Is Adaptive Estimation Useful for Panel Models with Heteroskedasticity in the Unit-Specific Error Component? Some Monte Carlo Evidence

Nilanjana Roy
Department of Economics, University of Victoria,
British Columbia, Canada, V8W 2Y2. Email: nroy@uvic.ca

December, 1999

Econometrics Working Paper EWP9913
ISSN 1485-6441

Abstract

This paper first derives an adaptive estimator when heteroskedasticity is present in the unit-specific error in an error component model and then compares the finite sample performance of the proposed estimator with various other estimators. While the Monte Carlo results show that the proposed estimator performs adequately in terms of relative efficiency, its performance on the basis of empirical size is quite similar to the other estimators considered. The results from using the different estimators in two applications highlight the importance of devising a test in future to distinguish between the source of heteroskedasticity.

Keywords: Heteroskedasticity, Kernel Estimation, Error Component Model.
JEL Classification: C14 and C23.

1 Introduction:

One of the common problems encountered in the econometrics literature is the estimation of linear regression models with heteroskedastic error of unknown functional form. There is an extensive literature on this issue in the context of cross-sectional and time-series data. It is widely recognized that we seldom know the form of the heteroskedasticity and that the application of estimated generalized least squares (EGLS) in the face of such misspecification would lead to inefficient estimators which in turn can lead to erroneous inferences. One approach to deal with this problem is to apply EGLS in the models using nonparametric methods since such estimators are robust to misspecification of the functional form; see Carroll (1982), Delgado (1992), Hidalgo (1992), Robinson (1987), among others. Even though the resulting EGLS estimators are asymptotically efficient, we need to know their finite sample performance as the data we encounter are always finite. Rilstone (1991) addresses one aspect of this question by undertaking a Monte Carlo study to compare the nonparametric EGLS estimators with various parametric estimators using both correct and incorrect forms for the heteroskedasticity. One of his results is that, apart from correctly specified EGLS, the semiparametric approach generally dominates other estimators in larger samples[1] with models exhibiting moderate and large amounts of heteroskedasticity.

Surprisingly, the issue of heteroskedasticity in the case of panel data has not been studied as extensively. In a recent survey, Baltagi (1998) mentions only four references (Mazodier and Trognon, 1978; Baltagi and Griffin, 1988; Randolph, 1988 and Li and Stengos, 1994). Randolph's work is based on unbalanced panel data while Baltagi and Griffin (1988) extend on Mazodier and Trognon's work for the balanced case. Baltagi and Griffin (1988) consider heteroskedasticity coming in through the unit specific error component while Li and Stengos (1994) look at heteroskedasticity in the unit-time varying

[1]Note that in Rilstone's experiment, "larger" refers to a sample size of 50 which is actually quite a small sample size, as he points out as well.

error component. The first study uses parametric methods while the second one uses a semiparametric approach[2].Both studies show that their proposed EGLS estimators have the same asymptotic distribution as the true GLS estimator. While Li and Stengos (1994) provide a Monte Carlo study to show that the finite sample behavior of their estimator is adequate as well, Baltagi and Griffin (1988) simply use an empirical example to provide some support for their estimators[3]. Also, the procedure proposed by Baltagi and Griffin (1988) requires a large time component for the panel[4], which may not always be available. In fact, Baltagi (1995, p.78) points out "this is not the typical labor or consumer panel data situation , but it is likely to be the case when pooling a few countries, states or regions over a long time period".

Given that both Rilstone (1991) in the non-panel context, and Li and Stengos (1994) in the panel context find semiparametric estimation in the face of heteroskedasticity encouraging in moderately sized samples, this paper first proposes a semiparametric estimation procedure in the presence of heteroskedasticity of unknown functional form in the unit-specific errors . Our procedure does not require a large time component unlike the estimator proposed by Baltagi and Griffin (1988). We then undertake a Monte Carlo experiment to study the finite sample behavior of the proposed estimator and one of the parametric estimators proposed by Baltagi and Griffin (1988) in the presence of heteroskedasticity of the unit-specific errors; we also include some other standard estimators. This is followed by the application of the various estimators to two empirical examples to illustrate our results. The last section has the conclusion.

[2]Their work is a generalization of the results of Carroll (1982) and Robinson (1987) to the panel data case.

[3]Note that the results in the latter may be specific to the data set and do not necessarily provide convincing evidence of the superiority of their estimators over the standard estimators in finite samples.

[4]For example, if the number of time periods (T) in the data and the number of parameters to be estimated (K) from the model are equal, then the formula given on p.748 in Baltagi and Griffin (1988) is undefined. Also, if T is less than K, which might well happen with short panels with T=2 or 3, the formula returns a negative estimate for the variance.

2 An Error Component Model with Heteroskedasticity:

A standard one-way error component model is given as

$$y_{it} = x_{it}\beta + \mu_i + v_{it} \quad i = 1, ..., N; \; t = 1, ..., T \tag{1}$$

where x_{it} is 1 x k, $\mu_i \sim i.i.d.(0, \sigma_\mu^2)$ is the unit-specific error component while $v_{it} \sim i.i.d.(0, \sigma_v^2)$ is the unit-time varying error component. For details of such models and its estimation, see Baltagi (1995), Hsiao (1986), among others.

