*Econometric Theory*, **15**, 1999, 814–823. Printed in the United States of America.

# UNEQUALLY SPACED PANEL DATA REGRESSIONS WITH AR(1) DISTURBANCES

Badi H. Baltagi
*Texas A&M University*

Ping X. Wu
*University of Melbourne*

This paper deals with the estimation of unequally spaced panel data regression models with AR(1) remainder disturbances. A feasible generalized least squares (GLS) procedure is proposed as a weighted least squares that can handle a wide range of unequally spaced panel data patterns. This procedure is simple to compute and provides natural estimates of the serial correlation and variance components parameters. The paper also provides a locally best invariant test for zero first-order serial correlation against positive or negative serial correlation in case of unequally spaced panel data.

## 1. INTRODUCTION

Some panel data sets cannot be collected every period as a result of lack of resources or cut in funding. Instead, these panels are collected over unequally spaced time intervals. For example, a panel of households could be collected over unequally spaced years rather than annually. This is also likely when collecting data on countries, states, or firms where in certain years, the data are not recorded, are hard to obtain, or are simply missing. Other common examples are panel data sets using daily data from the stock market, including stock prices, commodity prices, futures, etc. These panel data sets are unequally spaced when the market closes on weekends and holidays. The model considered in this paper allows for unequally spaced time-series data for each country, individual, or firm. This is particularly useful for housing resale data where the pattern of resales for each house occurs at different time periods and the panel is unbalanced because we observe different number of resales for each house.

Panel data with missing observations have been studied by Wansbeek and Kapteyn (1989) and Baltagi and Chang (1994). However, none of these studies consider the problem of serial correlation with unequally spaced panels. Estima-

We thank three anonymous referees and the editor for helpful comments and suggestions. We also thank Dong Li for his research assistance. An earlier version of this paper was given at the North American summer meetings of the Econometric Society held in Montreal, Canada, June 1998. Address correspondence to: Badi H. Baltagi, Department of Economics, Texas A&M University, College Station, TX 77843-4228, USA; e-mail: Badi@econ.tamu.edu.

  0266-4666/99 $9.50

tion of AR(1) disturbances in time-series regressions with missing observations has been studied by Wansbeek and Kapteyn (1985), whereas testing for AR(1) disturbances in this context has been considered by Shively (1993), Robinson (1985), Dufour and Dagenais (1985), and Savin and White (1978). This paper proposes a simple, feasible generalized least squares (GLS) estimation method for unbalanced panels that allows for a variety of patterns of missing data and serially correlated errors of the AR(1) type. In addition, this paper provides a locally best invariant (LBI) test for zero first-order serial correlation against positive or negative serial correlation. This extends the work of King (1985), Dufour and Dagenais (1985), and Shively (1993) to the context of an unequally spaced panel. In particular, we consider a random error component regression model with AR(1) disturbances (see Lillard and Willis, 1978; Bhargava, Franzini, and Narendranathan, 1982, Baltagi and Li, 1991). However, we allow for unequally spaced patterns for each individual in the time series dimension. Savin and White (1978) allowed for a gap of $m$ consecutive observations in time-series data. Here, we allow for a general type of unequally spaced panel data for each individual of the type considered by Shively (1993) in a time-series context.

## 2. THE MODEL

Consider the following unbalanced panel data regression model (see Wansbeek and Kapteyn, 1989):

$$y_{it} = x_{it}'\beta + u_{it} \qquad i = 1,\ldots,N; \qquad t = 1,\ldots,T_i, \tag{1}$$

where $\beta$ is a $K \times 1$ vector of regression coefficients including the intercept and $x_{it}$ is a $K \times 1$ vector of nonstochastic regressors. The disturbances follow a one-way error component model $u_{it} = \mu_i + \nu_{it}$ with individual effects $\mu_i \sim \text{IID}(0,\sigma_\mu^2)$ and the remainder disturbances $\nu_{it}$ following a stationary AR(1), i.e., $\nu_{it} = \rho\nu_{i,t-1} + \epsilon_{it}$ with $|\rho| < 1$ and $\epsilon_{it}$ is $\text{IID}(0,\sigma_\epsilon^2)$. The $\mu_i$'s are independent of the $\nu_{it}$'s, and $\nu_{i0} \sim (0,\sigma_\epsilon^2/(1-\rho^2))$. Each individual $i$ observes data at times $t_{i,j}$ for $j = 1,\ldots,n_i$ with $1 = t_{i,1} < \cdots < t_{i,n_i} = T_i$ with $n_i > K$ for $i = 1,2,\ldots,N$. For estimation of the equally spaced panel data regression model with AR(1) disturbances and no missing observations, see Baltagi and Li (1991). Note that the typical covariance element of $\nu_{it}$ for the observed periods $t_{i,j}$ and $t_{i,\ell}$ is given by $\text{cov}(\nu_{i,t_{i,j}},\nu_{i,t_{i,\ell}}) = \sigma_\epsilon^2\rho^{|t_{i,j}-t_{i,\ell}|}/(1-\rho^2)$ for $\ell, j = 1,\ldots,n_i$. In fact, by continuous substitution over the AR(1) process, one can show that

$$\nu_{i,t_{i,j}} = \rho^{t_{i,j}-t_{i,j-1}}\nu_{i,t_{i,j-1}} + \epsilon_{i,t_{i,j}} + \rho\epsilon_{i,t_{i,j}-1} + \cdots + \rho^{t_{i,j}-t_{i,j-1}-1}\epsilon_{i,t_{i,j-1}+1}. \tag{2}$$

