A SCORE TEST FOR INDIVIDUAL HETEROSCEDASTICITY IN A ONE-WAY ERROR COMPONENTS MODEL

by Alberto Holly and Lucien Gardiol

Cahier No. 9915

Version date : September 17, 1999

Preliminary version. Please do not quote without the authors' explicit consent. Comments most welcome.

authors' address

Département d'économétrie
et d'économie politique
Université de Lausanne
Ecole des HEC
1015-Lausanne
Switzerland

Abstract

The purpose of this paper is to derive a Rao's e¢cient score statistic for testing for heteroscedasticity in an error components model with only individual e¤ects. We assume that the individual e¤ect exists and therefore do not test for it. In addition, we assume that the individual e¤ects, and not the white noise term may be heteroscedastic. Finally, we assume that the error components are normally distributed.

We …rst establish, under a speci…c set of assumptions, the asymptotic distribution of the Score under contiguous alternatives. We then derive the expression for the Score test statistic for individual heteroscedasticity. Finally, we discuss the asymptotic local power of this Score test statistic.

Key words :Panel data, Error components model, Score test, Individual heteroscedasticity, Contiguous alternatives, Asymptotic local power.

JEL classi…cation : C23, C12

1 Introduction

In the analysis of error-components models it is custumary to assume that the individual e¤ects are homoscedastic. In some situations, however, it may be appropriate to generalize the error components model context to the heteroscedastic case, as …rst suggested by Mazodier and Trognon (1978). Misspeci…cation errors in presence of heteroscedasticity can produce misleading results. However, if no heteroscedasticity exists, standard estimation and speci…cation test procedures can be applied straightforwardly. It would therefore simplify the analysis considerably if one were to test for heteroscedasticity before implementing more elaborate inference procedures to deal with the possible heteroscedasticity situation.

To this testing purpose, a natural procedure consists in using Rao's e¢-cient score statistic [Rao (1948)], or it's Lagrange Multipier (LM) interpretation provided by Silvey (1959), as its computation is based on the usual error components model in the homoscedastic case. In another setting, Breusch and Pagan (1980) have considered the standard linear regression model with non–spherical disturbances and took the error-components model of Balestra and Nerlove (1966) as an example. They presented an LM test for the null hypothesis that the individual e¤ect is missing. Gourieroux, Holly and Monfort (1982) derived the asymptotic distribution of the LM test of Breusch and Pagan (1980) by taking into account the fact that the parameter de…ning the null hypothesis is on the boundary of the parameter set. They showed that the standard asymptotic distribution theory does not apply in this case and derived the appropiate nonstandard results.[1]

In a recent paper, Lejeune (1998) developed a pseudo-LM test procedure for jointly testing the null hypothesis of no individual e¤ects and homoscedasticity against the alternative of random individual e¤ects and heteroscedas-

[1] See also Baltagi, Chang and Li (1992) for an analysis of the behavior of one–sided LM tests.

ticity in the white noise error term.[2] The pseudo-LM test derived by Lejeune (1998) is distribution-free in the sense that it does not rely on any distribution assumption such as normality. In this paper we consider a di¤erent setting than in Lejeune (1998). Firstly, we assume that the individual e¤ect exists and therefore do not test for it. Secondly, we assume that the individual e¤ects, and not the white noise term may be heteroscedastic. Thirdly, we assume that the error components are normally distributed. In addition, not only the speci…cation considered in this paper di¤ers from that of Lejeune (1998) but also the method of derivation of the main results, which we believe to be useful in other contexts as well.

The paper is organized as follows. The speci…cation of the model as well as some preliminary assumptions are presented in Section 2. The derivation of the asymptotic distribution of the Score under contiguous alternatives is contained in Section 3. The expression of the heteroscedaticity test statistic is derived in Section 4 and its asymptotic local power is discussed in Section 5.

Throughout this paper, we tried to adhere to widely accepted set of notation in the context of Panel Data models. In particular, the unit vector (all elements = 1) of size T £ 1 is denoted by $¶_T$ and the unit matrix (all elements = 1) of size T £ T is denoted by J_T (= $¶_T ¶^0_T$). For a review of the main matrices used in this paper as well as their properties, see Crépon and Mairesse [(1996), Appendix].[3]

The notation D and AD are used throughout to mean the distribution and asymptotic distribution, respectively, of a random variable or a random vector. The noncentral chi–square distribution with p degrees of freedom and noncentrality parameter $±^2$ is de…ned as the distribution of the scalar product of a random p–variate normal vector with covariance equal to the identity

[2] We would like to thank B. Lejeune for making his unpublihed manuscripts available to us.

