 65 (Marzo, 2008). Walter Sosa Escudero and Anil K. Bera. "Tests for Unbalanced Error Component Models Under Local Misspecication".

- Nro. 64 (Febrero, 2008). Luis Casanova. "Trampas de Pobreza en Argentina: Evidencia Empírica a Partir de un Pseudo Panel".

- Nro. 63 (Enero, 2008). Francisco Franchetti y Diego Battistón. "Inequality in Health Coverage, Empirical Analysis with Microdata for Argentina 2006".

- Nro. 62 (Diciembre, 2007). Adriana Conconi, Guillermo Cruces, Sergio Olivieri y Raúl Sánchez. "E pur si muove? Movilidad, Pobreza y Desigualdad en América Latina".

- Nro. 61 (Diciembre, 2007). Mariana Marchionni, Germán Bet y Ana Pacheco. "Empleo, Educación y Entorno Social de los Jóvenes: Una Nueva Fuente de Información".

- Nro. 60 (Noviembre, 2007). María Gabriela Farfán y María Florencia Ruiz Díaz. "Discriminación Salarial en la Argentina: Un Análisis Distributivo".

- Nro. 59 (Octubre, 2007). Leopoldo Tornarolli y Adriana Conconi. "Informalidad y Movilidad Laboral: Un Análisis Empírico para Argentina".

- Nro. 58 (Septiembre, 2007). Leopoldo Tornarolli. "Metodología para el Análisis de la Pobreza Rural".

- Nro. 57 (Agosto, 2007). Adriana Conconi y Andrés Ham. "Pobreza Multidimensional Relativa: Una Aplicación a la Argentina".

- Nro. 56 (Agosto, 2007). Martín Cicowiez, Luciano Di Gresia y Leonardo Gasparini. "Politicas Públicas y Objetivos de Desarrollo del Milenio en la Argentina".

- Nro. 55 (Julio, 2007). Leonardo Gasparini, Javier Alejo, Francisco Haimovich, Sergio Olivieri y Leopoldo Tornarolli. "Poverty among the Elderly in Latin America and the Caribbean".

- Nro. 54 (Julio, 2007). Gustavo Javier Canavire-Bacarreza y Luís Fernando Lima Soria. "Unemployment Duration and Labor Mobility in Argentina: A Socioeconomic-Based Pre- and Post-Crisis Analysis".

- Nro. 53 (Junio, 2007). Leonardo Gasparini, Francisco Haimovich y Sergio Olivieri. "Labor Informality Effects of a Poverty-Alleviation Program".

- Nro. 52 (Junio, 2007). Nicolás Epele y Victoria Dowbley. "Análisis Ex-Ante de un Aumento en la Dotación de Capital Humano: El Caso del Plan Familias de Transferencias Condicionadas".

- Nro. 51 (Mayo, 2007). Jerónimo Carballo y María Bongiorno. "Vulnerabilidad Individual: Evolución, Diferencias Regionales e Impacto de la Crisis. Argentina 1995 – 2005".

- Nro. 50 (Mayo, 2007). Paula Giovagnoli. "Failures in School Progression".

- Nro. 49 (Abril, 2007). Sebastian Galiani, Daniel Heymann, Carlos Dabús y Fernando Tohmé. "Land-Rich Economies, Education and Economic Development".

- Nro. 48 (Abril, 2007). Ricardo Bebczuk y Francisco Haimovich. "MDGs and Microcredit: An Empirical Evaluation for Latin American Countries".

- Nro. 47 (Marzo, 2007). Sebastian Galiani y Federico Weinschelbaum. "Modeling Informality Formally: Households and Firms".

- Nro. 46 (Febrero, 2007). Leonardo Gasparini y Leopoldo Tornarolli. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata".

- Nro. 45 (Enero, 2007). Georgina Pizzolitto. "Curvas de Engel de Alimentos, Preferencias Heterogéneas y Características Demográficas de los Hogares: Estimaciones para Argentina".

- Nro. 44 (Diciembre, 2006). Rafael Di Tella, Sebastian Galiani y Ernesto Schargrodsky. "Crime Distribution and Victim Behavior during a Crime Wave".