Define $S_{i,t_{i,j}} = \nu_{i,t_{i,j}} - \rho^{t_{i,j}-t_{i,j-1}}\nu_{i,t_{i,j-1}} = \epsilon_{i,t_{i,j}} + \rho\epsilon_{i,t_{i,j}-1} + \cdots + \rho^{t_{i,j}-t_{i,j-1}-1} \times \epsilon_{i,t_{i,j-1}+1}$ for $j = 2,\ldots,n_i$, and let $S_{i,t_{i,1}} = (1-\rho^2)^{1/2}\nu_{i,t_{i,1}}$. For equally spaced data, with no missing observations, $S_{i,t_{i,j}}$ is equivalent to the Prais–Winsten transformation for the AR(1) model. In fact, for $\nu_i' = (\nu_{i,t_{i,1}},\ldots,\nu_{i,t_{i,n_i}})$

and $S_i' = (S_{i,t_{i,1}}, \ldots, S_{i,t_{i,n_i}})$, this transformation can be written in matrix form as $S_i(\rho) = C_i(\rho)\nu_i$ where

$$C_i(\rho) = \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 & \ldots & 0 & 0 & 0 \\ -\rho^{t_{i,2}-t_{i,1}} & 1 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\rho^{t_{i,n_i-1}-t_{i,n_i-2}} & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & -\rho^{t_{i,n_i}-t_{i,n_i-1}} & 1 \end{bmatrix} \tag{3}$$

However, for the unequally spaced data the transformed disturbances are still heteroskedastic. In fact, $\text{var}(S_{i,t_{i,1}}) = \sigma_\epsilon^2$ whereas $\text{var}(S_{i,t_{i,j}}) = \sigma_\epsilon^2(1 + \rho^2 + \cdots + \rho^{2(t_{i,j}-t_{i,j-1}-1)}) = \sigma_\epsilon^2(1 - \rho^{2(t_{i,j}-t_{i,j-1})})/(1-\rho^2)$ for $j = 2,\ldots,n_i$. Note that for the no missing observations case with $t_{i,j} - t_{i,j-1} = 1$ for all $j = 1,2,\ldots,n_i$, $\text{var}(S_{i,t_{i,j}}) = \sigma_\epsilon^2$ for all $j$. In general, the vector of disturbances $S_i$ can be easily made homoskedastic by premultiplying $S_i$ by a diagonal matrix

$$D_i(\rho) = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \left(\frac{1-\rho^2}{1-\rho^{2(t_{i,2}-t_{i,1})}}\right)^{1/2} & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & \left(\frac{1-\rho^2}{1-\rho^{2(t_{i,n_i}-t_{i,n_i-1})}}\right)^{1/2} \end{bmatrix}. \tag{4}$$

Hence, if we let $C_i^*(\rho) = D_i(\rho)C_i(\rho)$, and define $\nu_i^* = C_i^*(\rho)\nu_i$ then $\nu_i^* \sim (0, \sigma_\epsilon^2 I_{n_i})$. Note that for equally spaced data with no missing observations, the diagonal matrix $D_i(\rho)$ reverts back to the identity matrix as expected. The $n_i \times n_i$ matrix $C_i^*(\rho) = D_i(\rho)C_i(\rho)$ is given by[1]

$$C_i^*(\rho) = (1-\rho^2)^{1/2} \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ -\left(\frac{\rho^{2(t_{i,2}-t_{i,1})}}{1-\rho^{2(t_{i,2}-t_{i,1})}}\right)^{1/2} & \left(\frac{1}{1-\rho^{2(t_{i,2}-t_{i,1})}}\right)^{1/2} & \ldots & \vdots & \vdots \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & -\left(\frac{\rho^{2(t_{i,n_i}-t_{i,n_i-1})}}{1-\rho^{2(t_{i,n_i}-t_{i,n_i-1})}}\right)^{1/2} & \left(\frac{1}{1-\rho^{2(t_{i,n_i}-t_{i,n_i-1})}}\right)^{1/2} \end{bmatrix}. \tag{5}$$

Premultiply the panel data regression given in equation (1) by $\text{diag}[C_i^*(\rho)]$, which is a block-diagonal matrix with $C_i^*(\rho)$ in the $i$th block. This transforms the disturbances as follows:

$$u^* = \text{diag}[C_i^*(\rho)]u = \text{diag}[C_i^*(\rho)]\,\text{diag}(\iota_{n_i})\mu + \text{diag}[C_i^*(\rho)]\nu, \tag{6}$$

where $\iota_{n_i}$ is a vector of ones of dimension $n_i$, $u' = (u'_1,\ldots,u'_N)$, $\mu' = (\mu_1,\ldots,\mu_N)$, and $\nu' = (\nu'_1,\ldots,\nu'_N)$. Note that $\nu'_i$ was defined following (2) and $u'_i$ is similarly defined. It can be easily verified that

$$g_i = [C_i^*(\rho)]\iota_{n_i} = (1-\rho^2)^{1/2}\left(1, \frac{1-\rho^{(t_{i,2}-t_{i,1})}}{(1-\rho^{2(t_{i,2}-t_{i,1})})^{1/2}}, \ldots, \frac{1-\rho^{(t_{i,n_i}-t_{i,n_i-1})}}{(1-\rho^{2(t_{i,n_i}-t_{i,n_i-1})})^{1/2}}\right)'. \quad (7)$$