[3] Appendix based on an unpublished manuscript by Alain Trognon (1984).

matrix and mean vector having a norm of δ, and is denoted by $\chi^2_p(\delta^2)$.

2 Specification of the model and preliminary assumptions

We consider the one-way error components linear regression model

$$y_{nt} = x_{nt}'\beta^0 + u_{nt}^0 \quad n = 1, \ldots, N;\ t = 1, \ldots, T$$

where y_{nt} is the (scalar) observation of the dependent variable, x_{nt} a $K \times 1$ vector of nonstochastic explanatory variables, and u_{nt}^0 the unobservable error term which is decomposed as

$$u_{nt}^0 = \mu_n^0 + v_{nt}^0,$$

where μ_n^0 is the unobservable random variable of individual effects and v_{nt}^0 the usual unobservable error term.

We assume that

Assumption 1 μ_n^0 and v_{nt}^0 are independent for all n and t; the v_{nt}^0 are independent identically distributed as $N(0, \sigma_v^{02})$ and the μ_n^0 are independent and distributed as $N\left(0, \sigma_\mu^{02} h\left(z_n'\theta^0\right)\right)$ where z_n is a $p \times 1$ vector of explanatory variables such that $z_n'\theta^0$ does not contain a constant term; $h : \mathbb{R} \to \mathbb{R}$ is a strictly positive twice differentiable function satisfying $h(0) = 1$, $h^{(s)}(0) \neq 0$ for $s = 1, 2$ where $h^{(s)}$ denotes the derivative of order s of h.

Let

$$\delta^0 = (\beta^{0\prime}, \sigma_v^{02}, \sigma_\mu^{02})'$$

Assumption 2 $\delta^0 \in \Delta$ where Δ is a compact subset of $\mathbb{R}^K \times \mathbb{R}_+ \times \mathbb{R}_+$ and $\theta^0 \in \Theta$ where Θ is a compact subset of $\mathbb{R}^p$.

Assumption 3 (δ', θ') is an interior point of $\Delta \times \Theta$.

It is important to observe that we assume that $\sigma_1'^2$ is strictly positive - in other words, that $\sigma_1'^2$ is not on the boundary of the parameter set Θ. Therefore, we shall not question the existence of individual effects. Instead, we shall test for heteroscedasticity of the individual effects by testing H_0 : $\theta' = 0$ against $H_a : \theta' \neq 0$.

The T observations for individual n can be expressed in the following matrix form:

$$y_n = X_n \beta' + u_n',$$

where y_n is the $T \times 1$ vector of the y_{nt}, X_n is the $T \times K$ matrix whose n-th row is x_{nt}' and u_n' is the $T \times 1$ vector of the u_{nt}'.

In this paper we deal with the so-called semi-asymptotic case where T is fixed and N goes to infinity.

Assumption 4 The empiric distribution of (X_n, z_n), denoted by F_n, converges completely to a nondegenerate distribution function $F(X, z)$. The marginal (limiting) distribution of z will be denoted by F_z.

More assumptions will be introduced in the following section.

Stacking the individuals one after the other, we have:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{pmatrix} \beta' + \begin{pmatrix} u_1' \\ \vdots \\ u_n' \\ \vdots \\ u_N' \end{pmatrix}$$

or more compactly:

$$y = X\beta^{0} + u^{0}$$

where

$$D(u) = N\left(0; \Omega^{0}\right)$$

and

$$\Omega^{0} = \sigma_v^{02} I_{NT} + \sigma_1^{02}\,\mathrm{diag}\left(h\left(z_n'\theta^{0}\right)\right) \otimes J_T$$

which may be conveniently written as

$$\Omega^{0} = \sigma_v^{02} W_n + \mathrm{diag}[\sigma_v^{02} + T\sigma_1^{02} h\left(z_n'\theta^{0}\right)] \otimes \frac{J_T}{T}$$

where

$$W_n = I_N \otimes \left(I_T - \frac{J_T}{T}\right)$$

3 Asymptotic distribution of the Score under contiguous alternatives

The log–likelihood function is:

$$L = \text{constant} - \frac{1}{2}\ln\det(\Omega) - \frac{1}{2}u'\Omega^{-1}u$$

where

$$\Omega^{-1} = \frac{1}{\sigma_v^2}W_n + \mathrm{diag}\left(\frac{1}{\sigma_v^2 + T\sigma_1^2 h\left(z_n'\theta\right)}\right) \otimes \frac{J_T}{T} \qquad (1)$$