According to Baltagi and Griffin (1988, p.745), " particularly bothersome are the assumptions of the absence of autocorrelation in the error term v_{it} and homoskedasticity of the country [unit] specific variances". Like them, we concentrate on the second problem here but unlike them, we propose a semiparametric estimation method (as opposed to a parametric method) to estimate the regression coefficients. The justification of using a semiparametric method stems from the encouraging results of the Monte Carlo studies undertaken by Rilstone (1991) and Li and Stengos (1994) in similar contexts. The model under study is the same as in equation (1) but with a change in the assumption about the unit-specific error μ_i. More specifically, we assume that the μ_i's are i.i.d. with $E(\mu_i|\overline{x}_i) = 0$ and $Var(\mu_i|\overline{x}_i) = \omega(\overline{x}_i) \equiv \omega_i$ where $\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. In other words, the conditional variance of the unit-specific error term has heteroskedasticity of unknown functional form.

Rewriting (1) in vector-matrix form, we have

$$y = x\beta + Z\mu + v \tag{2}$$

where $Z = I_N \otimes e_T$, e_T is a T dimensional column vector of ones and $\mu = [\; \mu_1 \quad \mu_2 .....\mu_N \;]'$. y and v are NT x 1 column vectors of the dependent variable and the unit-time varying error component respectively while x is an NT x k matrix of regressors, all of which are formed by stacking the data using t as the fast index and i as the slow index. Then following Baltagi and Griffin (1988), the inverse of the conditional variance-covariance matrix of the error term in equation (2) (i.e. $Z\mu + v$) denoted by Ω^{-1} is given as

$$\Omega^{-1} = diag[1/\sigma_i^2] \otimes (J_T/T) + diag[1/\sigma_v^2] \otimes (I_T - J_T/T) \tag{3}$$

where $\sigma_i^2 = T\omega_i + \sigma_v^2 \; \forall \; i$ and J_T is a square matrix of ones of dimension T.

The true GLS estimator of β is then given as

$$\tilde{\beta} = (x'\Omega^{-1}x)^{-1}x'\Omega^{-1}y. \tag{4}$$

It should be noted that the formula above involves working with a NT x NT (Ω^{-1}) matrix which can be quite demanding if one has a large data set[5]. So we propose rewriting equation (4) as

$$\tilde{\beta} = \left(\sum_{i=1}^{N} x_i' A_i^{-1} x_i\right)^{-1} \left(\sum_{i=1}^{N} x_i' A_i^{-1} y_i\right) \tag{5}$$

where x_i is a T x K matrix of regressors for the i-th individual, y_i is T x 1 and A_i^{-1} is given as

$$A_i^{-1} = \frac{1}{\gamma_i(1-\rho_i)}\left[I_T - \frac{e_T e_T' \rho_i}{(1-\rho_i+T\rho_i)}\right] \tag{6}$$

with $\rho_i = \omega_i/\gamma_i$ and γ_i is defined below. The calculation of the estimator following the formula given in equation (5) involves using a T x T covariance matrix which for most panel data is quite managable.

Since γ_i and σ_v^2 are unknown, we need to find estimators of them to obtain an EGLS estimator of β. σ_v^2 can be estimated, following Hsiao (1986) in the standard case, as

$$\hat{\sigma}_v^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}[(y_{it} - \bar{y}_i) - \hat{\beta}_w'(x_{it} - \bar{x}_i)]^2}{N(T-1)-k} \tag{7}$$

where $\bar{y}_i$ is similarly defined as $\bar{x}_i$ and $\hat{\beta}_w$ is the within estimator.

Then we define $\gamma_i = E(u_{it}^2|\bar{x}_i) = \omega_i + \sigma_v^2$ and propose the following kernel estimator for γ_i :

[5]For example, in the empirical part, we used a data set from China where NT = 5950. Gauss Version 3.2.34 on a 266 megahertz pentinum machine with 128 megabytes of Ram (with 120 of that allocated to Gauss) reported insufficient workspace memory while trying to calculate $\tilde{\beta}$ by using equation (4). So this can be a real problem for the applied researcher.

$$\hat{\gamma}_i = \frac{\sum_{j=1}^{N}\sum_{t=1}^{T}\hat{u}_{jt}^2 K(\frac{\bar{x}_i - x_{jt}}{h})}{\sum_{j=1}^{N}\sum_{t=1}^{T} K(\frac{\bar{x}_i - x_{jt}}{h})} \tag{8}$$

where $\hat{u}_{jt}$ is the OLS residual from the regression of y_{jt} on x_{jt}, $K(.)$ is the kernel function with h as the smoothing parameter. Once we have $\hat{\gamma}_i$ we can obtain an estimator of ω_i as $\hat{\omega}_i = \hat{\gamma}_i - \hat{\sigma}_v^2$ and hence an estimator of A_i^{-1} by replacing the unknown parameters in (6) by their estimators. The EGLS estimator of β is then obtained as

$$\hat{\beta} = \left(\sum_{i=1}^{N} x_i' \hat{A}_i^{-1} x_i\right)^{-1} \left(\sum_{i=1}^{N} x_i' \hat{A}_i^{-1} y_i\right) \tag{9}$$

3 Monte Carlo Experiment

Here we describe the Monte Carlo experiment undertaken to study the finite sample properties of the proposed estimator and report the relative efficiency of the proposed estimator as compared to other estimators, including the iterative estimator proposed by Baltagi and Griffin (1988) under the same heteroskedasticity assumption. It should be noted that Baltagi and Griffin (1988) did not study the finite sample properties of their proposed estimators.