Therefore, $u^* = \mathrm{diag}(g_i)\mu + \mathrm{diag}[C_i^*(\rho)]\nu$, and $\Omega^* = E(u^*u^{*\prime}) = \sigma_\mu^2\,\mathrm{diag}(g_i g_i') + \sigma_\epsilon^2\,\mathrm{diag}(I_{n_i})$, because $\mathrm{diag}[C_i^*(\rho)]E(\nu\nu')\mathrm{diag}[C_i^*(\rho)]' = \sigma_\epsilon^2\,\mathrm{diag}(I_{n_i})$. This can be rewritten as

$$\Omega^* = \sigma_\mu^2\,\mathrm{diag}(g_i' g_i)\mathrm{diag}(P_{g_i}) + \sigma_\epsilon^2\,\mathrm{diag}(P_{g_i} + Q_{g_i}),$$

where $P_{g_i} = g_i(g_i'g_i)^{-1}g_i'$, $Q_{g_i} = I_{n_i} - P_{g_i}$, $g_i'g_i = \sum_{j=1}^{n_i} g_{i,j}^2 = (1-\rho^2) + (1-\rho^2)\times\sum_{j=2}^{n_i}(1-\rho^{(t_{i,j}-t_{i,j-1})})^2/(1-\rho^{2(t_{i,j}-t_{i,j-1})})$, and $g_{i,j}$ is the $j$th element of $g_i = [C_i^*(\rho)]\iota_{n_i}$ given in (7) for $j = 1,2,\ldots,n_i$. This replaces $I_{n_i}$ by $(Q_{g_i} + P_{g_i})$ (see Wansbeek and Kapteyn, 1982). Collecting terms with the same matrices, one obtains

$$\Omega^* = \mathrm{diag}(\omega_i^2)\mathrm{diag}(P_{g_i}) + \sigma_\epsilon^2\,\mathrm{diag}(Q_{g_i}), \quad (8)$$

where $\omega_i^2 = g_i'g_i\sigma_\mu^2 + \sigma_\epsilon^2$. Note that $P_{g_i}$ and $Q_{g_i}$ are idempotent, are orthogonal to each other, and sum to the identity matrix. Therefore,

$$\sigma_\epsilon\Omega^{*-1/2} = \sigma_\epsilon\,\mathrm{diag}(1/\omega_i)\mathrm{diag}(P_{g_i}) + \mathrm{diag}(Q_{g_i}) = \mathrm{diag}(I_{n_i}) - \mathrm{diag}(\theta_i P_{g_i}), \quad (9)$$

where $\theta_i = 1 - (\sigma_\epsilon/\omega_i)$. For equally spaced panel data with no missing observations and no serial correlation, i.e., $\rho = 0$, this reduces to the familiar Fuller and Battese (1974) transformation. Premultiplying $y^* = \mathrm{diag}[C_i^*(\rho)]y$ by $\sigma_\epsilon\Omega^{*-1/2}$ one gets $y^{**} = \sigma_\epsilon\Omega^{*-1/2}y^*$. The typical elements of $y^{**}$ are given by

$$y_{i,t_{i,j}}^{**} = y_{i,t_{i,j}}^{*} - \theta_i g_{i,j}\left(\sum_{s=1}^{n_i} g_{i,s}y_{i,t_{i,s}}^{*}\right)\Bigg/\left(\sum_{s=1}^{n_i} g_{i,s}^2\right). \quad (10)$$

Quadratic unbiased estimators of the variance components arise naturally from (8) (for the balanced panel data case, see Baltagi and Li, 1991). In fact, $\mathrm{diag}(P_{g_i})u^* \sim (0, \mathrm{diag}(\omega_i^2 P_{g_i}))$ and $\mathrm{diag}(Q_{g_i})u^* \sim (0, \sigma_\epsilon^2\,\mathrm{diag}(Q_{g_i}))$. Therefore, quadratic unbiased estimators of $\sigma_\omega^2 = \sum_{i=1}^N \omega_i^2$ and $\sigma_\epsilon^2$ are given by

$$\hat\sigma_\omega^2 = u^{*\prime}\,\mathrm{diag}(P_{g_i})u^* \quad \text{and} \quad \hat\sigma_\epsilon^2 = u^{*\prime}\,\mathrm{diag}(Q_{g_i})u^*\Bigg/\sum_{i=1}^N (n_i - 1), \quad (11)$$

where $\mathrm{tr}(P_{g_i}) = 1$ and $\mathrm{tr}(Q_{g_i}) = n_i - 1$. Using $\sum_{i=1}^N \omega_i^2 = \sum_{i=1}^N g_i'g_i\sigma_\mu^2 + N\sigma_\epsilon^2$, an estimator for $\sigma_\mu^2$ can be obtained as follows:

$$\hat\sigma_\mu^2 = (u^{*\prime}\,\mathrm{diag}(P_{g_i})u^* - N\hat\sigma_\epsilon^2)\Bigg/\sum_{i=1}^N g_i'g_i. \quad (12)$$

Using a consistent estimate of $\rho$, one can estimate this model by feasible GLS using the following steps.