Let

$$\gamma = \left(\beta', \sigma_v^2, \sigma_1^2, \theta'\right)' \tag{2}$$

and consider contiguous alternatives of the form :

$$\gamma_N^a = \gamma^0 + N^{-1/2}\gamma^a \tag{3}$$

where, since we consider contiguous alternatives only for the heteroscedasticity coefficients,

$$\gamma^a = \left(0', \; 0, \; 0, \; \theta^{a\prime}\right)' \tag{4}$$

The purpose of this section is to show that under specific regularity assumptions, $N^{-1/2}\partial L(\gamma_N^a)/\partial\gamma$ is asymptotically normaly distributed. This is the key result for the asymptotic distribution of the heteroscedasticity test to be derived in the following section.

The first differential of L is:

$$dL = -\frac{1}{2}\mathrm{tr}\left(\Sigma^{-1} d\Sigma\right) - u'\Sigma^{-1}du + \frac{1}{2}u'\Sigma^{-1}d\Sigma\,\Sigma^{-1}u \tag{5}$$

where

$$\begin{aligned} du &= -X\,d\beta \\ d\Sigma &= I_{NT}\,d\sigma_v^2 + T\left(\mathrm{diag}\left(h(z_n'\theta)\right) \otimes \frac{J_T}{T}\right) d\sigma_1^2 \\ &\quad + T\sigma_1^2\left(\mathrm{diag}\left(h'(z_n'\theta)\, z_n'\, d\theta\right) \otimes \frac{J_T}{T}\right) \end{aligned}$$

By using the properties of the matrices W_n and J_T, it is not difficult to show that dL may be written as

$$dL = \frac{\partial L}{\partial \beta}(\gamma)d\beta + \frac{\partial L}{\partial \sigma_v^2}(\gamma)d\sigma_v^2 + \frac{\partial L}{\partial \sigma_1^2}(\gamma)d\sigma_1^2 + \frac{\partial L}{\partial \theta}(\gamma)d\theta$$

where

$$\frac{\partial L}{\partial \beta}(\gamma) = X'\Omega^{-1}u; \tag{6}$$

$$\begin{aligned}\frac{\partial L}{\partial \sigma_v^2}(\gamma) = \frac{1}{2}\Bigg\{ & -\frac{N(T-1)}{\sigma_v^2} - \sum_{n=1}^{N}\left(\frac{1}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]}\right) + \frac{1}{\sigma_v^4}u'W_n u \\ & + u'\left[\operatorname{diag}\left(\frac{1}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2}\right) \otimes \frac{J_T}{T}\right]u\Bigg\};\end{aligned} \tag{7}$$

$$\begin{aligned}\frac{\partial L}{\partial \sigma_1^2}(\gamma) = \frac{T}{2}\Bigg\{ & -\sum_{n=1}^{N}\left(\frac{h(z_n'\theta)}{\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)}\right) \\ & + u'\left[\operatorname{diag}\left(\frac{h(z_n'\theta)}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2}\right) \otimes \frac{J_T}{T}\right]u\Bigg\}\end{aligned} \tag{8}$$

$$\begin{aligned}\frac{\partial L}{\partial \theta}(\gamma) = \frac{1}{2}\sum_{n=1}^{N}\Bigg\{ & -\frac{T\sigma_1^2 h^{(1)}(z_n'\theta)\, z_n}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]} + \\ & + \frac{1}{2}u'\left[\operatorname{diag}\left(\frac{T\sigma_1^2 h^{(1)}(z_n'\theta)\, z_n}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2}\right) \otimes \frac{J_T}{T}\right]u\Bigg\}\end{aligned} \tag{9}$$

In order to derive the asymptotic distribution of $N^{-1/2}\partial L(\gamma_N^a)/\partial\gamma$, it is necessary to evaluate the second differential of L.

