

Improved Inference of Heteroscedastic Fixed Effects Models

Afshan Saeed

Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan
phdstatbzu@gmail.com

Muhammad Aslam

Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan
aslamasadi@bzu.edu.pk

Abstract

Heteroscedasticity is a stern problem that distorts estimation and testing of panel data model (PDM). Arellano (1987) proposed the White (1980) estimator for PDM with heteroscedastic errors but it provides erroneous inference for the data sets including high leverage points. In this paper, our attempt is to improve heteroscedastic consistent covariance matrix estimator (HCCME) for panel dataset with high leverage points. To draw robust inference for the PDM, our focus is to improve kernel bootstrap estimators, proposed by Racine and MacKinnon (2007). The Monte Carlo scheme is used for assertion of the results.

Keywords: Bootstrap; HCCME; Kernel smoothing; Leverage points; Size distortion.

1. Introduction

Panel data is the combination of time series and cross-sectional units collected from the same group of cross-sectional units over time. In econometric research, it has many advantages over time series and cross-sectional data. The estimation of PDM includes two sorts of cases, the fixed effect model (FEM) and random effect model (REM). In the first stated case, the individual-specific effect or heterogeneity is assumed as fixed while in the last mentioned, it is expected to be random. The FEM has been frequently used in panel data analysis for the issue of individual heterogeneity.

An important assumption of classical linear regression model (CLRM) is homoscedasticity that the variance of error term remains constant and thus, the error term is identically distributed. The results are inadequate if this assumption is not met. The ordinary least square (OLS) estimates are not biased and inconsistent but they do not remain best linear unbiased estimator (BLUE) when the assumption of homoscedasticity is violated. Heteroscedasticity is a usual problem in the PDM, like unit-specific heteroscedasticity (USH) and unit-time varying heteroscedasticity (UTVH) and it is desirable to concentrate on it for making some robust inference.

Earlier studies about the problem of heteroscedasticity in the PDM were carried out by Mazodier and Trognon (1978). Eicker (1963) and White (1980) proposed heteroscedastic consistent covariance matrix estimator (HCCME) for non-panel data to tackle the problem of heteroscedasticity, which makes it conceivable to draw asymptotically robust inference. In the existing literature, it can be seen that Arellano (1987) builds White's estimator for the PDM. Uchôa et al. (2014) used another variant of the White estimator for the case of FE. However, the HCCME proposed by Cribari-Neto et al. (2007) for cross-sectional data has not been used for the study of panel data yet by any researcher.

Thus, in current study, this estimator is being used for the improvement in inference of PDM.

Besides the HCCME, some bootstrap estimators have also been developed to draw correct inference about the PDM. Cameron et al. (2008) and MacKinnon and Webb (2013) used the bootstrap technique in order to draw robust inference for the PDM with clustered errors. Unfortunately, a wide literature on the use of bootstrap estimators is not observed for the PDM, especially when there are also some high leverage points.

Another statistical approach is also available in the literature to make robust inference of linear regression model. This is the kernel smoothing approach proposed by Racine and MacKinnon (2007). This approach has not been studied in context of the PDM yet by any researcher and it is tried to fill this gap in our current work.

The paper is outlined as follows. Section 2 provides the model description and the HCCMEs. Section 3 discusses the HCCME based quasi t -test statistic. Section 4 presents the bootstrap estimators and Section 5 presents the kernel based versions of bootstrap. Section 6 displays all the numerical results and Section 7 concludes the results.

2. The Model and HCCMEs

The concerned model is the heteroscedastic FEM and its basic framework is

$$y_{it} = \mu_i + x_{it} \beta + v_{it}, i = 1, \dots, n; T = 1, \dots, T, \quad (1)$$

where μ_i is unobserved heterogeneity. Model (1) can be embodied in the form of matrix by stacking the data over time dimension T

$$y_i = X_i \beta + \mu_i + v_i,$$

where $y_i = (y_{i1}, \dots, y_{iT})'$, $X_i = (x_{i1}, \dots, x_{iT})'$ and $v_i = (v_{i1}, \dots, v_{iT})'$. Following model is obtained by assembling the whole data set cross-sectionally

$$y = \mu \otimes e_T + X\beta + v, \quad (2)$$

where $y = (y_1, \dots, y_n)'$, $X = (X_1, \dots, X_n)'$ and $v = (v_1, \dots, v_n)'$. Here, e_T is $T \times 1$ vector of ones, unobserved individual heterogeneity is captured by $\mu = (\mu_1, \dots, \mu_n)'$ and \otimes is the Kronecker product. Additionally, y is $nT \times 1$ response vector, X is $nT \times q$ implicitly fixed regressor ($q < nT$), β is $q \times 1$ vector of unknown parameters and v is $nT \times 1$ vector of error term (for more details, see Uchôa et al. 2014). Here, we have some assumptions about error term that $E(v_{it}) = 0$, $\text{var}(v_{it}) = \sigma_v^2$.

The estimation of (2) can be achieved by within group estimator (WGE). It can be accomplished by pre-multiplication of (2) with the following matrix

$$M = I_{nT} - P,$$

where $P=Q(Q'Q)^{-1}Q'$, $Q=I_n \otimes \bar{e}_T$ is $nT \times n$ matrix of n dummy variables associated with each cross-sectional unit, I_{nT} and I_n is the identity matrix of order nT and n , respectively. $\bar{e}_T=T^{-1} \times e_T$ is a vector of ones of index T (for details; see Greene, 2003). Model (2) becomes

$$\tilde{y} = \tilde{X}\tilde{\beta} + \tilde{v}, \quad (3)$$

where $\tilde{X}=MX$, $\tilde{y}=My$ and $\tilde{v}=Mv$, the appropriate estimator for (3) is the OLS. The unbiased estimator of β is

$$\hat{\tilde{\beta}} = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\tilde{y}.$$

The variance of estimator is

$$\text{var}(\hat{\tilde{\beta}}) = \hat{\sigma}^2 (\tilde{X}'\tilde{X})^{-1}, \quad (4)$$

where $\hat{\sigma}^2 = \frac{\hat{v}'\hat{v}}{n(T-1)-tr(\tilde{H})}$, $\hat{v}=(\tilde{y}-\tilde{X}\hat{\tilde{\beta}})$ and the hat matrix can be defined as $\tilde{H} = \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$.