Step 1. Perform the $\mathrm{diag}[C_i^*(\rho)]$ transformation given in (5) on equation (1) to get rid of serial correlation and ensure homoskedasticity. This yields $y_i^* = [C_i^*(\rho)]y_i$, where $y_i' = (y_{it_{i,1}}, \ldots, y_{it_{i,n_i}})$.

Step 2. Compute the residuals $\hat{u}^*$ obtained from the ordinary least squares (OLS) regression of $y^*$ on $X^*$ in step 1. Using these $\hat{u}^*$'s, compute estimates of $\sigma_\mu^2$ and $\sigma_\epsilon^2$ from (11) and (12). Deduce $\hat{\theta}_i = 1 - (\hat{\sigma}_\epsilon / \hat{\omega}_i)$, where $\hat{\omega}_i^2 = g_i' g_i \hat{\sigma}_\mu^2 + \hat{\sigma}_\epsilon^2$.

Step 3. Obtain $y^{**}$ as described in (10) using the $y^*$'s in step 1 and $\hat{\theta}_i$ from step 2. Perform OLS of $y^{**}$ on $X^{**}$ to obtain the feasible GLS estimate of $\beta$. In the next section, we consider the problem of testing for zero first-order serial correlation and in the process provide a natural estimator for $\rho$.

## 3. LOCALLY BEST INVARIANT (LBI) TEST

In this section, we derive a locally best invariant (LBI) test for $H_0$; $\rho = 0$ versus $H_a^+$; $\rho > 0$ or $H_a^-$; $\rho < 0$, for the unequally spaced panel data regression model described in Section 2. We assume normality of the disturbances and rewrite (1) in matrix form as

$$y = X\beta + \mathrm{diag}(\iota_{n_i})\mu + \nu, \tag{13}$$

where $\iota_{n_i}$, $\mu$, and $\nu$ have been defined following (6). Also $\Sigma_\nu(\rho) = E(\nu\nu') = \sigma_\epsilon^2 \,\mathrm{diag}(V_i) = \sigma_\epsilon^2 \Omega_\nu(\rho)$, where $\sigma_\epsilon^2 V_i$ has a typical element given preceding (2). Consider the orthogonal matrix $O_{n_i}$ of dimension $n_i$ and let $O_{n_i} = (\iota_{n_i}/\sqrt{n_i}, B_i)$, where $B_i$ is an $n_i \times (n_i - 1)$ matrix that by definition satisfies the following properties: $B_i'\iota_{n_i} = 0$, $B_i' B_i = I_{n_i - 1}$, and $B_i B_i' = I_{n_i} - \bar{J}_{n_i}$, where $\bar{J}_{n_i} = J_{n_i}/n_i$ and $J_{n_i}$ is a matrix of ones of dimension $n_i$. Premultiplying the preceding model by $\mathrm{diag}(B_i')$ yields $\mathrm{diag}(B_i')y = \mathrm{diag}(B_i')X\beta + \mathrm{diag}(B_i')\nu$. Note that $\mu$ is swept away because $B_i'\iota_{n_i} = 0$. If we let $\tilde{y} = \mathrm{diag}(B_i')y$, $\widetilde{X} = \mathrm{diag}(B_i')X$, and $\tilde{\nu} = \mathrm{diag}(B_i')\nu$, then this transformed model can be rewritten as

$$\tilde{y} = \widetilde{X}\beta + \tilde{\nu}, \tag{14}$$

where $\Sigma_{\tilde{\nu}}(\rho) = E(\tilde{\nu}\tilde{\nu}') = \sigma_\epsilon^2 \,\mathrm{diag}(B_i' V_i B_i) = \sigma_\epsilon^2 \Omega_{\tilde{\nu}}(\rho)$. Under $H_0$; $\rho = 0$, we have $V_i = I_{n_i}$, and $\Sigma_{\tilde{\nu}}(\rho)$ reduces to $\sigma_\epsilon^2 \,\mathrm{diag}(I_{n_i - 1})$. This means that under $H_0$, $\tilde{\nu} \sim N(0, \sigma_\epsilon^2 I_{\Sigma n_i - N})$.