For comparative purposes, the design of our Monte Carlo experiment is similar to that of Li and Stengos (1994) and Rilstone (1991). The following simple model is considered:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + v_{it} \quad i = 1, ..., N;\ t = 1, ..., T \tag{8}$$

where $x_{it} = 0.5w_{i,t-1} + w_{it}$. We generate w_{it} following two different data generating processes, namely

(i) $w_{it} \sim i.i.d.\ U(0,2)$. We denote this as DGP1.

(ii) $w_{it} \sim i.i.d.\ e^{\nu_{it}}$ and $\nu_{it} \sim i.i.d.\ N(0,(0.4)^2)$ i.e. w_{it} is lognormal. We denote this as DGP2.

The parameters β_0 and β_1 are assigned values 5 and 0.5 respectively. We generate $v_{it} \sim N(0,\sigma_v^2)$ and $\mu_i \sim N(0,\omega_i)$ where $\omega_i = \omega(\bar{x}_i) = \alpha^2(1+\lambda\bar{x}_i)^2$; i.e., we assume heteroskedasticity of additive form. Given our earlier definition of the total variance as $\gamma_i = \omega_i + \sigma_v^2$ and denoting the expected variance of μ_i by $\bar{\omega}$, we set the expected total variance, $\bar{\omega} + \sigma_v^2 = 8$ to make it comparable across the different data generating processes. We let σ_v^2 take values 2, 4 and 6; that is, we vary the share of the variance of the unit-specific error term in the total variance. For each fixed value of σ_v^2, λ is assigned values 0, 1, 2 and 3 where 0 denotes homoskedastic unit-specific error, while the degree of heteroskedasticity increases as the value becomes larger. For a fixed σ_v^2, we obtain a value of $\bar{\omega}$ and using the different values of λ, we get the values for α for each λ value using the additive heteroskedasticity specification above. Then we obtain the values of ω_i for each σ_v^2 under the four different λ values. For the estimation part, the kernel used is the normal kernel[6] since it is well-known in the nonparametric literature that the choice of the kernel function is not crucial as long as it satisfies certain regulatory conditions and the sample size is not small[7]. The choice of the smoothing parameter, h, however, is not as straightforward. One can use the cross-validation approach or the plug-in approach or a "quick and simple" approach[8]; see Wand and Jones (1995) for reference. There are no theoretical results in the literature that clearly indicate the superiority of any one of these approaches in finite samples. The first two approaches are computationally quite intensive and even though they may be desirable asymptotically, their finite sample performances are often not very good. As a result, many researchers have tended to report estimates using different degrees of smoothing by choosing different values of h. For instance, Rilstone (1991) follows this approach of choosing h and reports his results for h = 0.5, h =1 and h = 1.5. Unlike Li and Stengos (1994), he finds results that are sensitive to the choice of h. We also report results

[6]This may be written as:

$$K(\psi_{it}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\psi_{it}^2}{2}\right)$$

where $\psi_{it} = (x - x_{it})/h$.

[7]This result is also reflected in Rilstone's (1991) study where the performances of the estimators were not sensitive to the choice of the kernel function.

[8]An example of a "quick and simple" approach is to choose the smoothing parameter according to the following formula

$$h = cs_x n^{-0.2}$$

where s_x is the sample standard deviation of the regressor, n is the sample size and c is a constant. For example, Li and Stengos (1994) computed h using c = 0.8, c = 1, and c = 1.2.

for h = 0.5, 1 and 1.5 respectively. Our experiments involve 1,000 replications, though we have tried 2,000 replications for some cases and found similar results. We use the following sample sizes:

(I) N = 50, T = 3; i.e., N*T = 150.

(II) N=100, T = 3; i.e., N*T = 300.

We report the relative efficiency[9] of the following estimators:

(a) The proposed EGLS estimator (EGLS).

(b) The iterative EGLS estimator proposed by Baltagi and Griffin (1988) (EGLSB).

(c) The standard GLS estimator for a one-way error components model; that is, the assumption is one of homoskedastic error components (GLSH).

(d) The within or fixed effects estimator which is obtained by sweeping out the unit-specific errors through a transformation (WITHIN). See Baltagi (1995) for details.

(e) The OLS estimator (OLS).

Table 1 gives the relative efficiency of the different estimators for the case where the regressor is generated from a uniform distribution (DGP1), with a sample size of 150, formed with N=50 and T=3. It should be noted that the attractiveness of nonparametric estimators come from their asymptotic properties and this sample size is not very large. The second column of the table reports the relative efficiency of the different estimators when both the error components are homoskedastic. So a priori we expect the standard GLS estimator (GLSH) for a one-way error component model to perform better than any of the other estimators. The results from table 1 shows that this is indeed true, but our proposed EGLS estimator also performs well for all σ_v^2 values. This is not so for the other three estimators. The WITHIN estimator performs somewhat similar to the GLSB with a decrease in relative efficiency as σ_v^2 increases. The OLS estimator performs poorly, and suffers an efficiency loss for smaller values of σ_v^2 (larger values of $\bar{\omega}$) since this estimator does not take into account the unit-specific effect.

(TABLE 1 GOES HERE)

[9] Relative efficiency is defined here as the ratio of the mean square error of the estimator under consideration to the mean square error of the true GLS estimator.