The usual covariance matrix of $\hat{\tilde{\beta}}$ can be described as

$$\tilde{\varphi} = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\tilde{\Omega}\tilde{X}(\tilde{X}'\tilde{X})^{-1} = D\tilde{\Lambda}D', \quad (5)$$

where $D = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'$. In case of homoscedasticity, Eq. (5) reduces to variance given in (4). For the heteroscedastic case, the error term is independent but non-identically distributed, hence there is a need of consistent estimator of $\tilde{\varphi}$. Arellano (1987) improved White (1980) estimator for the FEM which can be defined as

$$HC0 = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\hat{\tilde{\Omega}}_d\tilde{X}(\tilde{X}'\tilde{X})^{-1} = D\hat{\tilde{\Lambda}}_0D' = \hat{\tilde{\varphi}}_0,$$

where $\hat{\tilde{\Lambda}}_0 = \text{diag} \{ \hat{v}_1^2, \hat{v}_2^2, \dots, \hat{v}_n^2 \}$.

For the FEM, HC3 can be found by replacement of $\hat{\tilde{\Lambda}}_0$ with following diagonal elements in (5)

$$\hat{\tilde{\Lambda}}_3 = \text{diag} \left\{ \frac{\hat{v}_1^2}{(1-\tilde{h}_1)^2}, \frac{\hat{v}_2^2}{(1-\tilde{h}_2)^2}, \dots, \frac{\hat{v}_n^2}{(1-\tilde{h}_n)^2} \right\},$$

where \tilde{h}_{it} is being the it^{th} diagonal element of \tilde{H} .

Uchôa et al. (2014) used the HC4 for FEM having high leverage points. For this estimator,

$$\hat{\tilde{\Lambda}}_4 = \text{diag} \left\{ \frac{\hat{v}_1^2}{(1-\tilde{h}_1)^{\pi_1}}, \frac{\hat{v}_2^2}{(1-\tilde{h}_2)^{\pi_2}}, \dots, \frac{\hat{v}_n^2}{(1-\tilde{h}_n)^{\pi_n}} \right\},$$

where $\pi_i = \min \left\{ 4, \frac{\tilde{h}_{i1}}{\tilde{h}_{i1}} \right\}, \dots, \min \left\{ 4, \frac{\tilde{h}_{iT}}{\tilde{h}_{iT}} \right\}$. Since $0 < \tilde{h}_i < 1$ and $\pi_i > 0$, therefore $0 < (1 - \tilde{h}_i)^{\pi_i} < 1$. In

the available literature, the HC5 version of HCCME has not been used for the FEM which has proposed by Cribari-Neto et al. (2007) for non-panel data to study the influence of maximal leverage. In our present work, we propose to use the HC5 for the FEM. For this purpose, we use

$$\hat{\Delta}_5 = \text{diag} \left\{ \frac{\hat{\tilde{v}}_1^2}{(1 - \tilde{h}_1)^{\delta_1}}, \frac{\hat{\tilde{v}}_2^2}{(1 - \tilde{h}_2)^{\delta_2}}, \dots, \frac{\hat{\tilde{v}}_n^2}{(1 - \tilde{h}_n)^{\delta_n}} \right\},$$

$$\text{where } \delta_i = \min \left[\left\{ \frac{\tilde{h}_{i1}}{\tilde{h}_{i1}} \right\}, \dots, \left\{ \frac{\tilde{h}_{iT}}{\tilde{h}_{iT}} \right\}, \max \left\{ \left\{ 4, \frac{c\tilde{h}_{\max}}{\tilde{h}_{i1}} \right\}, \dots, \left\{ 4, \frac{c\tilde{h}_{\max}}{\tilde{h}_{iT}} \right\} \right] \right], \quad 0 < c < 1.$$

The condition implicates that $\delta_i > 0$ and $0 < \tilde{h}_i < 1$, it entails that $0 < (1 - \tilde{h}_i)^{\delta_i} < 1$

The general structure of the HCCME is

$$HC(l) : \hat{\tilde{\varphi}}_l = D \hat{\Delta}_l D', \quad (6)$$

where $l = 0, 3, \dots, 5$, $l = 0$ indicates HC0, $l = 3$ is for HC3 and so on.

3. The HCCME based t -test and the Heteroscedastic-Consistent Covariance Interval Estimator (HCCIE)

The quasi- t test statistic can be intended with the concept of normality of regression parameter, such that $\hat{\tilde{\beta}}_r \sim N(\beta_r, \tilde{\varphi}_{(rr)})$.

Consider the hypothesis

$$H_0 : \beta_r = \beta_r^0 \text{ against } H_1 : \beta_r \neq \beta_r^0, \text{ where } \beta_r^0 \text{ is hypothesized value of } \beta_r.$$

The t -test statistic under the null hypothesis is

$$\hat{\tilde{t}}_r = \frac{(\hat{\tilde{\beta}}_r - \beta_r^0)}{\sqrt{\hat{\tilde{\varphi}}_{(rr)}}}. \quad (7)$$

where $\hat{\tilde{\varphi}}_{(rr)}$ is the r^{th} diagonal element of $\hat{\tilde{\varphi}}$ and $r = 0, 1, \dots, q - 1$. The $\hat{\tilde{t}}_r$ is likely to follow a $t_{\lfloor n(T-1) - p(\tilde{H}) \rfloor}$ distribution. In case of large sample size, Halawa and El-Bassiouni (2000) noted that the asymptotic distribution of this statistic is normal. Thus a test of asymptotic significance α rejects H_0 if $|\hat{\tilde{t}}_{WG}| > Z_{1-\frac{\alpha}{2}}$, where $Z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ quantile of standard normal distribution (SND). Thus, the true size of test can be built up as

$$P(\text{reject } H_0 \mid H_0) = P\left(\left| \hat{\tilde{t}}_{(r)} \right| > Z_{\frac{\alpha}{2}} \mid \beta_r = \beta_r^0 \right). \quad (8)$$

The power of test can be measured as

$$P(\text{reject } H_0 \mid H_1) = P\left(\left|\hat{t}_{(r)}\right| > Z_{\frac{\alpha}{2}} \mid \beta_r = \beta_r^1\right). \quad (9)$$

Now for the case of heteroscedastic errors, the Statistic (7) can be improved as

$$\hat{t}_r = \frac{(\hat{\beta}_r - \beta_r^0)}{\sqrt{\hat{\phi}_{l(\pi)}}}, \text{ where } l = 0, 3, \dots, 5.$$

Cribari-Neto and Lima (2009) constructed confidence interval based on the $\hat{\beta}_{OLS}$ and HCCMEs for linear regression models. But, we are going to construct confidence interval for the FEM. Similarly, for large dataset, the HCIE can be derived and $(1-\alpha) \times 100 \%$ (two sided) confidence interval for β is

$$\hat{\beta}_r \pm Z_{1-\frac{\alpha}{2}} \sqrt{\hat{\phi}_{l(\pi)}}. \quad (10)$$

4. The Bootstrap Estimator

To draw robust inference about the heteroscedastic linear FEM, some bootstrap estimators have also been observed. The residual bootstrap (RB), WB and pair bootstrap (PB) estimators have been frequently used for the PDM (for more details; see Cameron et al., 2008). Cameron et al. (2008) have considered different bootstrap procedures for clustered errors like bootstrap- T method and bootstrap- se method. MacKinnon (2007) discussed the bootstrap p -value procedures for non-panel data. The bootstrap schemes are given below.