This testing problem is invariant to transformations of the form $y \to \gamma_0 y + X\gamma$, where $\gamma_0$ is a positive scalar and $\gamma$ is $K \times 1$. This means that if we change the scale of $y$ and add a known linear combination of the regressors to the rescaled $y$, this does not change the truth of either $H_0$ or $H_a$. This is the transformation used by Durbin and Watson (1971) to establish optimal properties of the Durbin–Watson test (see Dufour and King, 1991). Let $m = \sum_{i=1}^N (n_i - 1) - K$, $\bar{P}_{\widetilde{X}} = I_{\Sigma(n_i - 1)} - P_{\widetilde{X}}$, where $P_{\widetilde{X}} = \widetilde{X}(\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'$, and $z = \bar{P}_{\widetilde{X}}\tilde{y}$ be the OLS residuals from $\tilde{y}$ on $\widetilde{X}$. Let $R$ be the $m \times \Sigma(n_i - 1)$ matrix such that $RR' = I_m$ and $R'R = \bar{P}_{\widetilde{X}}$. Note that the $m \times 1$ vector $Rz$ is a linear unbiased residual vector with a scalar covariance matrix (see Theil, 1971, Ch. 5) with zero mean and variance-covariance matrix $\sigma_\epsilon^2 I_m$ under $H_0$. As noted by King and Hillier (1985), the vector $s = Rz/(z'R'Rz)^{1/2}$ is a maximal invariant under the preceding group of transformations for our testing problem. Also, the LBI test, which is also known as the locally most powerful invariant test, is of the form $d = z'A_0 z/z'z < c_\alpha$, where $A_0 = \{\partial\Omega_{\tilde{\nu}}^{-1}(\rho)/\partial\rho\}|_{\rho=0} = -(\partial\Omega_{\tilde{\nu}}(\rho)/\partial\rho)|_{\rho=0}$ and $c_\alpha$ is the $\alpha$-level critical value of $d$. This test has the steepest sloping power curve at $H_0$; $\rho = 0$ within the class of invariant tests of the same significance level. It can also be viewed as the test that has optimal power within the neighborhood of $H_0$. King and Hillier (1985) also demonstrated the equivalence of this test to a one-sided version of the Lagrange multiplier test. Note that the denominator of the test statistic $d$ is easy to compute. In fact, $z'z = \tilde{y}'\bar{P}_{\tilde{X}}\tilde{y} = \tilde{y}'\tilde{y} - \tilde{y}'P_{\tilde{X}}\tilde{y} = y'Qy - y'QX(X'QX)^{-1}X'Qy$, where $Q = \operatorname{diag}(I_{n_i} - \bar{J}_{n_i})$. This is the within residuals sum of squares obtained from the fixed effects model, i.e., assuming the $\mu_i$'s are fixed parameters to be estimated. It can be obtained from the OLS residuals sum of squares of (13) or in case $N$ is large from the regression of $(y_{i,t_{i,j}} - \bar{y}_{i.})$ on $(X_{i,t_{i,j},k} - \bar{X}_{i.,k})$ for $k = 2, \ldots, K$. In this case, $\bar{y}_{i.} = \sum_{j=1}^{n_i} y_{i,t_{i,j}}/n_i$ and $\bar{X}_{i.,k}$ is defined similarly.

For the numerator of $d$, we note that $\partial V_i/\partial\rho|_{\rho=0} = V_i^0$, where the typical elements of $V_i^0$ are

$$V_i^0(j,\ell) = \begin{cases} 1 & \text{if } |t_{i,j} - t_{i,\ell}| = 1 \\ 0 & \text{if } |t_{i,j} - t_{i,\ell}| > 1 \end{cases} \quad \text{or} \quad \text{if } j = \ell \quad \text{for } j,\ell = 1,\ldots,n_i, \tag{15}$$

so that $z'A_0 z = -z' \operatorname{diag}(B_i' V_i^0 B_i)z$. Let $\tilde{z} = \operatorname{diag}(B_i)z$; then $z'A_0 z = \tilde{z}' \operatorname{diag}(V_i^0)\tilde{z}$. It can be easily shown that OLS on (14) yields $\tilde{\beta} = (X'QX)^{-1}X'Qy$, the within estimator of $\beta$, and $z = \tilde{y} - \tilde{X}\tilde{\beta} = \operatorname{diag}(B_i')(y - X\tilde{\beta})$. This means that $\tilde{z} = \operatorname{diag}(B_i B_i')(y - X\tilde{\beta}) = Qy - QX\tilde{\beta}$, which is exactly the within residuals described previously using the deviation from individual means regression. This also proves that $z'z = \tilde{z}'\tilde{z}$. Therefore, the LBI test statistic can be written as

$$d^* = d + 2 = \tilde{z}' \operatorname{diag}(2I_{n_i} - V_i^0)\tilde{z}/\tilde{z}'\tilde{z}, \tag{16}$$

and our test statistic can be expressed as the sum of four terms:

$$d_1 = \sum_{i=1}^{N}\sum_{j=1}^{n_i} [\tilde{z}_{i,t_{i,j}} - \tilde{z}_{i,t_{i,j-1}} I(t_{i,j} - t_{i,j-1} = 1)]^2 \bigg/ \sum_{i=1}^{N}\sum_{j=1}^{n_i} \tilde{z}_{i,t_{i,j}}^2,$$

$$d_2 = \sum_{i=1}^{N}\sum_{j=1}^{n_i-1} \tilde{z}_{i,t_{i,j}}^2 [1 - I(t_{i,j+1} - t_{i,j} = 1)] \bigg/ \sum_{i=1}^{N}\sum_{j=1}^{n_i} \tilde{z}_{i,t_{i,j}}^2,$$