Using the fact that the second differential of u is equal to zero, the second differential of L is equal to:

$$
\begin{aligned}
d^2 L = {} & \frac{1}{2}\operatorname{tr}\left[\Omega^{-1} d\Omega\, \Omega^{-1} d\Omega\right] - \frac{1}{2}\operatorname{tr}\left[\Omega^{-1} d^2 \Omega\right] \\
& - u'\Omega^{-1} d\Omega\, \Omega^{-1} d\Omega\, \Omega^{-1} u + \frac{1}{2} u'\Omega^{-1} d^2\Omega\, \Omega^{-1} u \\
& + 2u'\Omega^{-1} d\Omega\, \Omega^{-1} du - du'\Omega^{-1} du
\end{aligned}
\qquad (10)
$$

By taking the expectation of $d^2 L$, we obtain after obvious simplification, that:

$$
E\left[-d^2 L\right] = \frac{1}{2}\operatorname{tr}\left[\Omega^{-1} d\Omega\, \Omega^{-1} d\Omega\right] + d\beta' X'\Omega^{-1} X d\beta \qquad (11)
$$

It is not difficult to show that

$$
\begin{aligned}
\operatorname{tr}\left[\Omega^{-1} d\Omega\, \Omega^{-1} d\Omega\right] = {} & d\sigma_v^2 \left(\frac{N(T-1)}{\sigma_v^4} + \sum_{n=1}^{N} \frac{1}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} \right) d\sigma_v^2 \\
& + d\sigma_v^2 \left(2T \sum_{n=1}^{N} \frac{h(z_n'\theta)}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} \right) d\sigma_1^2 \\
& + d\sigma_v^2 \left(2T\sigma_1^2 \sum_{n=1}^{N} \frac{h^{(1)}(z_n'\theta)}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} z_n' \right) d\theta \\
& + d\sigma_1^2 \left(T^2 \sum_{n=1}^{N} \frac{[h(z_n'\theta)]^2}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} \right) d\sigma_1^2 \\
& + d\sigma_1^2 \left(2T^2\sigma_1^2 \sum_{n=1}^{N} \frac{h(z_n'\theta)\, h^{(1)}(z_n'\theta)}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} z_n' \right) d\theta
\end{aligned}
$$

$$+ d\theta' \left(T^2\sigma_1^4 \sum_{n=1} \frac{\left[h^{(1)}(z_n'\theta)\right]^2}{\left[\sigma_v^2 + T\sigma_1^2 h(z_n'\theta)\right]^2} z_n z_n' \right) d\theta$$

We shall first prove the following result:

Lemma 1 The loglikelihood L is regular with respect to its first and second derivatives, i.e.

$$E\left[-d^2 L\right] = E(dL)^2$$

Proof. [4]

From (??) we have

$$\begin{aligned}(dL)^2 = {} & \frac{1}{4}\left[u'\Omega^{-1}d\Omega\,\Omega^{-1}u - E(u'\Omega^{-1}d\Omega\,\Omega^{-1}u)\right]^2 + du'\Omega^{-1}uu'\Omega^{-1}du \\ & - \left[u'\Omega^{-1}d\Omega\,\Omega^{-1}u - E(u'\Omega^{-1}d\Omega\,\Omega^{-1}u)\right]u'\Omega^{-1}du\end{aligned}$$

The expectation of the third term of the right-hand side is equal to zero. Thus,

$$\begin{aligned}E(dL)^2 &= \frac{1}{4}V(u'\Omega^{-1}d\Omega\,\Omega^{-1}u) + du'\Omega^{-1}du \\ &= \frac{1}{2}\mathrm{tr}\left[\Omega^{-1}d\Omega\,\Omega^{-1}d\Omega\right] + du'\Omega^{-1}du \\ &= E\left[-d^2L\right]\end{aligned}$$

as stated. ■

Let z_j be the $j-$th component of z, $j = 1, \ldots, p$. We introduce the following additional assumptions.

Assumption 5 $\int \left|\frac{h^{(1)}(z'\theta)}{h(z'\theta)} z_j\right| dF_z(z) < \infty$ for every $j = 1, \ldots, p$.

[4] This result holds in a more general situation than the specific one considered in this paper. The proof we provide is simpler than the corresponding proof in Magnus (1978).