4.1 The Bootstrap- T Procedures

4.1.1 The RB Estimator (t) (RBE (t))

The bootstrap scheme for the residual resampling is as follows

1. Create pseudo-sample of residuals \tilde{v} by resampling of residual $\hat{\tilde{v}}$.
2. Make a bootstrap sample (y_{RB}, \tilde{X}) , as $y_{RB} = \tilde{X}\hat{\beta} + \tilde{v}$,
where $\hat{\beta}$ and \tilde{v} are the WG coefficient and residual, respectively.
3. Estimate $\hat{\beta}_{RB} = (X'MX)^{-1}X'My_{RB}$.
4. Construct $\hat{\tau}_{RB} = \frac{(\hat{\beta}_{RB} - \hat{\beta})}{\sqrt{\text{var}(\hat{\beta}_{RB})}}$.
5. Repeat Steps 1 to 4 for large number of (say, B) times.

Reject the null hypothesis at level α if and only if

$$|\hat{\tau}| > \hat{\tau}_{RB[1-\frac{\alpha}{2}]},$$

where $t_{RB[q]}$ represents the q^{th} quantile of $\hat{\tau}_{RB_1}, \dots, \hat{\tau}_{RB_B}$.

The percentile- t two-sided confidence interval at level $1 - \alpha$ is

$$\hat{\beta} \pm \hat{\tau}_{RB \left[\frac{1-\alpha}{2} \right]} \sqrt{\hat{\sigma}^2 (\tilde{X}' \tilde{X})^{-1}}.$$

4.1.2 The WB Estimator (t) (WBE (t))

It has proposed by Liu (1988), who followed the suggestions of Wu (1986) and Beran (1986). Its scheme is given below.

1. For each m , $m = 1, \dots, n$, draw a random number R_m from a population that has zero mean and unit variance.
2. Construct a bootstrap sample (y_{WB}, \tilde{X}) , as

$$y_{WB} = \tilde{X} \hat{\beta} + \frac{\hat{v}}{1 - \tilde{h}_i} R_m, \quad (11)$$

where \tilde{h}_i is the i^{th} diagonal element of hat matrix, $\hat{\beta}$ and \hat{v} are the WG coefficient and residual, respectively.

Steps 3 to 5 are similar to the scheme of ARB (t). The weight in (11) is based on HC3.

4.2 The Bootstrap- p Procedures

MacKinnon (2007) presented the way of computing bootstrap p -value for linear regression model.

4.2.1 The RBE (p)

For the RB, the bootstrap- p technique is as follows

$$\hat{p}_{RB(p)} = 1 - \hat{F}(\hat{\tau}) = \frac{1}{B} \sum_{j=1}^B I(|\hat{\tau}_{j(RB)}| > |\hat{\tau}|), \quad j=1, \dots, B,$$

where $\hat{F}(\hat{\tau}) = \frac{1}{B} \sum_{j=1}^B I(|\hat{\tau}_{j(RB)}| > |\hat{\tau}|)$ denotes the empirical distribution function (EDF),

$I(\cdot)$ is the indicator function, $\hat{\tau}_{RBWG}$ is the t -statistic obtained from RB and $\hat{\tau}$ is the t -statistic acquired from WG. It is the way of computing symmetric bootstrap p -value.

4.2.2 The WBE (p)

The method of the computation of the symmetric bootstrap p -value on the basis of the WB is as follows

$$\hat{p}_{WB(p)} = 1 - \hat{F}(\hat{\tau}) = \frac{1}{B} \sum_{j=1}^B I(|\hat{\tau}_{j(WB)}| > |\hat{\tau}|),$$

where $\hat{\tau}_{PBWG}$ is the t -statistic comes from WB and $\hat{F}(\hat{\tau}) = \frac{1}{B} \sum_{j=1}^B I(|\hat{\tau}_{j(WB)}| > |\hat{\tau}|)$.

5. The Kernel Estimators

The kernel estimator of Racine and MacKinnon (2007) has proposed for non-panel model. In the present work, we propose to use this estimator while considering the FEM. The scheme of kernel estimators is given below:

5.1 The RB Kernel Estimator

The kernel estimator is going to be used for the RBE and can be defined as

$$\hat{P}_{RBK} = 1 - \hat{F}_d(\hat{t}) = 1 - \frac{1}{B} \sum_{j=1}^B K(\hat{t}_{j(RB)}, \hat{t}, d),$$

where $\hat{F}_d(\hat{t}) = \frac{1}{B} \sum_{j=1}^B K(\hat{t}_{j(RB)}, \hat{t}, d)$, $j=1, \dots, B$, $\hat{F}_d(\hat{t})$ is the cumulative distribution function (CDF), d is the bandwidth, \hat{t} is the t test statistic from WGE and \hat{t}_{RB} is the quasi- t test statistic acquired from the RBE. It is termed as the “RBE (k)”.

5.2 The WB Kernel Estimator

Racine and MacKinnon (2007) calculated the p -value for simple linear regression model but in the current work, it is computed for the FEM as

$$\hat{P}_{WBK} = 1 - \hat{F}_w(\hat{t}) = 1 - \frac{1}{B} \sum_{j=1}^B K(\hat{t}_{j(WB)}, \hat{t}, w),$$

where $\hat{F}_w(\hat{t}) = \frac{1}{B} \sum_{j=1}^B K(\hat{t}_{j(WB)}, \hat{t}, w)$, $\hat{F}_w(\hat{t})$ is the CDF, w is the bandwidth and \hat{t}_{WB} is the quasi- t test statistic obtained from the WBE. It is termed as the “WBE (k)”.