For the cases of heteroskedasticity ($\lambda = 1$ to $\lambda = 3$), the EGLS outperforms all of the other estimators except when $\lambda = 2$ and $\lambda = 3$ under $\sigma_v^2 = 2$; then the GLSB and both the GLSB and the WITHIN estimators are preferable respectively. This may be arising from the fact that the sample size is not very large and the nonparametric estimators are known to be biased in small samples. So we expect that as the sample size increases, the estimator's performance would improve. It is interesting to note that while the performances of the EGLS, the GLSH and the OLS are slipping as we move to higher degrees of heteroskedasticity, from $\lambda = 1$ to $\lambda = 3$, the performance of the EGLSB and the WITHIN actually improve slightly. Overall our results suggest that the performance of the EGLS does not seem too sensitive to the degree of heteroskedasticity, though we do find some sensitivity to the choice of the smoothing parameter. For example, for the homoskedasticity case, under $\sigma_v^2 = 6$, depending on h, the relative efficiency of the EGLS ranges from 1.021 to 1.063. It is interesting to note that even though the EGLSB and the WITHIN perform very well when $\sigma_v^2 = 2$ and $\lambda = 3$,their relative efficiencies slip considerably when $\sigma_v^2 = 6$. In other words, when the degree of heteroskedasticity is high and the share the variance of the unit-specific error in the expected total variance is quite high compared to that of the unit and time specific error, the EGLSB and the WITHIN perform well but not so, when that share is quite low; the EGLS or the GLSH estimators do not show such sensitivity. These results suggest from an applied point of view that when one does not have information on the degree of heteroskedasticity and the share of the different variances of the error terms in the total variance, which would indeed be the case with real data, it might be better to use either the EGLS or the GLSH estimator.

(TABLE 2 GOES HERE)

Table 2 gives the relative efficiency of the different estimators under DGP1 as in Table 1 except that the sample size is now 300 (N=100 and T=3). The results in this table are very similar to that in Table 1, though as expected, the performance of the EGLS has mostly improved with the increase in the sample size. The EGLS performs quite adequately under all cases, which is not the case with the EGLSB, especially when the share of the expected variance of the unit-specific error term in the total expected variance is relatively small;

i.e., $\sigma_v^2 = 6$. The EGLS is again somewhat sensitive to the choice of the window-width.

(TABLE 3 GOES HERE)

Tables 3 and 4 present the relative efficiency of the different estimators when the regressor is generated from a lognormal distribution (DGP2) with sample sizes equal to 150 and 300 respectively. We find results similar to those under DGP1; that is, the EGLS[10] performs reasonably well under the different cases, while the EGLSB does not perform well for higher σ_v^2 values particularly for $\sigma_v^2 = 6$. The simulations also show that the relative efficiency of the GLSH is not too far behind that of the EGLS. Further, as found for DGP1, the EGLS results are sensitive to the choice of window-width, h. For example, in Table 4, when $\sigma_v^2 = 6$ and $\lambda = 3$, the relative efficiency of the EGLS for h = 0.5 is 1.023 while that of the EGLS for h = 1.5 is 1.040, which is the same as the relative efficiency of the GLSH.

(TABLE 4 GOES HERE)

Efficiency of estimators is not our only concern when estimating models. We are also interested in the performance of hypothesis tests regarding the coefficients. To illustrate the impact of using different estimators in a one-way error-component model on hypothesis testing, we consider a test of $H_0 : \beta_1 = 0.5$ against $H_a : \beta_1 \neq 0.5$. Tables 5 and 6 report the empirical size performance of a t-test for the above mentioned hypothesis for DGP1[11]; that is, the tables report the percentage of times the absolute value of the t-ratio is greater than the critical values of a standard normal variable corresponding to the nominal levels of significance of 1%, 5% and 10%. The results show that all of the estimators we consider for an error-component model perform similarly with a tendency to overreject (especially at the 10% level)[12]. As expected, the size distortion is smaller for the larger sample. Our results, on the basis of this Monte Carlo exercise, suggest that we cannot recommend any one of the estimators over the others in terms of their empirical size for the hypothesis test of $H_0 : \beta_1 = 0.5$ against $H_a : \beta_1 \neq 0.5$.

[10]We have used a trimmed version of the estimator with the trimming constant being equal to 0.001, to avoid the problem of getting distorted results from the values of x which are in the tails of the distribution.

[11]The results for DGP2 are similar and are available from the author on request.

[12]We have selected a few entries in Table 5 and report the standard errors associated with them. For example, the standard error associated with the entry 0.054 in Table 5 is 0.007 while that associated with 0.163 is 0.011.

(TABLE 5 and 6 GO HERE)

4 Application:

In this section, we estimate a simple bivariate relationship between calorie intake and income for two different panel data sets using the estimators we examined above. Our purpose is to determine if there is a gain in applying the more computationally intensive estimator proposed in Section 2. It should be noted that according to the discussion in Section 1, we cannot use the estimator proposed by Baltagi and Griffin (1988) for the examples here since both the data sets are short panels with T = 2. We are aware of the fact that there are other variables such as age, household size etc. which influence individual calorie intake as well, but we choose to use income as the only regressor since it is undoubtedly the most influential factor in an individual's consumption decision. Some authors in the calorie-income literature have also done the same; e.g., Subramanian and Deaton (1996). As we have chosen this application for illustrative purposes only, and are not trying to add to the debate on what is the correct magnitude of the income/expenditure elasticity of calorie intake, our modelling assumption is reasonable. The two data sets differ in their sample sizes, one being relatively smaller than the other; details follow.