- Nro. 43 (Noviembre, 2006). Martín Cicowiez, Leonardo Gasparini, Federico Gutiérrez y Leopoldo Tornarolli. "Areas Rurales y Objetivos de Desarrollo del Milenio en America Latina y El Caribe".

- Nro. 42 (Octubre, 2006). Martín Guzmán y Ezequiel Molina. "Desigualdad e Instituciones en una Dimensión Intertemporal".

- Nro. 41 (Septiembre, 2006). Leonardo Gasparini y Ezequiel Molina. "Income Distribution, Institutions and Conflicts: An Exploratory Analysis for Latin America and the Caribbean".

- Nro. 40 (Agosto, 2006). Leonardo Lucchetti. "Caracterización de la Percepción del Bienestar y Cálculo de la Línea de Pobreza Subjetiva en Argentina".

- Nro. 39 (Julio, 2006). Héctor Zacaria y Juan Ignacio Zoloa. "Desigualdad y Pobreza entre las Regiones Argentinas: Un Análisis de Microdescomposiciones".

- Nro. 38 (Julio, 2006). Leonardo Gasparini, Matías Horenstein y Sergio Olivieri. "Economic Polarisation in Latin America and the Caribbean: What do Household Surveys Tell Us?".

- Nro. 37 (Junio, 2006). Walter Sosa-Escudero, Mariana Marchionni y Omar Arias. "Sources of Income Persistence: Evidence from Rural El Salvador".

- Nro. 36 (Mayo, 2006). Javier Alejo. "Desigualdad Salarial en el Gran Buenos Aires: Una Aplicación de Regresión por Cuantiles en Microdescomposiciones".

- Nro. 35 (Abril, 2006). Jerónimo Carballo y María Bongiorno. "La Evolución de la Pobreza en Argentina: Crónica, Transitoria, Diferencias Regionales y Determinantes (1995-2003)".

- Nro. 34 (Marzo, 2006). Francisco Haimovich, Hernán Winkler y Leonardo Gasparini. "Distribución del Ingreso en América Latina: Explorando las Diferencias entre Países".

- Nro. 33 (Febrero, 2006). Nicolás Parlamento y Ernesto Salinardi. "Explicando los Cambios en la Desigualdad: Son Estadísticamente Significativas las Microsimulaciones? Una Aplicación para el Gran Buenos Aires".

- Nro. 32 (Enero, 2006). Rodrigo González. "Distribución de la Prima Salarial del Sector Público en Argentina".

- Nro. 31 (Enero, 2006). Luis Casanova. "Análisis estático y dinámico de la pobreza en Argentina: Evidencia Empírica para el Periodo 1998-2002".

- Nro. 30 (Diciembre, 2005). Leonardo Gasparini, Federico Gutiérrez y Leopoldo Tornarolli. "Growth and Income Poverty in Latin America and the Caribbean: Evidence from Household Surveys".

- Nro. 29 (Noviembre, 2005). Mariana Marchionni. "Labor Participation and Earnings for Young Women in Argentina".

- Nro. 28 (Octubre, 2005). Martín Tetaz. "Educación y Mercado de Trabajo".

- Nro. 27 (Septiembre, 2005). Matías Busso, Martín Cicowiez y Leonardo Gasparini. "Ethnicity and the Millennium Development Goals in Latin America and the Caribbean".

- Nro. 26 (Agosto, 2005). Hernán Winkler. "Monitoring the Socio-Economic Conditions in Uruguay".

- Nro. 25 (Julio, 2005). Leonardo Gasparini, Federico Gutiérrez y Guido G. Porto. "Trade and Labor Outcomes in Latin America's Rural Areas: A Cross-Household Surveys Approach".

- Nro. 24 (Junio, 2005). Francisco Haimovich y Hernán Winkler. "Pobreza Rural y Urbana en Argentina: Un Análisis de Descomposiciones".

- Nro. 23 (Mayo, 2005). Leonardo Gasparini y Martín Cicowiez. "Equality of Opportunity and Optimal Cash and In-Kind Policies".