6. Empirical Results

For the empirical results, we used the same Monte Carlo scheme as used in some previous studies like Li and Stengos (1994), Roy (2002) and Aslam and Pasha (2007). The considered model is

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + \nu_{it} \quad i=1, \dots, n; \quad t=1, \dots, T, \quad \text{where } x_{it} = 0.5w_{i,t-1} + w_{it}.$$

We have used two data generating processes (DGPs)

- (i) w_{it} as $w_{it} \sim i.i.d.U(0, 2)$.
- (ii) w_{it} as $w_{it} \sim i.i.d.exp(N(0, 0.4^2))$, w_{it} follows lognormal distribution.

The values assigned to β_0 and β_1 are 5 and 0.5 respectively. ν_{it} and μ_i can be generated as $\nu_{it} \sim i.i.d.N(0, \sigma_v^2)$, $\mu_i \sim i.i.d.N(0, \omega_i)$ where $\omega_i = \omega(\bar{x}_i) = \delta^2 (1 + \lambda \bar{x}_i)^2$. It is supposed that heteroscedasticity is of additive form. Let the total variance γ_i and expected variance of μ_i is $\gamma_i = \omega_i + \sigma_v^2$ and $\bar{\omega}$, respectively. For comparison across different DGPs, the expected total variance is set to be $\bar{\omega} + \sigma_v^2 = 8$. The values of λ are 0, 1, 2 and 3, where 0 indicates homoscedastic unit-specific error and other shows different levels of heteroscedasticity for the fixed value of σ_v^2 and the values assigned to σ_v^2 are 2, 4 and 6. Increase in λ cause increase in degree of heteroscedasticity. Moreover, the value of $\bar{\omega}$ can be obtained using

different values of λ for each value of σ_v^2 and α is obtained using the additive heteroscedastic design specified above for given λ . Thus, the values of ω_i for each σ_v^2 under the four different values of λ are obtained. In our work, we have used normal kernel which has cited in Roy (2002) that is $k(x) = \sqrt{2\pi}^{-1/2} \exp(-\frac{x^2}{2})$,

where the value of smoothing parameter given in Racine and MacKinnon (2007) is $g = cB^{(-1/5)}$, where $c = 2.418, 1.575$ and 1.3167 for 10%, 5% and 1%, respectively. The simulations are 5000 with two sample size scheme

- I. Scheme I: $n = 50; T = 3; nT = 150$
- II. Scheme II: $n = 100; T = 3; nT = 300$

The number of bootstraps are 399. We use the following estimators

- a. The pooled OLS estimators
- b. The WGE
- c. The HCCME
- d. The Bootstrap- t estimators (RBE (t), WBE (t))
- e. The Bootstrap- p estimators (RBE (p), WBE (p))
- f. The Kernel Bootstrap estimator (RBE (k), WBE (k))

For study of the finite sample behaviour of the given estimators under heteroscedasticity, following measures are used

- Mean and MSE
- Empirical coverage
- Empirical size
- Empirical power

Empirical coverage of confidence interval and size of test are given in percentage form. Nominal coverage is taken to be 95% and empirical size is investigated at 1%, 5% and 10% nominal level of significance (LOS). Value of λ indicates level of heteroscedasticity for instance, 0 shows homoscedasticity, 1 reveals mild, 2 shows moderate and 3 indicates severe heteroscedasticity.

Table 1 (a) and 1 (b) contain the mean and MSE for Scheme I and II under DGPI, respectively. Intercept is excluded in the WG estimation (see Aslam, 2006), therefore it is not given in these tables and discussion is focused only on the slope estimates. Table 1 (a) shows that all the estimators are efficient for all homoscedastic and heteroscedastic cases but the OLSE is inefficient for smaller UTVH. For $\sigma_v^2 = 2$, the MSE of OLSE is more than twice of WGE. However, the OLSE performs equally well for $\sigma_v^2 = 4$. For larger UTVH ($\sigma_v^2 = 6$), performance of the OLSE is improved and it shows smaller MSE

than that of WGE. The similar behaviour of OLSE and WGE is noticed in Table 1 (b). For large samples, it is expected that performance of the OLSE is improved and it happens that for smaller UTVH, the OLSE performance is improved but still not better than the WGE.

Under DGPII, the results are given in Table 2 (a) and 2 (b) for given Scheme I and II of sample size, respectively. The obtained results do not differ from those generated under DGPI. The WGE outperformed OLSE for $\sigma_v^2 = 2$ but MSE of the OLSE is smaller than WGE for larger UTVH.

The average length and coverage is measured for study of finite sample properties of estimators under heteroscedasticity. The estimation of confidence interval is done as illustrated in (10). We have performed experiment for $\sigma_v^2 = 2$, though we have performed for $\sigma_v^2 = 4$ and 6 and found similar results. Under DGP I and II, empirical coverage is presented in Fig. 1 (a) and 1 (b), respectively. Under DGPI, the OLSE curve shows under-coverage while the curve of WGE is closer to the nominal coverage (95%). There is overlapping of the curves of HC3 and HC5. The curve of WBE (t) is closer to 95%. The similar behaviour is observed in Fig. 1 (b). Under DGPI, Table 3 (a) and 3 (b) carry empirical coverage and average length for Scheme I and II, respectively. Performance of the OLSE is not satisfactory in Table 3 (a) as it shows under-coverage for homoscedastic as well as heteroscedastic cases. The empirical coverage of WGE is closer to nominal coverage for all degrees of heteroscedasticity and it outperforms the OLSE. On the other side, it is noticed that the best empirical coverage among the HCCMEs are provided by HC4 and HC5. Among the bootstrap estimators, the WBE (t) exhibits best empirical coverage for heteroscedastic cases. This is verification of findings of Liu (1988) for linear regression models. Performance of the estimators in Table 3 (b) is similar to that observed in Table 3 (a). With the increment in sample size, there is no improvement in performance of the OLSE. Again, performance of the WBE (t) is remarkable and it remains an attractive choice.

Table 4 (a) and 4 (b) show empirical coverage and average interval length for DGP II for Scheme I and II, respectively. Table 4 (a) shows that the OLSE confidence interval does not exhibit good empirical coverage. For all degrees of heteroscedasticity ($\lambda = 0, 1, 2$ and 3), HC5 shows the best coverage among all the HCCMEs. It is noticed in Table 4 (b) that the WBE (t) confidence interval displays coverage that is closer to the nominal coverage (95%). The similar behaviour of estimators is observed in Table 4 (b) as observed in Table 4 (a).