$$d_3 = \sum_{i=1}^{N} \tilde{z}_{i,t_{i,1}}^2 \bigg/ \sum_{i=1}^{N}\sum_{j=1}^{n_i} \tilde{z}_{i,t_{i,j}}^2 \quad \text{and} \quad d_4 = \sum_{i=1}^{N} \tilde{z}_{i,t_{i,n_i}}^2 \bigg/ \sum_{i=1}^{N}\sum_{j=1}^{n_i} \tilde{z}_{i,t_{i,j}}^2,$$

where $I(t_{i,j} - t_{i,j-1} = 1)$ is the indicator function that takes the value 1 when the proposition in parentheses is true and zero elsewhere. In the time-series context, $d^*$ reduces to the LBI test suggested by Dufour and Dagenais (1985). Some comments are in order: (i) When there are no missing observations, i.e., $t_{i,n_i} - t_{i,n_i-1} = t_{i,n_i-1} - t_{i,n_i-2} = \cdots = t_{i,2} - t_{i,1} = 1$, then $d_2 = 0$ and our test statistic reduces to $d_1 + d_3 + d_4$. Note that in this case, $d_1$ reduces to the Durbin–Watson test statistic proposed by Bhargava et al. (1982, p. 535) for the fixed effect AR(1) model. Bhargava et al. claimed that $d_1$ is a locally most powerful invariant test in the neighborhood of $\rho = 0$. They argued that exact critical values for $d_1$ are both impractical and unnecessary because they involve the computation of nonzero eigenvalues of a large $NT \times NT$ matrix. Instead, the 5% upper and lower bounds for $d_1$ were tabulated for various values of $N$, $T$, and $K$. Note that the Bhargava et al. test statistic ignores $d_3$ and $d_4$ and is therefore approximately locally best invariant. This is the same criticism for the Durbin–Watson statistic in the time-series context pointed out by King (1981). Whereas terms such as $d_3$ and $d_4$ are negligible when $T$ is very large, they are less likely to be so in typical panels where $T$ is usually small. (ii) In practice, one can insert zeros in between $\tilde{z}_{i,t_{i,j}}$ and $\tilde{z}_{i,t_{i,j+1}}$ if $t_{i,j+1} - t_{i,j} > 1$. Then one gets a new residuals series that may look like $e_i' = (\tilde{z}_{i,t_{i,1}}, 0, 0, \tilde{z}_{i,t_{i,2}}, 0, \ldots, \tilde{z}_{i,t_{i,j}}, \ldots, \tilde{z}_{i,t_{i,n_i}})$ if $t_{i,1} = 1$, $t_{i,2} = 4, \ldots,$ and $t_{i,n_i} = T$ (see Dufour and Dagenais, 1985, p. 375; Shively, 1993, p. 248). The term $e_i$ is a $T \times 1$ vector of within residuals for the $i$th individual whose element is zero if the data for that period are not available. Now it is easy to compute the test statistic in the usual way, $d_c = \sum_{i=1}^{N} \sum_{t=2}^{T} (e_{i,t} - e_{i,t-1})^2 / \sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t}^2$. This also suggests a natural estimator of $\rho$, i.e., $\hat{\rho} = (\sum_{i=1}^{N} \sum_{t=2}^{T} e_{i,t} e_{i,t-1} / m_c) / (\sum_{i=1}^{N} \sum_{t=1}^{T} e_{i,t}^2 / n)$, where $m_c$ is the number of consecutive pairs of nonzero $e_{it}$'s. (iii) Note that our test statistic $d^*$ in (16) can be rewritten as $d^* = 2 - \tilde{z}' \operatorname{diag}(V_i^0) \tilde{z} / \tilde{z}'\tilde{z} = 2 - z' \operatorname{diag}(B_i' V_i^0 B_i) z / z'z$ where $z = \bar{P}_{\tilde{X}} \tilde{y} = \bar{P}_{\tilde{X}} \tilde{\nu}$ from (14) with $\tilde{\nu} \sim N(0, \sigma_\epsilon^2 I_{\Sigma(n_i-1)})$ under $H_0$; $\rho = 0$. Substituting this in the test statistic, we get $d^* = 2 - (\tilde{\nu}' \bar{P}_{\tilde{X}} A \bar{P}_{\tilde{X}} \tilde{\nu} / \tilde{\nu}' \bar{P}_{\tilde{X}} \tilde{\nu})$, where $A = \operatorname{diag}(B_i' V_i^0 B_i)$. This statistic is in the form of a ratio of quadratic forms of normal variates $(\tilde{\nu})$. For a given $A$, bounds for critical values can be constructed just like the Bhargava et al. or Durbin–Watson test statistics. However, different patterns of missing observations imply a different $V_i^0$ matrix and therefore a different $A$ matrix.[2] A reasonable and practical way to approximate these critical values is to standardize this test statistic. Using the results in Evans and King (1985), we compute $E(d^*) = 2 - \operatorname{tr}(\bar{P}_{\tilde{X}} A)/m$ and $\operatorname{var}(d^*) = 2[m \operatorname{tr}(\bar{P}_{\tilde{X}} A)^2 - \{\operatorname{tr}(\bar{P}_{\tilde{X}} A)\}^2] / m^2(m+2)$, where $m = \sum_{i=1}^{N} (n_i - 1) - K$, and then use the test statistic