Assumption 6 $\int \left|\frac{h^{(2)}(z'\theta)}{h(z'\theta)} z_j z_k\right| dF_z(z) < \infty$ for every $j, k = 1, \ldots, p$.

Assumption 7 $\int \frac{h^{(1)2}(z'\theta)}{h^2(z'\theta)} |z_j z_k| \, dF_z(z) < \infty$ for every $j, k = 1, \ldots, p$.

Proposition 1 Let Assumptions 1 through ?? hold. Then
a) $E[-(1/N)\partial^2 L(\gamma)/\partial\gamma\partial\gamma']$ converges uniformly on Γ to the asymptotic information matrix $I(\gamma)$;
b) $-(1/N)\partial^2 L(\gamma)/\partial\gamma\partial\gamma'$ converges almost surely and uniformly on Γ to $I(\gamma)$.

Proof. Let $e^0 = \Omega^{0\,-1/2}u^0$ and suppose that Assumptions 1 through ?? hold. Then, by inspection, one may easily verify that all the elements of $-(1/N)\partial^2 L(\gamma)/\partial\gamma\partial\gamma'$ which, to save space, are not reproduced here, are of the form $(1/N)\sum_{n=1}^{N} f(X_n; z_n; e^0_n; \gamma)$ where the functions $f(X; z; e^0; \gamma)$ are either uniformly bounded or dominated by a function independent of γ which is integrable with respect to the product measure

$$\nu(A) = \int\int \mathbb{1}_A(X; z; e^0)\, dF(X; z)\, d\Phi(e^0)$$

where $\Phi(e^0)$ is the $N(0; 1)$ distribution, and $\mathbb{1}_A(X; z; e^0) = 1$ if $(X; z; e^0) \in A$, 0 otherwise.

The assertion of the proposition follows from the version of the Uniform strong law of large numbers proved in Gallant [(1987), Theorem 1, p. 159–162]. ■

Note that the asymptotic information matrix $I(\gamma)$ is of the form

$$I(\gamma) = \begin{pmatrix} I_{\beta\beta}(\gamma) & 0 & 0 & 0 \\ 0 & I_{\sigma_v^2\sigma_v^2}(\gamma) & I_{\sigma_v^2\sigma_1^2}(\gamma) & I_{\sigma_v^2\theta}(\gamma) \\ 0 & I_{\sigma_1^2\sigma_v^2}(\gamma) & I_{\sigma_1^2\sigma_1^2}(\gamma) & I_{\sigma_1^2\theta}(\gamma) \\ 0 & I_{\theta\sigma_v^2}(\gamma) & I_{\theta\sigma_1^2}(\gamma) & I_{\theta\theta}(\gamma) \end{pmatrix}$$

where, when evaluated at γ^0,

$$I_{\beta\beta}(\gamma^0) = \lim_{N\to\infty} \frac{1}{N} X' \Omega^{0\,-1} X;$$

$$I_{\sigma_\nu^2\sigma_\nu^2}(\gamma^0) = \frac{1}{2}\left[\frac{(T-1)}{\sigma_\nu^{04}} + \frac{1}{\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}\right];$$

$$I_{\sigma_\nu^2\sigma_1^2}(\gamma^0) = I_{\sigma_1^2\sigma_\nu^2}(\gamma^0) = \frac{T}{2\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}$$

$$I_{\sigma_\nu^2\theta}(\gamma^0) = I'_{\theta\sigma_\nu^2}(\gamma^0) = \frac{T\sigma_1^{02}h^{(1)}(0)}{2\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}\lim_{N\to\infty}\frac{1}{N}\iota_N' Z$$

$$I_{\sigma_1^2\sigma_1^2}(\gamma^0) = \frac{T^2}{2\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}$$

$$I_{\sigma_1^2\theta}(\gamma^0) = I'_{\theta\sigma_1^2}(\gamma^0) = \frac{T^2\sigma_1^{02}h^{(1)}(0)}{2\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}\lim_{N\to\infty}\frac{1}{N}\iota_N' Z$$

$$I_{\theta\theta}(\gamma^0) = \frac{T^2\sigma_1^{04}\left[h^{(1)}(0)\right]^2}{2\left(\sigma_\nu^{02} + T\sigma_1^{02}\right)^2}\lim_{N\to\infty}\frac{1}{N}Z'Z$$

Let

$$\underline{Z} = \left(I_N - \frac{J_N}{N}\right)Z \qquad (12)$$

be the matrix obtained by centering each column of Z.