In this work, hypothesis of interest is

$$H_0 : \beta_1 = 0.5, H_1 : \beta_1 \neq 0.5.$$

Empirical size can be measured according to (8). We have performed experiment for $\sigma_v^2 = 2$, though we have performed for $\sigma_v^2 = 4$ and 6 and found similar results. Empirical size is

presented in Fig. 2 (a) and 2 (b) under DGP I and II, respectively. The OLSE curve reveals high size distortion in Fig. 2 (a). The WGE curve is closer to nominal level (5%). The bootstrap and kernel bootstrap estimators curves are also closer to nominal size. The similar trend is observed in Fig. 2 (b). Table 5 (a) and 5 (b) carry empirical results under DGP I for the estimators described above for Scheme I and II, respectively. It is noticed in Table 5 (a) that empirical sizes executed by the OLSE are very poor and there is high size distortion under both homoscedastic ($\lambda = 0$) and heteroscedastic cases ($\lambda = 1, 2$ and 3). However, the OLSE gets improvement in performance. While performance of all the other estimators is remarkable in case of heteroscedasticity. The HC4 and HC5 provide show the best null rejection rate (NRR) among all the HCCMEs. The HC3 also shows admirable rejection rates. Good performance of HC4 verifies the results of Uchao et al. (2014). The HC5 carries out a sound performance from mild ($\lambda = 1$) to severe heteroscedasticity ($\lambda = 3$) which justifies our new design for the FEM. The RBE also performs well in terms of NRR. The WB (p) approach provides remarkable NRR and gives confirmation to the results of MacKinnon (2007) for cross-sectional data. Among all the bootstrap estimators, WB performs excellently and the similar results are observed in the literature like in the work of Cameron et al. (2008). Among all the considered estimators, the best NRR is provided by kernel bootstrap estimators in the presence of heteroscedasticity ($\lambda = 1, 2$ and 3) and it becomes an attractive choice for heteroscedastic PDM at all nominal LOS. This verifies findings of Racine and MacKinnon (2007) for linear regression models. Our proposed kernel bootstrap estimator performs the best among all the estimators under consideration in the presence of heteroscedasticity. It also justifies our proposal for the PDM. The results given in Table 5 (b) indicate that performance of all the estimators is analogous to those given in Table 3.9. Performance of the OLSE is expectedly very poor. The HC5 shows substantial performance under all cases of heteroscedasticity ($\lambda = 1, 2$ and 3). The WB (k) approach justifies new formulation by providing NRR closer to all the nominal LOS (1%, 5% and 10%). Table 6 (a) and 6 (b) show empirical sizes under DGP II for Scheme I and II, respectively. The results given in these tables striking the same mark as in DGP I.

Empirical power can be estimated according to (9). Under DGP I, Fig. 3 (a) shows empirical power curves based on all the estimators under consideration for Scheme I. For the case of homoscedasticity ($\lambda = 0$), it is observed that power curves of all the estimators are identical except that of OLSE which shows high power distortion for smaller UTVH ($\sigma_v^2 = 2$). The curves of all other estimators are identical and they perform equally well. It verifies reported results of Aslam (2006). Fig. 3 (b) displays empirical power curves for Scheme II under DGP I. It is expected that with the increment in sample size, the power curves tend to slumber. The curve of OLSE does not improve and shows power distortion for homoscedasticity ($\lambda = 0$) and heteroscedasticity ($\lambda = 1, 2$ and 3).

Fig. 4 (a) and 4 (b) show power curves for Scheme I and II, respectively under DGP II. The similar results are noticed under DGP II as given under DGP I.