4.1 Data Descriptions:

The first data set is from the International Crops Research Institute for the Semi-Arid Tropics' (ICRISAT) Village Level Studies which covers over 240 households in six carefully selected representative villages in three districts representing three different agroclimatic and soils regions of India. In each village a random sample of 10 households was selected from the agricultural labor and nonland holding households and another 30 households were a stratified sample of the cultivating households. We use individual calorie intake data[13] and real per capita income data for the years 1976 and 1977. The results are based

[13]Four rounds of nutrition survey were undertaken for the 1976-77 and 1977-78 agricultural years. Data were collected on individual nutrition intakes in the past 24 hours. For details on ICRISAT VLS data, see Binswanger and Jodha (1978), Ryan et al. (1984) and Walker and Ryan (1989). Since the income data is annual, the daily calorie intake data used here are obtained by averaging the rounds observations within

on 730 observations; that is, 365 individuals, each observed over 2 years.

The second data set is a subset from the China Health and Nutrition Survey (CHNS), jointly undertaken by the Carolina Population Center at the University of North Carolina at Chapel Hill, the Institute of Nutrition and Food Hygiene and the Chinese Academy of Preventive Medicine. The data were collected using a multistage, random cluster sampling process to draw a sample of about 3800 households in eight diverse provinces. This paper uses data on individual calorie intake and deflated per capita income for 2975 males[14] for the years 1989 and 1991 giving us a total of 5950 observations.

4.2 Results:

For the ICRISAT data, the null hypothesis of homoskedasticity is tested using White's (1980) test, first for the individual years separately and then for the pooled data. Under the null hypothesis, White's test statistic is distributed χ^2 with k - 1 degrees of freedom. The observed values of the test statistic are 4.760, 8.965 and 12.863 with p-values of approximately 0.0291, 0.0028 and 0.0003 for 1976, 1977 and the pooled data respectively. The results reject homoskedasticity for this data set. The next test undertaken is the Hausman (1978) specification test which returns a p -value of 0.3095, thereby failing to reject the null hypothesis of no correlation between the regressor and the error term[15]. Under this scenario, we use the different estimators considered in Section 3 (except the EGLSB) and report the results in Table 7. For EGLS, the kernel used is the normal kernel and h, the window-width is equal to $cs_xN^{-0.2}$, where s_x is the sample standard deviation of the variable x, $N = 365$ and c is set equal to 0.8, 1 and 1.2 respectively. We find similar estimates of the standard errors and the t-ratios for the statistical significance of the slope parameter from the different estimators, which suggests that there is no gain in using the proposed heteroskedasticity consistent estimator. One reason for this could be

each year.

[14]The Hausman test applied to the complete data and also to the data on only females rejects the null hypothesis of no correlation between the regressor and the error term, making the use of the random effects model inappropriate.

[15]If there exists a correlation between the regressor and the error term, then the GLSH will be an inconsistent estimator.

that the heteroskedasticity is in the unit-time specific error component and not the unit-specific error component. Unfortunately the currently available heteroskedasticity tests cannot distinguish between the sources of the heteroskedasticity as they are based on the combined error term, u_{it}.

For the CHNS data, the results are very similar to that for the ICRISAT data (see Table 7); that is, even when White's test strongly rejects the homoskedasticity hypothesis[16], there seems to be no gain in using the proposed heteroskedasticity consistent estimator. Again, it could be the case that the heteroskedasticity is not coming from the unit-specific error term.

(TABLE 7 GOES HERE)

5 Conclusion:

Given that various studies in the econometrics literature have shown that adaptive estimators work well in the presence of heteroskedasticity even in moderately sized samples for both panel data[17] and standard cases, this paper first derives an adaptive estimator using nonparametric methods when the the heteroskedasticity is present in the unit-specific error in a one-way error component model. A Monte Carlo exercise is then undertaken to compare the finite sample performances of various estimators including the proposed one. We then use the different estimators in two empirical applications to examine whether they result in practical differences.

The Monte Carlo results show that the relative efficiency of the proposed estimator is adequate, but it is somewhat sensitive to the choice of the window-width. However, in terms of the size performance, all the estimators considered behave in a similar fashion; that is, none of them overreject or underreject substantially. So on the basis of the size performance, we cannot claim the superiority of our proposed estimator over the other ones. This result is quite unlike that of Li and Stengos (1994) who find that their proposed adaptive estimator in the presence of conditional heteroskedasticity of unknown form in the unit-time specific

[16] The observed value of the test statistic is 7.85 with a p-value of 0.005.

[17] This is the case when the variance of the unit-time specific error component is heteroskedastic.

error term is the only one which has the correct size under heteroskedasticity, while the others overreject to a large extent. Also, unlike them, our size performance results are also sensitive to the choice of the window-width.

We use two data sets of quite different sample sizes to estimate a simple bivariate relation between calorie intake and income to illustrate the usefulness of using one estimator over another in the presence of heteroskedasticity. Even though the hypothesis of homoskedasticity is soundly rejected in both cases, all the estimators give very similar results in terms of the coefficient estimates and the standard errors. One inference we can draw is that if the heteroskedasticity is present in the unit-specific term, then one can still use the standard GLS estimator for an error component model. Practically, this means that one has to first distinguish between the source of heteroskedasticity; that is, whether it is from the unit-specific error term or from the unit-time specific error term before deciding on the correct estimator. This suggests another way to view our results; the heteroskedasticity in the two examples considered in this paper are indeed coming from the unit-time specific error term and hence an estimator which takes into account heteroskedasticity in the other error component may not make a difference. This suggests a need for a test for heteroskedasticity that will distinguish between heteroskedasticity coming in from the unit-specific error component and that from the unit-time specific error component. This is beyond the scope of this paper and is the subject of current research.