$$d_s^* = [d^* - E(d^*)] \Big/ \sqrt{\operatorname{var}(d^*)}. \tag{17}$$

For testing $H_0$; $\rho = 0$ versus $H_a$; $\rho > 0$ (or $\rho < 0$), one compares $d^*$ to the critical value from the lower (upper) tail of an $N(0,1)$ distribution. Although $B_i$ appears in the preceding standardized test statistic, this matrix is not needed for the actual computations. In fact,

$$\begin{aligned} \operatorname{tr}(\bar{P}_{\tilde{X}} A) &= \operatorname{tr}[\bar{P}_{\tilde{X}} \operatorname{diag}(B_i') \operatorname{diag}(V_i^0) \operatorname{diag}(B_i)] \\ &= \operatorname{tr}\{[Q - QX(X'QX)^{-1}X'Q] \operatorname{diag}(V_i^0)\}. \end{aligned}$$

A weakness of the preceding LBI test is that for unequally spaced observations that are two or more periods apart, the LBI test statistic in (16) degenerates to 2 because $\partial V_i'/\partial\rho$ evaluated at the maximum likelihood estimator (MLE) under $H_0$ is equal to zero. Even in situations that are not as extreme as this, if a big portion of the sample has observations that are two or more periods apart, then the LBI test will still be ineffective because it ignores these observations in the computations. Clearly, this difficulty can be removed if we do not evaluate $\partial V_i/\partial\rho$ at the null hypothesis in the test construction. A test that is particularly suitable for this situation is the POI test suggested by King (1985). In fact, King's point optimal invariant test for the AR(1) case has been generalized to the missing data case in time-series regressions by Shively (1993). This POI test can be easily generalized for the unequally spaced panel data model and is available upon request from the authors. Also, our proposed estimator of $\rho$ breaks down if there are no consecutive observations. In this case, one may want to use maximum likelihood methods that do not throw away nonconsecutive observations (for MLE in the serially correlated time-series model with missing observations, see Wansbeek and Kapteyn, 1985).

## 4. EMPIRICAL ILLUSTRATION

Table 1 applies the LBI test statistic given by $d^*$ in (16) and the Bhargava et al. modified Durbin–Watson test statistic given by $d_1$ following (16) to the Grunfeld data on investment. This is the same data set used for time-series illustration by Wansbeek and Kapteyn (1985). It consists of 10 large U.S. manufacturing firms observed from 1935 to 1954. Real gross investment of firm $i$ in year $t$ $(I_{it})$ is regressed on the real value of the firm $(F_{it})$ and the real value of the capital stock $(C_{it})$:

$$I_{it} = \alpha + \beta_1 F_{it} + \beta_2 C_{it} + u_{it}. \qquad \textbf{(18)}$$

The assumption of normality of the disturbances and nonstochastic regressors may be untenable for this empirical example, but we use these data for illustrative purposes. The null hypothesis is $H_0; \rho = 0$ versus $H_a^+; \rho > 0$, and various patterns of missing observations are considered. As clear from the table, the LBI statistic is greater than the Bhargava et al. statistic for all patterns considered. The difference between the LBI and BFN test statistics highlights the contributions of $d_2$, $d_3$, and $d_4$ defined following (16) to the LBI statistic. This difference is substantial depending on the pattern of missing observations. Table 1 also gives the standardized LBI statistic given by $d_s^*$ defined in (17). The null hypothesis of no serial correlation is rejected in all cases.

## 5. SUMMARY AND CONCLUSION

This paper derives a simple method of correcting for serial correlation and random individual effects in the context of unequally spaced panel data models. It

**TABLE 1.** Testing for Zero Serial Correlation in Unequally Spaced Panels: Grunfeld Data[a]

| Pattern | Missing periods | | | | | | | | LBI | Bhargava et al. | $d_s^*$ | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 9 | 10 | | | | | | | 1.022 | 0.706 | −7.870 | 180 |
| B | 17 | 18 | | | | | | | 1.139 | 0.807 | −6.994 | 180 |
| C | 3 | 4 | 5 | | | | | | 1.162 | 0.738 | −6.751 | 170 |
| D | 7 | 8 | 9 | | | | | | 1.013 | 0.701 | −7.796 | 170 |
| E | 13 | 14 | 15 | | | | | | 0.982 | 0.674 | −7.986 | 170 |
| F | 3 | 4 | 5 | 6 | | | | | 1.188 | 0.733 | −6.455 | 160 |
| G | 12 | 13 | 14 | 15 | | | | | 0.920 | 0.612 | −8.254 | 160 |
| H | 2 | 4 | 5 | 14 | | | | | 1.237 | 0.694 | −6.493 | 160 |
| I | 8 | 9 | 16 | 17 | 19 | | | | 1.499 | 0.968 | −4.447 | 150 |
| J | 2 | 3 | 15 | 16 | 17 | 19 | | | 1.580 | 0.911 | −3.842 | 140 |
| K | 2 | 3 | 15 | 18 | 19 | 20 | | | 1.174 | 0.813 | −6.471 | 140 |
| L | 2 | 3 | 5 | 7 | 15 | 20 | | | 1.330 | 0.689 | −5.899 | 140 |
| M | 3 | 5 | 8 | 9 | 16 | 17 | 19 | | 1.807 | 1.031 | −2.290 | 130 |
| N | 2 | 4 | 5 | 14 | 15 | 16 | 19 | | 1.641 | 0.901 | −3.459 | 130 |
| O | 2 | 3 | 4 | 8 | 9 | 16 | 17 | 19 | 1.709 | 1.005 | −2.998 | 120 |
| P | 2 | 3 | 5 | 7 | 15 | 18 | 19 | 20 | 1.589 | 0.866 | −3.881 | 120 |
| Q | 2 | 4 | 5 | 8 | 14 | 15 | 16 | 19 | 1.656 | 0.873 | −3.430 | 120 |