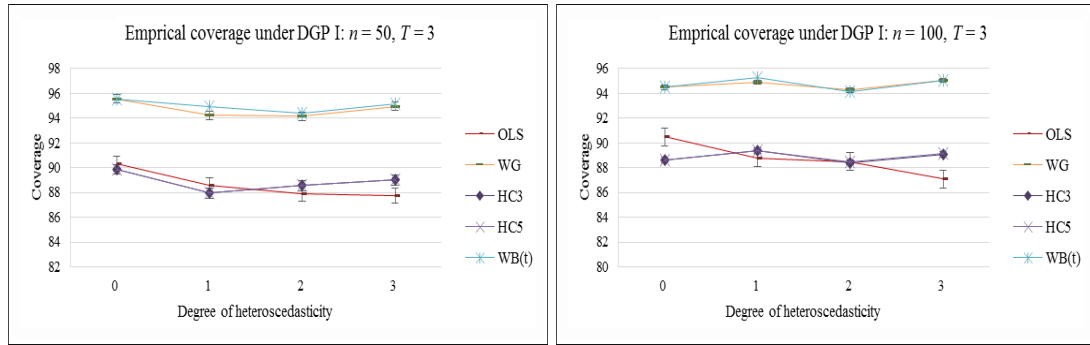


Figure 1 (a): Empirical coverage (%) of confidence interval for β_1 under DGP I

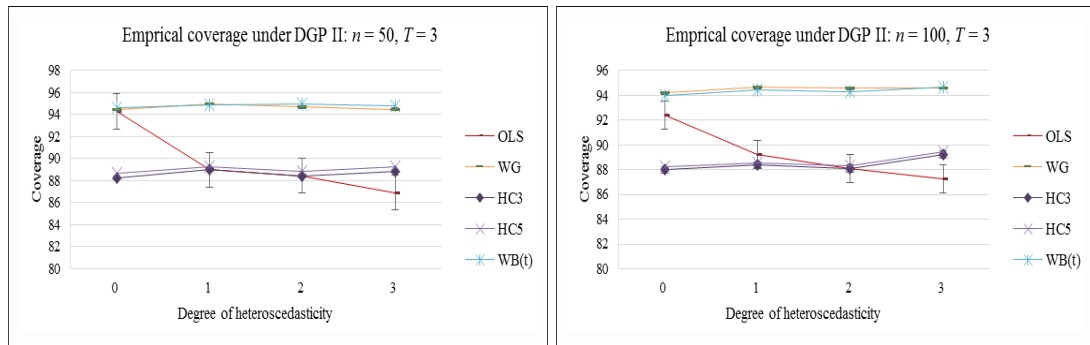


Figure 1 (b): Empirical coverage (%) of confidence interval for β_1 under DGP II

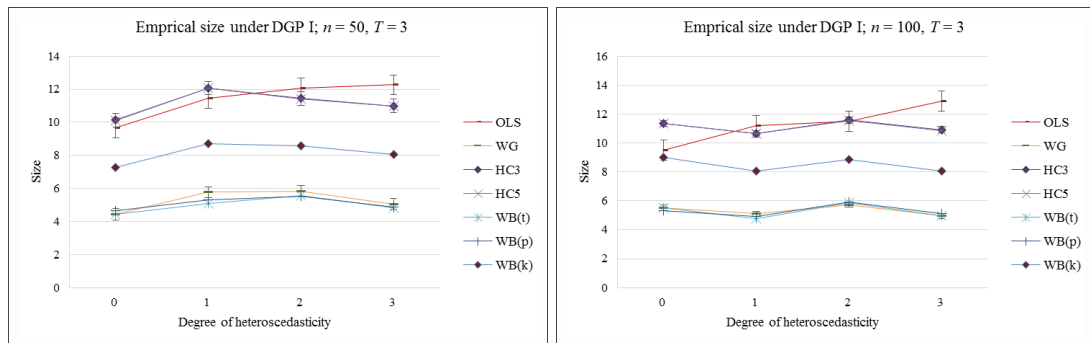


Figure 2 (a): Empirical size (%) for β_1 under DGP I

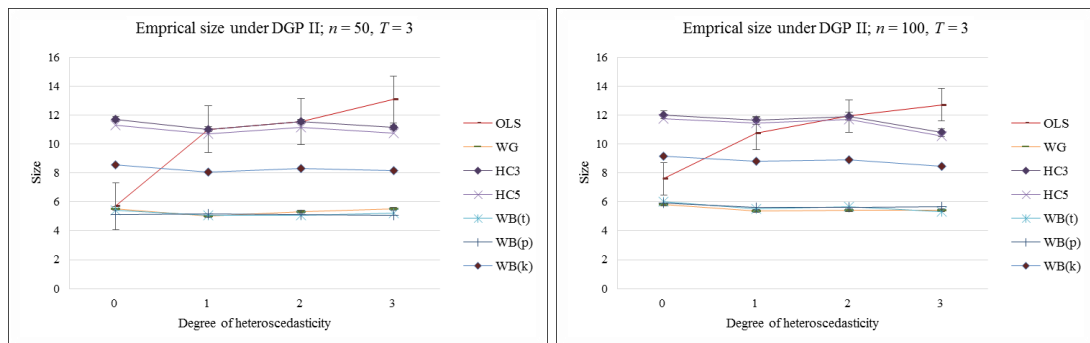


Figure 2 (b): Empirical size (%) for β_1 under DGP II

Table 1 (a): Mean and MSE under DGP1 ($n = 50, T = 3$)

$\lambda = 0$							coefficient
σ_v^2	2		4		6		β_0 (true value 5)
	Mean	MSE	Mean	MSE	Mean	MSE	
OLS	4.99	0.24	5.00	0.24	5.00	0.24	β_1 (true value 0.5)
OLS	0.50	0.11	0.50	0.11	0.50	0.11	
WG	0.50	0.05	0.50	0.11	0.49	0.16	
$\lambda = 1$							coefficient
OLS	4.99	0.25	4.99	0.25	5.00	0.25	β_0 (true value 5)
OLS	0.50	0.11	0.50	0.11	0.50	0.11	β_1 (true value 0.5)
WG	0.50	0.05	0.50	0.11	0.50	0.16	
$\lambda = 2$							coefficient
OLS	5.00	0.25	4.98	0.25	5.01	0.25	β_0 (true value 5)
OLS	0.49	0.12	0.51	0.11	0.50	0.11	β_1 (true value 0.5)
WG	0.49	0.05	0.51	0.11	0.50	0.16	
$\lambda = 3$							coefficient
OLS	5.00	0.26	5.01	0.25	5.00	0.25	β_0 (true value 5)
OLS	0.51	0.12	0.49	0.11	0.50	0.11	β_1 (true value 0.5)
WG	0.50	0.05	0.50	0.11	0.49	0.16	

Table 1 (b): Mean and MSE under DGP1 ($n = 100, T = 3$)

$\lambda = 0$							coefficient
σ_v^2	2		4		6		β_0 (true value 5)
	Mean	MSE	Mean	MSE	Mean	MSE	
OLS	5.01	0.12	5.00	0.12	5.00	0.12	β_1 (true value 0.5)
OLS	0.49	0.06	0.50	0.06	0.50	0.06	
WG	0.50	0.03	0.50	0.05	0.50	0.08	
$\lambda = 1$							coefficient
OLS	5.00	0.13	5.00	0.12	5.00	0.12	β_0 (true value 5)
OLS	0.50	0.06	0.50	0.06	0.50	0.06	β_1 (true value 0.5)
WG	0.50	0.03	0.50	0.05	0.50	0.08	
$\lambda = 2$							coefficient
OLS	5.00	0.13	5.00	0.13	5.00	0.12	β_0 (true value 5)
OLS	0.49	0.06	0.50	0.06	0.50	0.06	β_1 (true value 0.5)
WG	0.50	0.03	0.50	0.05	0.50	0.08	
$\lambda = 3$							coefficient
OLS	4.99	0.13	5.00	0.13	5.00	0.12	β_0 (true value 5)
OLS	0.50	0.06	0.50	0.06	0.50	0.06	β_1 (true value 0.5)
WG	0.50	0.03	0.50	0.05	0.50	0.08	

Table 2 (a): Mean and MSE under DGP1I ($n = 50, T = 3$)

$\lambda = 0$							coefficient
σ_v^2	2		4		6		β_0 (true value 5)
	Mean	MSE	Mean	MSE	Mean	MSE	
OLS	5.01	0.49	5.00	0.49	5.02	0.49	
OLS	0.49	0.21	0.49	0.21	0.49	0.21	β_1 (true value 0.5)
WG	0.49	0.09	0.50	0.19	0.50	0.28	
$\lambda = 1$							coefficient
OLS	4.99	0.49	5.01	0.49	5.01	0.50	β_0 (true value 5)
OLS	0.51	0.21	0.50	0.21	0.50	0.21	β_1 (true value 0.5)
WG	0.50	0.09	0.51	0.19	0.50	0.28	
$\lambda = 2$							coefficient
OLS	5.00	0.50	5.00	0.50	5.00	0.49	β_0 (true value 5)
OLS	0.50	0.21	0.50	0.21	0.51	0.21	β_1 (true value 0.5)
WG	0.50	0.09	0.50	0.19	0.51	0.28	
$\lambda = 3$							coefficient
OLS	5.02	0.50	5.00	0.50	5.01	0.50	β_0 (true value 5)
OLS	0.49	0.22	0.50	0.21	0.50	0.21	β_1 (true value 0.5)
WG	0.49	0.09	0.50	0.19	0.51	0.28	