We introduce the following additional assumptions:

Assumption 8 $\lim_{N\to\infty}\frac{1}{N}X'\Omega^{0\,-1}X$ is nonsingular

Assumption 9 $\lim_{N\to\infty}\frac{1}{N}\underline{Z}'\underline{Z}$ is nonsingular

Lemma 2 Under Assumptions ?? and ??, $I(\gamma^0)$ is nonsingular

Proof. We may write $I(\gamma^0)$ as

$$I(\gamma^0) = \begin{pmatrix} I_{\delta\delta}(\gamma^0) & I_{\delta\theta}(\gamma^0) \\ I_{\theta\delta}(\gamma^0) & I_{\theta\theta}(\gamma^0) \end{pmatrix}$$

It is easy to verify that if Assumption ?? is satisfied, then $I_{\delta\delta}(\gamma^0)$ is nonsingular. Therefore, $I(\gamma^0)$ is nonsingular if and only if $I_{\delta\delta}(\gamma^0) - I_{\delta\theta}(\gamma^0) I_{\theta\theta}(\gamma^0)^{-1} \times I_{\theta\delta}(\gamma^0)$ is nonsingular. In turn, this property is implied by Assumption ?? ■

Proposition 2 Under Assumptions 1 through ??

$$AD[N^{-1/2}\partial L(\gamma_N^a)/\partial\gamma] = N(0; I(\gamma^0))$$

Proof. If Assumptions 1 through ?? hold, then one can verify that, without additional assumptions, the Central limit theorem for contiguous alternatives proved in Gallant and Holly (1980) applies. Hence, $N^{-1/2}\partial L(\gamma_N^a)/\partial\gamma$ converges in distribution to the stated normal distribution. ■

4 The heteroscedasticity Score test statistic

The necessary first–order conditions system for the maximization of the log–likelihood function subject to the constraint $\theta = 0$ boils down to the familiar estimating equation for the homoscedastic one–way error components model; that is:

$$\hat\beta^{(c)} = \left(X'\hat\Omega^{(c)-1}X\right)^{-1}\left(X'\hat\Omega^{(c)-1}y\right)$$

$$\hat\sigma_v^{2(c)} = \frac{\hat u^{(c)\prime} W_n \hat u^{(c)}}{N(T-1)}$$

$$\hat\sigma_\mu^{2(c)} = \frac{\hat u^{(c)\prime} \overline{B}_n \hat u^{(c)}}{N(T-1)} - \frac{\hat u^{(c)\prime}\hat u^{(c)}}{NT(T-1)}$$

where

$$\hat{u}^{(c)} = y - X\hat{\beta}^{(c)}$$

$$\overline{B}_n = I_N - \frac{J_T}{T}$$

It is useful to note that

$$\hat{\sigma}_\nu^{2(c)} + T\hat{\sigma}_\mu^{2(c)} = \frac{\hat{u}^{(c)\prime}\overline{B}_n\hat{u}^{(c)}}{N} \tag{13}$$

Let

$$\hat{\theta}^{(c)} = \left(\hat{\beta}^{(c)}; \hat{\sigma}_\nu^{2(c)}; \hat{\sigma}_\mu^{2(c)}; 0\right)'$$

All the components of the score vector $\partial L(\gamma)/\partial\gamma$ evaluated at the constrained estimator $\hat{\theta}^{(c)}$ are equal to zero, except $\partial L(\hat{\theta}^{(c)})/\partial\theta$ which is equal to:

$$\frac{\partial L}{\partial\theta}(\hat{\theta}^{(c)}) = \frac{T}{2}\frac{\hat{\sigma}_\mu^{2(c)}h^{(1)}(0)}{\left(\hat{\sigma}_\nu^{2(c)} + T\hat{\sigma}_\mu^{2(c)}\right)^2} \times \left[\sum_{n=1}\left(u_n^{(c)\prime}\frac{J_T}{T}u_n^{(c)}\right)z_n - \left(\hat{\sigma}_\nu^{2(c)} + T\hat{\sigma}_\mu^{2(c)}\right)\sum_{n=1}z_n\right] \tag{14}$$

It is convenient to write $\partial L(\hat{\theta}^{(c)})/\partial\gamma$ more compactly in matrix notation. To this purpose, let $\hat{e}^{(c)}$ be the $N \times 1$ vector of the $\hat{e}_n^{(c)}$ where $\hat{e}_n^{(c)} = \hat{u}_n^{(c)\prime}(J_T/T)\hat{u}_n^{(c)}$.