Table 1
Relative Efficiency: DGP1, N=50, T=3.

		$\sigma_v^2 = 2$		
$\lambda =$	0	1	2	3
EGLS (h = 0.5)	1.027	1.081	1.103	1.108
EGLS (h = 1)	1.020	1.072	1.099	1.105
EGLS (h = 1.5)	1.017	1.076	1.107	1.117
EGLSB	1.082	1.095	1.076	1.076
GLSH	1.013	1.117	1.130	1.135
WITHIN	1.101	1.119	1.091	1.084
OLS	2.843	4.871	5.696	5.962
		$\sigma_v^2 = 4$		
EGLS (h = 0.5)	1.048	1.099	1.136	1.149
EGLS (h = 1)	1.028	1.087	1.126	1.140
EGLS (h = 1.5)	1.022	1.095	1.138	1.154
EGLSB	1.189	1.189	1.162	1.158
GLSH	1.016	1.158	1.177	1.186
WITHIN	1.266	1.264	1.210	1.198
OLS	1.483	2.227	2.480	2.565
		$\sigma_v^2 = 6$		
EGLS (h = 0.5)	1.063	1.083	1.105	1.115
EGLS (h = 1)	1.035	1.067	1.093	1.104
EGLS (h = 1.5)	1.021	1.071	1.098	1.110
EGLSB	1.349	1.321	1.285	1.274
GLSH	1.014	1.120	1.135	1.142
WITHIN	1.545	1.487	1.416	1.399
OLS	1.091	1.344	1.414	1.439

Table 2
Relative Efficiency: DGP1, N=100, T=3.

		$\sigma_v^2 = 2$		
$\lambda =$	0	1	2	3
EGLS (h = 0.5)	1.000	1.071	1.052	1.048
EGLS (h = 1)	1.000	1.074	1.065	1.062
EGLS (h = 1.5)	1.000	1.083	1.084	1.084
EGLSB	1.102	1.097	1.083	1.077
GLSH	1.000	1.123	1.118	1.116
WITHIN	1.135	1.130	1.098	1.093
OLS	2.948	5.245	6.052	6.309
		$\sigma_v^2 = 4$		
EGLS (h = 0.5)	1.006	1.085	1.089	1.092
EGLS (h = 1)	1.002	1.088	1.101	1.105
EGLS (h = 1.5)	1.000	1.100	1.125	1.134
EGLSB	1.225	1.199	1.174	1.171
GLSH	0.998	1.162	1.175	1.181
WITHIN	1.317	1.272	1.232	1.226
OLS	1.545	2.387	2.650	2.736
		$\sigma_v^2 = 6$		
EGLS (h = 0.5)	1.015	1.073	1.095	1.103
EGLS (h = 1)	1.005	1.069	1.096	1.105
EGLS (h = 1.5)	1.001	1.078	1.112	1.123
EGLSB	1.379	1.305	1.301	1.296
GLSH	0.997	1.123	1.150	1.157
WITHIN	1.603	1.484	1.454	1.449
OLS	1.128	1.413	1.505	1.532

Table 3
Relative Efficiency: DGP2, N=50, T=3.

	$\sigma_v^2 = 2$			
$\lambda =$	0	1	2	3
EGLS (h = 0.5)	1.005	1.009	1.010	1.020
EGLS (h = 1)	1.005	1.013	1.015	1.016
EGLS (h = 1.5)	1.005	1.015	1.017	1.018
EGLSB	1.089	1.060	1.051	1.046
GLSH	1.004	1.017	1.020	1.021
WITHIN	1.108	1.081	1.071	1.066
OLS	3.048	4.254	5.019	5.500
	$\sigma_v^2 = 4$			
EGLS (h = 0.5)	1.016	1.020	1.022	1.022
EGLS (h = 1)	1.013	1.026	1.031	1.032
EGLS (h = 1.5)	1.012	1.028	1.034	1.037
EGLSB	1.204	1.162	1.146	1.137
GLSH	1.011	1.031	1.038	1.041
WITHIN	1.272	1.209	1.185	1.173
OLS	1.553	1.952	2.207	2.368
	$\sigma_v^2 = 6$			
EGLS (h = 0.5)	1.024	1.028	1.031	1.032
EGLS (h = 1)	1.016	1.028	1.035	1.039
EGLS (h = 1.5)	1.014	1.030	1.038	1.042
EGLSB	1.344	1.308	1.289	1.275
GLSH	1.014	1.032	1.042	1.047
WITHIN	1.543	1.445	1.404	1.382
OLS	1.115	1.233	1.312	1.363

Table 4
Relative Efficiency: DGP2, N=100, T=3.