[a]This tests $H_0$; $\rho = 0$ versus $H_a$; $\rho > 0$. The LBI statistic is given in (16). The Bhargava et al. statistic is given by $d_1$ following (16). The term $d_s^*$ is given in (17). Here $n$ denotes the total number of observations in the panel.

also provides an LBI test for zero first-order serial correlation against positive or negative serial correlation. Although exact critical values can be computationally prohibitive with panel data, advances in computing exact critical values following Shively et al. (1990) can be used. Alternatively, one can approximate these critical values by standardizing the LBI statistic using the results of Evans and King (1985) on quadratic normals. Finally, the LBI test statistic is illustrated for various missing observations panels using the Grunfeld investment data. Future work should generalize the results to allow for stochastic regressors.

## NOTES

1. An alternative derivation of this matrix in a time-series context is given by Wansbeek and Kapteyn (1985).
2. To take advantage of the power optimality of the LBI test ($d^*$), one should compute the exact critical values. Unfortunately, this needs the nonzero eigenvalues of $AP_{\bar{X}}$, which is a large $\Sigma(n_i - 1) \times \Sigma(n_i - 1)$ matrix. This requires $O([\Sigma(n_i - 1)]^3)$ operations. Here, advances in computing exact critical values by Shively, Ansley, and Kohn (1990) can be used. Alternatively, one can easily compute bounds for the exact critical values as described in the Appendix that is available upon request from the authors.

## *REFERENCES*


Baltagi, B.H. & Y.J. Chang (1994) Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model. *Journal of Econometrics* 62, 67–89.

Baltagi, B.H. & Q. Li (1991) A transformation that will circumvent the problem of autocorrelation in an error component model. *Journal of Econometrics* 48, 385–393.

Bhargava, A., L. Franzini, & W. Narendranathan (1982) Serial correlation and the fixed effects model. *Review of Economic Studies* 49, 533–549.

Dufour, J.M. & M.G. Dagenais (1985) Durbin-Watson tests for serial correlation in regressions with missing observations. *Journal of Econometrics* 27, 371–381.

Dufour, J.M. & M. King (1991) Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary and nonstationary AR(1) errors. *Journal of Econometrics* 47, 115–143.

Durbin, J. & G.S. Watson (1971) Testing for serial correlation in least squares regression III. *Biometrika* 58, 1–19.

Evans, M.A. & M.L. King (1985) Critical value approximations for tests of linear regression disturbances. *Australian Journal of Statistics* 27, 68–83.

Fuller, W.A. & G.E. Battese (1974) Estimation of linear models with cross-error structure. *Journal of Econometrics* 2, 67–78.

King, M.L. (1981) The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic. *Journal of Econometrics* 17, 51–66.

King, M.L. (1985) A point optimal test for autoregressive disturbances. *Journal of Econometrics* 27, 21–37.

King, M.L. & G.H. Hillier (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society, Series B* 47, 98–102.

Lillard, L.A. & R.J. Willis (1978) Dynamic aspects of earning mobility. *Econometrica* 46, 985–1012.

Robinson, P.M. (1985) Testing for serial correlation in regression with missing observations. *Journal of the Royal Statistical Society, Series B* 47, 429–437.

Savin, N.E. & K.J. White (1978) Testing for autocorrelation with missing observations. *Econometrica* 46, 59–67.

Shively, T.S. (1993) Testing for autoregressive disturbances in a time series regression with missing observations. *Journal of Econometrics* 57, 233–255.

Shively, T.S., C.F. Ansley, & R. Kohn (1990) Fast evaluation of the distribution of the Durbin-Watson and other invariant test statistics in time series regression. *Journal of the American Statistical Association* 85, 676–685.

Theil, H. (1971) *Principles of Econometrics.* New York: Wiley.

Wansbeek, T. & A. Kapteyn (1982) A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics A* 11, 2105–2112.

Wansbeek, T. & A. Kapteyn (1985) Estimation in a linear model with serially correlated errors when observations are missing. *International Economic Review* 26, 469–490.

Wansbeek, T. & A. Kapteyn (1989) Estimation of the error components model with incomplete panels. *Journal of Econometrics* 41, 341–361.