Using (??), it is easy to verify that $\hat{\sigma}_\nu^{2(c)} + T\hat{\sigma}_\mu^{2(c)}$ is the mean of the $\hat{e}_n^{(c)}$. We may thus write,

$$\sum_{n=1}\left[\hat{u}_n^{(c)\prime}\frac{J_T}{T}\hat{u}_n^{(c)}\left(z_n - \frac{1}{N}\sum_{n=1}z_n\right)\right] = \underline{Z}'\hat{e}^{(c)}$$

We may thus write $\partial L(\hat{\theta}^{(c)})/\partial\gamma$ more compactly as:

$$\frac{\partial L}{\partial \gamma}(\hat{\gamma}^{(c)}) = \left(0; 0; 0; \frac{\partial L}{\partial \theta}(\hat{\gamma}^{(c)})\right)' \tag{15}$$

where

$$\frac{\partial L}{\partial \theta}(\hat{\gamma}^{(c)}) = \frac{1}{2} \frac{T\hat{\sigma}_1^{2(c)} h^{(1)}(0)}{\left(\hat{\sigma}_v^{2(c)} + T\hat{\sigma}_1^{2(c)}\right)^2} \underline{Z}' \hat{u}^{(c)} \tag{16}$$

As usual, the information matrix evaluated at γ is defined as

$$I_N(\gamma) = -E\left\{\frac{\partial^2 L(\gamma)}{\partial \gamma \partial \gamma'}\right\}$$

We may write $I_N^{-1}(\gamma)$ as

$$I_N^{-1}(\gamma) = \begin{pmatrix} I_N^{\delta\delta}(\gamma) & I_N^{\delta\theta}(\gamma) \\ I_N^{\theta\delta}(\gamma) & I_N^{\theta\theta}(\gamma) \end{pmatrix} \tag{17}$$

By using (??) and (??), we easily verify that

$$I_N^{\theta\theta}(\hat{\gamma}^{(c)}) = \frac{2\left(\hat{\sigma}_v^{2(c)} + T\hat{\sigma}_1^{2(c)}\right)^2}{T^2 \hat{\sigma}_1^{4(c)} [h^{(1)}(0)]^2} (\underline{Z}'\underline{Z})^{-1} \tag{18}$$

We are now in position to derive the expression for the Score test statistic ξ^S given by:

$$\xi^S = \frac{\partial L}{\partial \gamma}\left(\hat{\gamma}^{(c)}\right)' I_N^{-1}(\hat{\gamma}^{(c)}) \frac{\partial L}{\partial \gamma}\left(\hat{\gamma}^{(c)}\right) \tag{19}$$

Straightforward computation shows, by using (??), (??) and (??), that

$$\xi^S = \frac{1}{2(\hat{\sigma}_v^{2(c)} + T\hat{\sigma}_1^{2(c)})^2} \hat{u}^{(c)\prime} \underline{Z} (\underline{Z}'\underline{Z})^{-1} \underline{Z}' \hat{u}^{(c)} \tag{20}$$

Alternatively, by using the fact that $\hat{\sigma}_v^{2(c)} + T\hat{\sigma}_1^{2(c)}$ is the mean of the $\hat{e}_n^{(c)}$, we may write the expression for the Score test statistic ξ^S as follows,

$$\xi^S = \frac{1}{2}\left(\frac{\hat{e}^{(c)}}{\bar{s}} - \iota_N\right)' \underline{Z}\,(\underline{Z}'\underline{Z})^{-1}\,\underline{Z}'\left(\frac{\hat{e}^{(c)}}{\bar{s}} - \iota_N\right) \tag{21}$$

where $\bar{s}$ is the mean of $\hat{e}^{(c)}$.