		$\sigma_v^2 = 2$		
$\lambda =$	0	1	2	3
EGLS (h = 0.5)	0.998	1.011	1.014	1.016
EGLS (h = 1)	0.999	1.074	1.020	1.022
EGLS (h = 1.5)	0.999	1.083	1.023	1.025
EGLSB	1.089	1.077	1.071	1.068
GLSH	0.999	1.020	1.025	1.027
WITHIN	1.119	1.093	1.083	1.078
OLS	2.994	4.258	5.043	5.535
		$\sigma_v^2 = 4$		
EGLS (h = 0.5)	1.000	1.014	1.019	1.021
EGLS (h = 1)	1.001	1.021	1.029	1.032
EGLS (h = 1.5)	1.001	1.024	1.033	1.037
EGLSB	1.202	1.165	1.148	1.139
GLSH	1.001	1.026	1.037	1.041
WITHIN	1.294	1.229	1.204	1.192
OLS	1.543	1.961	2.225	2.391
		$\sigma_v^2 = 6$		
EGLS (h = 0.5)	1.009	1.018	1.022	1.023
EGLS (h = 1)	1.006	1.019	1.027	1.032
EGLS (h = 1.5)	1.006	1.022	1.031	1.040
EGLSB	1.370	1.315	1.289	1.275
GLSH	1.006	1.024	1.034	1.040
WITHIN	1.587	1.482	1.439	1.416
OLS	1.113	1.235	1.318	1.372

Table 5
Empirical Size: DGP1, N=50, T=3.

	$\boldsymbol{\sigma_v^2 = 2}$			$\boldsymbol{\sigma_v^2 = 4}$			$\boldsymbol{\sigma_v^2 = 6}$		
Nominal Significance level	1%	5%	10%	1%	5%	10%	1%	5%	10%
					$\boldsymbol{\lambda = 0}$				
EGLS (h = 0.5)	1.6	5.4	12.0	1.6	6.4	12.6	1.3	7.2	13.2
EGLS (h = 1)	1.4	5.6	11.8	1.5	5.5	12.5	1.2	7.1	12.3
EGLS (h = 1.5)	1.4	5.5	11.9	1.5	5.0	12.4	1.1	6.7	12.3
EGLSB	2.0	7.3	13.6	2.0	7.4	13.1	2.1	7.6	12.5
GLSH	1.3	5.7	11.8	1.2	4.9	12.3	0.9	5.6	11.9
					$\boldsymbol{\lambda = 1}$				
EGLS (h = 0.5)	1.7	6.3	13.6	1.8	7.9	15.0	1.8	7.9	13.9
EGLS (h = 1)	1.7	6.2	13.1	2.0	7.3	14.7	1.7	8.2	13.7
EGLS (h = 1.5)	1.7	6.3	12.7	2.1	7.3	14.3	1.6	7.6	13.6
EGLSB	2.0	7.0	13.9	1.9	8.1	13.6	2.1	7.8	14.0
GLSH	1.7	7.0	12.9	2.2	7.8	14.5	2.2	8.6	13.8
					$\boldsymbol{\lambda = 2}$				
EGLS (h = 0.5)	1.8	6.9	14.2	2.5	7.9	16.3	2.3	9.5	15.1
EGLS (h = 1)	1.8	6.5	14.3	2.3	8.0	15.4	2.3	8.9	14.5
EGLS (h = 1.5)	2.0	6.5	14.3	2.4	8.0	15.5	2.3	9.1	14.8
EGLSB	2.0	6.5	14.0	2.1	7.5	13.7	2.0	7.5	13.6
GLSH	1.9	6.9	14.1	2.5	9.2	15.7	2.6	8.8	15.4
					$\boldsymbol{\lambda = 3}$				
EGLS (h = 0.5)	2.1	6.9	14.3	2.6	8.0	16.3	2.7	9.7	15.5
EGLS (h = 1)	2.0	6.7	14.3	2.7	8.0	15.4	2.7	9.2	14.6
EGLS (h = 1.5)	1.9	6.8	14.5	2.6	8.5	16.3	2.5	9.4	14.9
EGLSB	2.0	6.4	13.7	2.3	7.9	13.6	2.1	7.5	13.0
GLSH	2.1	7.1	14.8	2.7	9.3	16.2	2.8	9.1	15.6

Table 6
Empirical Size: DGP1, N=100, T=3.