- Nro. 22 (Abril, 2005). Leonardo Gasparini y Santiago Pinto. "Equality of Opportunity and Optimal Cash and In-Kind Policies".

- Nro. 21 (Abril, 2005). Matías Busso, Federico Cerimedo y Martín Cicowiez. "Pobreza, Crecimiento y Desigualdad: Descifrando la Última Década en Argentina".

- Nro. 20 (Marzo, 2005). Georgina Pizzolitto. "Poverty and Inequality in Chile: Methodological Issues and a Literature Review".

- Nro. 19 (Marzo, 2005). Paula Giovagnoli, Georgina Pizzolitto y Julieta Trías. "Monitoring the Socio-Economic Conditions in Chile".

- Nro. 18 (Febrero, 2005). Leonardo Gasparini. "Assessing Benefit-Incidence Results Using Decompositions: The Case of Health Policy in Argentina".

- Nro. 17 (Enero, 2005). Leonardo Gasparini. "Protección Social y Empleo en América Latina: Estudio sobre la Base de Encuestas de Hogares".

- Nro. 16 (Diciembre, 2004). Evelyn Vezza. "Poder de Mercado en las Profesiones Autorreguladas: El Desempeño Médico en Argentina".

- Nro. 15 (Noviembre, 2004). Matías Horenstein y Sergio Olivieri. "Polarización del Ingreso en la Argentina: Teoría y Aplicación de la Polarización Pura del Ingreso".

- Nro. 14 (Octubre, 2004). Leonardo Gasparini y Walter Sosa Escudero. "Implicit Rents from Own-Housing and Income Distribution: Econometric Estimates for Greater Buenos Aires".

- Nro. 13 (Septiembre, 2004). Monserrat Bustelo. "Caracterización de los Cambios en la Desigualdad y la Pobreza en Argentina Haciendo Uso de Técnicas de Descomposiciones Microeconometricas (1992-2001)".

- Nro. 12 (Agosto, 2004). Leonardo Gasparini, Martín Cicowiez, Federico Gutiérrez y Mariana Marchionni. "Simulating Income Distribution Changes in Bolivia: a Microeconometric Approach".

- Nro. 11 (Julio, 2004). Federico H. Gutierrez. "Dinámica Salarial y Ocupacional: Análisis de Panel para Argentina 1998-2002".

- Nro. 10 (Junio, 2004). María Victoria Fazio. "Incidencia de las Horas Trabajadas en el Rendimiento Académico de Estudiantes Universitarios Argentinos".

- Nro. 9 (Mayo, 2004). Julieta Trías. "Determinantes de la Utilización de los Servicios de Salud: El Caso de los Niños en la Argentina".

- Nro. 8 (Abril, 2004). Federico Cerimedo. "Duración del Desempleo y Ciclo Económico en la Argentina".

- Nro. 7 (Marzo, 2004). Monserrat Bustelo y Leonardo Lucchetti. "La Pobreza en Argentina: Perfil, Evolución y Determinantes Profundos (1996, 1998 Y 2001)".

- Nro. 6 (Febrero, 2004). Hernán Winkler. "Estructura de Edades de la Fuerza Laboral y Distribución del Ingreso: Un Análisis Empírico para la Argentina".

- Nro. 5 (Enero, 2004). Pablo Acosta y Leonardo Gasparini. "Capital Accumulation, Trade Liberalization and Rising Wage Inequality: The Case of Argentina".

- Nro. 4 (Diciembre, 2003). Mariana Marchionni y Leonardo Gasparini. "Tracing Out the Effects of Demographic Changes on the Income Distribution. The Case of Greater Buenos Aires".

- Nro. 3 (Noviembre, 2003). Martín Cicowiez. "Comercio y Desigualdad Salarial en Argentina: Un Enfoque de Equilibrio General Computado".

- Nro. 2 (Octubre, 2003). Leonardo Gasparini. "Income Inequality in Latin America and the Caribbean: Evidence from Household Surveys".

- Nro. 1 (Septiembre, 2003). Leonardo Gasparini. "Argentina's Distributional Failure: The Role of Integration and Public Policies".