The Score test statistic ξ^S is thus one half of the explained sum of squares of the OLS regression of $\hat{e}^{(c)}/\bar{s} - 1$ against $\underline{Z}$ as in Breusch and Pagan (1979).[5]

5 Asymptotic local power

Since, according to Proposition ??, $AD[N^{-1/2}\partial L(\gamma_N^a)/\partial\gamma] = N(0; I(\gamma^0))$, one can show that, under contiguous alternatives, the distribution of the Score test statistic ξ^S converges to the noncentral chi–square distribution with p degrees of freedom and noncentrality parameter $\gamma^{a\prime} A \gamma^a$, that is,

$$AD\left(\xi^S\right) = \chi_p^2\left(\gamma^{a\prime} A \gamma^a\right)$$

where

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I^{\mu\mu}(\gamma^0) \end{pmatrix}$$

For a proof see, for example, Holly (1987).

The asymptotic power of the test is given by the noncentrality parameter $\gamma^{a\prime} A \gamma^a$. Its expression is given by:

[5] Notice also the difference and similarity with the particular expression of the Pseudo-LM test in Lejeune (1998) when the normality assumption is assumed to hold.

$$\gamma^{a\prime} A \gamma^{a} = \lim_{N \to \infty} \frac{1}{2} \frac{T^2 \sigma_1^4 \left(h^{(1)}(0)\right)^2}{\left(\sigma_v^2 + T\sigma_1^2\right)^2} \theta^{a\prime} (\underline{Z}'\underline{Z}/N)\theta^{a}$$

The asymptotic power is influenced by three factors. Firstly, not surprisingly, the power is influenced by θ^a, the test is more powerful to detect alternatives which are away from the null hypothesis: Secondly, although the Score test statistic itself does not depend on $h(0)$ or $h^{(1)}(0)$, the asymptotic power is an increasing function of $\left(h^{(1)}(0)\right)^2$ for any given alternative. Thirdly, the power increases with T, as the multiplicative constant $T\sigma_1^2 / \left(\sigma_v^2 + T\sigma_1^2\right)$ converges to 1 when T goes to infinity. This last effect shows that the test is improved when the number of observations for each individual sample increases. One could also note that the local power of the test tends to zero when σ_1^2 tends to zero and will tend to be small if $T\sigma_1^2$ is small compared to σ_v^2. Thus, as one should expect, the test will be powerful in situations where the individual heteroscedasticity is high.

References

Balestra, P. and Nerlove, M. (1966), "Pooling cross-section and time series data in the estimation of a dynamic model," Econometrica, 34, pp 585–612.

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, pp 95–120.

Breusch, T. S. and Pagan, A. R. (1979), "A simple test of heteroscedasticity and random coe¢cient variation," Econometrica, 47, pp 1287– 1294.

Breusch, T. S. and Pagan, A. R. (1980), "The Lagrange multiplier test and its applications to model speci…cation in econometrics," Review of Economic Studies, 47, pp 239–253.

Crépon, B. and Mairesse J. (1996), "The Chamberlain approach," in The econometrics of panel data. A handbook of the theory with applications, L. Mátyás and P. Sevestre eds The Netherlands: Kluwer.

Gallant, A. R. (1987), Nonlinear statistical models, New York: Wiley.

Gallant, A. R. and Holly A. (1980), "Statistical inference in an implicit, nonlinear, simultaneous equation model in the context of maximum likelihood estimation," Econometrica, 48, pp 697–720.

Gourieroux, Ch., Holly A. and Monfort A. (1982), "Kuhn-Tucker, Likelihood Ratio and Wald tests for nonlinear models with inequality constraints on the parameters," Discussion Paper Series , 770, Harvard University.

Holly, A. (1987), "Speci…cation tests: an overview," in Advances in econometrics –Fifth World Congress, Volume I, T. F. Bewley ed. New York: Cambridge University Press.

Lejeune, B. (1998), "A distribution-free joint test for random individual effects and heteroscedasticity allowing for incomplete panels," Mimeo.

Magnus, J. R. (1978), "Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix," Journal of Econometrics, 7, pp 281–312. Corrigenda, Journal of Econometrics, 10, 261.

Mazodier, P. and Trognon, A. (1978), "Heteroscedasticity and strati…cation in error components models," Annales de l'INSEE, 30–31, pp 451– 509.

Rao, C. R. (1948), "Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation," Proceedings of the Cambridge Philosophical Society, 44, pp 50–57.

Silvey, S. D. (1959), "The Lagrangian multiplier test," The Annals of Mathematical Statistics , 30, 389-407.

Trognon, A. (1984), Econométrie II, Lecture notes, INSEE–ENSAE