Table 2 (b): Mean and MSE under DGP1I ($n = 100, T = 3$)

$\lambda = 0$							coefficient
σ_v^2	2		4		6		β_0 (true value 5)
	Mean	MSE	Mean	MSE	Mean	MSE	
OLS	5.00	0.25	5.01	0.25	5.00	0.25	
OLS	0.50	0.11	0.49	0.11	0.50	0.11	β_1 (true value 0.5)
WG	0.50	0.05	0.49	0.09	0.50	0.14	
$\lambda = 1$							coefficient
OLS	4.99	0.25	5.01	0.25	5.00	0.25	β_1 (true value 0.5)
OLS	0.51	0.11	0.49	0.11	0.50	0.11	
WG	0.50	0.05	0.49	0.09	0.51	0.14	
$\lambda = 2$							coefficient
OLS	5.01	0.25	5.01	0.25	4.99	0.25	β_1 (true value 0.5)
OLS	0.49	0.11	0.50	0.11	0.51	0.11	
WG	0.50	0.05	0.50	0.09	0.50	0.14	
$\lambda = 3$							coefficient
OLS	4.99	0.25	5.00	0.25	5.01	0.25	β_1 (true value 0.5)
OLS	0.51	0.11	0.50	0.11	0.49	0.11	
WG	0.50	0.05	0.50	0.09	0.50	0.14	

Table 3 (a): 95% Confidence Interval: Coverage and Length ($n = 50, T = 3$) under DGP I

$\sigma_v^2 = 2$								
$\lambda = 0$			$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
Coverage	Length		Coverage	Length	Coverage	Length	Coverage	Length
OLS	90.34	0.91	88.56	0.91	87.92	0.91	87.72	0.91
WG	95.56	0.73	94.22	0.73	94.16	0.73	94.94	0.73
HC0	89.00	0.73	86.90	0.73	87.84	0.74	88.40	0.73
HC3	89.86	0.75	87.94	0.75	88.56	0.75	89.00	0.75
HC4	89.88	0.75	87.94	0.75	88.60	0.75	89.02	0.75
HC5	89.88	0.75	87.94	0.75	88.60	0.75	89.02	0.75
RB(t)	95.74	0.92	94.54	0.93	94.62	0.93	95.12	0.93
WB(t)	95.56	0.93	94.92	0.93	94.42	0.93	95.16	0.93

Table 3 (b): 95% Confidence Interval: Coverage and Length ($n = 100, T = 3$) under DGP I

$\sigma_v^2 = 2$								
$\lambda = 0$			$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
Coverage	Length		Coverage	Length	Coverage	Length	Coverage	Length
OLS	90.48	0.92	88.78	0.93	88.50	0.94	87.08	0.94
WG	94.48	0.64	94.90	0.64	94.26	0.64	95.04	0.64
HC0	88.26	0.52	88.98	0.52	88.08	0.52	88.68	0.52
HC3	88.64	0.53	89.36	0.53	88.42	0.53	89.08	0.52
HC4	88.64	0.53	89.34	0.53	88.44	0.53	89.14	0.53
HC5	88.64	0.53	89.34	0.53	88.44	0.53	89.14	0.53
RB(t)	94.54	0.66	95.20	0.66	94.36	0.66	95.24	0.66
WB(t)	94.50	0.66	95.24	0.66	94.12	0.66	95.04	0.66

Table 4 (a): 95% Confidence Interval: Coverage and Length ($n = 50, T = 3$) under DGP II

$\sigma_v^2 = 2$								
$\lambda = 0$			$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
Coverage	Length		Coverage	Length	Coverage	Length	Coverage	Length
OLS	94.30	1.79	88.98	1.80	88.46	1.81	86.92	1.81
WG	94.46	1.19	94.96	1.19	94.68	1.19	94.48	1.19
HC0	87.18	0.95	88.08	0.95	87.60	0.95	87.82	0.95
HC3	88.28	0.98	88.98	0.97	88.44	0.98	88.84	0.98
HC4	88.70	0.99	89.26	0.99	88.82	0.99	89.22	1.00
HC5	88.70	1.00	89.28	0.99	88.84	0.99	89.24	1.00
RB(t)	94.64	1.22	95.26	1.22	95.14	1.21	94.94	1.22
WB(t)	94.60	1.24	94.92	1.23	94.94	1.23	94.80	1.24

Table 4 (b): 95% Confidence Interval: Coverage and Length ($n = 100, T = 3$) under DGP II

$\sigma_v^2 = 2$								
$\lambda = 0$			$\lambda = 1$		$\lambda = 2$		$\lambda = 3$	
Coverage	Length		Coverage	Length	Coverage	Length	Coverage	Length
OLS	92.40	1.27	89.26	1.28	88.06	1.29	87.28	1.29
WG	94.20	0.84	94.64	0.84	94.56	0.84	94.58	0.84
HC0	87.32	0.68	87.90	0.68	87.64	0.68	88.78	0.68
HC3	87.98	0.69	88.36	0.69	88.08	0.69	89.22	0.69
HC4	88.22	0.70	88.52	0.69	88.28	0.69	89.44	0.70
HC5	88.24	0.70	88.56	0.70	88.30	0.69	89.46	0.70
RB(t)	94.36	0.86	94.92	0.86	94.52	0.86	94.80	0.86
WB(t)	94.00	0.86	94.46	0.86	94.32	0.86	94.68	0.86