Nominal Significance level	$\boldsymbol{\sigma_v^2 = 2}$ 1%	5%	10%	$\boldsymbol{\sigma_v^2 = 4}$ 1%	5%	10%	$\boldsymbol{\sigma_v^2 = 6}$ 1%	5%	10%
					$\boldsymbol{\lambda = 0}$				
EGLS (h = 0.5)	0.7	4.6	8.6	0.6	4.3	8.8	0.6	4.9	11.0
EGLS (h = 1)	0.7	4.6	8.5	0.4	4.3	8.6	0.5	4.8	10.4
EGLS (h = 1.5)	0.7	4.6	8.6	0.3	4.3	8.7	0.6	4.8	10.1
EGLSB	1.5	5.9	11.8	1.9	6.7	12.2	1.9	5.8	11.2
GLSH	0.8	4.5	8.7	0.3	4.5	9.0	0.7	4.8	10.1
					$\boldsymbol{\lambda = 1}$				
EGLS (h = 0.5)	1.0	5.4	10.8	1.0	6.1	11.8	1.1	6.3	12.1
EGLS (h = 1)	1.1	5.4	10.6	0.9	5.8	11.4	1.1	6.3	12.2
EGLS (h = 1.5)	1.2	5.2	10.7	1.3	5.4	12.0	1.1	6.4	12.1
EGLSB	1.4	6.3	12.2	1.8	6.2	12.6	2.1	5.9	11.3
GLSH	1.5	5.9	10.5	1.9	6.3	12.1	1.4	6.8	11.8
					$\boldsymbol{\lambda = 2}$				
EGLS (h = 0.5)	1.1	5.7	11.1	1.3	6.4	11.4	1.2	7.5	13.5
EGLS (h = 1)	1.2	5.8	10.8	1.7	6.8	12.6	1.1	7.1	13.6
EGLS (h = 1.5)	1.4	5.5	11.1	2.1	6.7	13.1	1.5	6.9	13.7
EGLSB	1.7	6.5	12.1	1.4	7.1	11.9	1.9	6.0	11.6
GLSH	1.5	5.8	11.4	2.3	6.7	12.9	1.8	7.4	13.8
					$\boldsymbol{\lambda = 3}$				
EGLS (h = 0.5)	1.2	5.7	10.5	1.5	6.5	12.0	1.2	7.6	13.4
EGLS (h = 1)	1.5	5.8	10.9	1.9	6.7	12.7	1.4	7.0	13.9
EGLS (h = 1.5)	1.6	5.9	11.0	2.2	7.1	13.4	1.7	6.9	14.5
EGLSB	1.7	6.2	11.5	1.5	7.1	12.0	1.9	5.8	11.0
GLSH	1.7	6.5	11.5	2.3	7.4	13.8	1.7	7.1	14.4

Table 7
Estimates of the income elasticity of calorie intake
in the presence of heteroskedasticity

Estimator	Coefficient estimate	standard error	t- ratio
	ICRISAT		
EGLS (c = 0.8)	0.119	0.018	6.513
EGLS (c = 1)	0.120	0.018	6.503
EGLS (c = 1.2)	0.120	0.019	6.498
GLSH	0.126	0.019	6.664
WITHIN	0.109	0.026	4.255
OLS	0.138	0.019	7.104
	CHNS		
EGLS (c = 0.8)	0.007	0.006	1.190
EGLS (c = 1)	0.007	0.006	1.182
EGLS (c = 1.2)	0.007	0.006	1.174
GLSH	0.006	0.006	1.075
WITHIN	0.007	0.007	0.966
OLS	0.005	0.007	0.876

ACKNOWLEDGEMENTS

I am most grateful to Judith Giles for her careful reading of the paper and the helpful comments. I would like to thank Aman Ullah and Qi Li for encouraging me to look into this issue. I have also benefited from the comments of David Giles and other seminar participants at the University of Victoria. I gratefully acknowledge the financial support from the internal grant provided by the Univeristy of Victoria. All remaining errors are mine.

References

[1] Baltagi, B. H. (1985), *Econometric Analysis of Panel Data* (John Wiley and Sons).

[2] Baltagi, B. H. (1988), Panel Data Methods, *Handbook of Applied Economic Statistics*, A. Ullah and D.E.A. Giles (eds), Marcel Dekker, New York.

[3] Baltagi, B.H. and M. Griffin (1988), A Generalized Error Component Model with Heteroskedastic Disturbances, *International Economic Review*, 29: 745-753.

[4] Binswanger, H. P. and N. S. Jodha (1978), *Manuals of Instructions for Economic Investigators in ICRISAT's Village Level Studies*, Vol. 2 Village Level Studies Series, Economics Program. Patancheru: ICRISAT.

[5] Carroll, R. D. (1982), Adapting for Heteroskedasticity in Linear Models, *The Annals of Statistics*, 10: 1224-1233.

[6] Delgado, M. A. (1992), Semiparametric Generalized Least Squares Estimation in the Multivariate Nonlinear Regression Model, *Econometric Theory*, 8: 203-222.

[7] Hausman, J. A. (1978), Specification Tests in Econometrics, *Econometrica*, 46: 1251-1271.

[8] Hidalgo, J. (1992), Adaptive Estimation in Time Series Regression Models with Heteroskedasticity of Unknown Form, *Econometric Theory*, 8: 161-187.

[9] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press, New York.

[10] Li, Q. and T. Stengos (1994), Adaptive Estimation in the Panel Data Error Component Model with Heteroskedasticity of Unknown Form, *International Economic Review*, 35: 981-1000.

[11] Mazodier, P. and A. Trognon (1978), Heteroskedasticity and Stratification in Error Components Models, *Annales de l'Insee*, 30-31: 451-482.

[12] Randolph, W. C. (1988), A Tranformation for Heteroskedastic Error Components Regression Models, *Economics Letters*, 27: 349-354.

[13] Rilstone, P. (1991), Some Monte Carlo Evidence on the Relative Efficiency of Parametric and Semiparametric EGLS Estimators, *Journal of Business and Economic Statistics*, 9: 179-187.

[14] Robinson, P. M. (1987), Asymptotic Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form, *Econometrica*, 55: 817-891.

[15] Ryan, J. G., P. D. Bidinger, P. Pushpamma and N. Prahlad Rao, 1984, *The Determinants of Individual Diets and Nutritional Status in Six Villages of Southern India*, Research Bulletin No. 7, Patancheru: ICRISAT.

[16] Walker, T. S. and J. G. Ryan, 1989, *Village and Household Economies in India's Semi-Arid Tropics* (The Johns Hopkins University Press).

[17] Wand, M.P. and Jones, M.C. (1995), *Kernel Smoothing*, St. Edmundsbury Press, Suffolk.