Table 5 (a): NRR of quasi- t test for $n = 50$, $T = 3$ under DGP I

$\sigma_v^2 = 2$												
α	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
σ_v^2	$\lambda = 0$			$\lambda = 1$			$\lambda = 2$			$\lambda = 3$		
OLS	2.88	9.66	16.56	3.58	11.44	18.46	4.24	12.08	19.02	4.08	12.28	20.06
WG	0.90	4.44	9.70	1.06	5.78	11.12	1.42	5.84	10.66	1.14	5.06	9.98
HC0	3.64	11.00	18.30	4.40	13.10	19.68	4.98	12.16	19.70	4.04	11.60	18.70
HC3	3.14	10.14	17.38	4.00	12.06	18.70	4.46	11.44	18.68	3.64	11.00	17.74
HC4	3.14	10.12	17.34	3.98	12.06	18.62	4.44	11.40	18.64	3.64	10.98	17.74
HC5	3.14	10.12	17.34	3.98	12.06	18.62	4.44	11.40	18.64	3.64	10.98	17.74
RB(t)	0.82	4.26	9.18	1.06	5.46	10.88	1.26	5.38	10.32	0.68	4.88	9.66
WB(t)	0.88	4.44	8.92	0.90	5.08	10.56	1.16	5.58	10.28	1.06	4.84	9.82
RB(p)	0.76	4.10	9.30	0.88	5.36	10.84	1.10	5.58	10.36	0.86	4.88	9.70
WB(p)	0.98	4.64	8.84	0.90	5.30	10.74	1.36	5.52	10.36	1.08	4.88	9.74
RB(k)	1.20	7.62	10.24	1.48	8.92	11.60	1.92	9.12	11.18	1.22	7.90	10.54
WB(k)	1.22	7.28	9.22	1.40	8.70	11.32	1.62	8.58	11.02	1.36	8.04	10.40

Table 5 (b): NRR of quasi- t test for $n = 100$, $T = 3$ under DGP I

$\sigma_v^2 = 2$												
α	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
σ_v^2	$\lambda = 0$			$\lambda = 1$			$\lambda = 2$			$\lambda = 3$		
OLS	2.92	9.52	15.90	3.32	11.22	18.10	3.88	11.50	18.84	4.60	12.92	20.32
WG	1.16	5.52	10.46	1.16	5.10	9.76	1.18	5.74	10.70	0.90	4.96	10.04
HC0	4.18	11.74	18.48	3.78	11.02	17.92	4.72	11.92	19.42	3.84	11.32	19.24
HC3	3.90	11.36	18.10	3.62	10.64	17.36	4.50	11.58	18.76	3.66	10.92	18.58
HC4	3.88	11.36	18.06	3.62	10.66	17.32	4.48	11.56	18.78	3.66	10.86	18.58
HC5	3.88	11.36	18.06	3.62	10.66	17.32	4.48	11.56	18.78	3.66	10.86	18.58
RB(t)	1.00	5.46	9.92	1.00	4.80	9.68	1.24	5.64	10.18	0.88	4.76	9.96
WB(t)	1.06	5.50	10.28	1.30	4.76	9.44	1.22	5.88	10.34	1.08	4.96	10.06
RB(p)	0.94	5.40	10.30	1.00	4.76	9.42	1.28	5.66	10.52	0.84	4.90	9.56
WB(p)	1.12	5.32	10.46	1.34	4.92	9.60	1.16	5.90	10.20	1.04	5.14	9.86
RB(k)	1.44	8.74	11.00	1.42	8.14	10.20	1.56	9.02	11.12	1.20	8.32	10.84
WB(k)	1.54	9.02	11.00	1.56	8.04	10.10	1.74	8.86	10.88	1.28	8.06	10.62

Table 6 (a): NRR of quasi- t test for $n = 50$, $T = 3$ under DGP II

$\sigma_v^2 = 2$												
α	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
σ_v^2	$\lambda = 0$			$\lambda = 1$			$\lambda = 2$			$\lambda = 3$		
OLS	1.08	5.70	11.36	3.22	11.02	17.76	3.82	11.54	19.22	4.36	13.08	20.74
WG	1.10	5.54	10.60	1.16	5.04	10.00	1.08	5.32	10.40	1.26	5.52	10.32
HC0	4.80	12.82	19.76	4.52	11.92	18.70	4.54	12.40	19.68	4.62	12.18	19.58
HC3	4.32	11.72	18.54	4.02	11.02	17.36	4.04	11.56	18.50	4.10	11.16	18.16
HC4	4.18	11.30	18.22	3.84	10.74	16.84	3.86	11.18	17.98	3.86	10.78	17.50
HC5	4.18	11.30	18.22	3.84	10.72	16.76	3.82	11.16	17.94	3.84	10.76	17.42
RB(t)	1.16	5.36	10.60	0.98	4.74	9.54	0.84	4.86	10.30	1.06	5.06	10.08
WB(t)	1.32	5.40	10.18	1.16	5.08	9.76	1.00	5.06	9.96	1.34	5.20	9.76
RB(p)	1.00	5.32	10.34	1.08	4.92	9.78	0.94	4.84	10.04	1.00	5.14	10.06
WB(p)	1.16	5.10	9.96	1.14	5.16	9.64	1.02	5.10	10.00	1.54	5.08	9.74
RB(k)	1.44	8.74	11.20	1.58	8.50	10.28	1.30	8.78	10.72	1.50	8.48	10.60
WB(k)	1.52	8.54	10.70	1.66	8.08	10.26	1.28	8.30	10.56	1.68	8.16	10.22

Table 6 (b): NRR of quasi- t test for $n = 100, T = 3$ under DGP II

α	$\sigma_v^2 = 2$											
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
σ_v^2	$\lambda = 0$			$\lambda = 1$			$\lambda = 2$			$\lambda = 3$		
OLS	1.84	7.60	14.28	3.52	10.74	17.94	3.96	11.94	18.74	4.60	12.72	19.50
WG	1.02	5.80	11.14	0.82	5.36	10.96	1.08	5.44	10.58	1.12	5.42	9.86
HC0	4.62	12.68	19.44	4.06	12.10	19.48	4.60	12.36	19.22	4.26	11.22	17.46
HC3	4.20	12.02	18.82	3.64	11.64	18.94	4.30	11.92	18.40	3.90	10.78	16.70
HC4	4.08	11.78	18.38	3.56	11.48	18.56	4.22	11.72	17.98	3.72	10.56	16.50
HC5	4.08	11.76	18.38	3.56	11.44	18.50	4.22	11.70	17.94	3.70	10.54	16.48
RB(t)	0.92	5.64	10.88	0.58	5.08	10.46	1.04	5.48	9.98	0.86	5.20	9.30
WB(t)	1.04	6.00	10.88	1.00	5.54	10.82	1.34	5.68	10.64	0.98	5.32	9.82
RB(p)	0.82	5.56	10.84	0.68	5.24	10.60	0.98	5.36	9.94	0.90	5.24	9.52
WB(p)	1.20	5.94	10.72	0.84	5.60	10.70	1.24	5.64	10.72	1.12	5.68	9.64
RB(k)	1.36	9.64	11.70	1.12	9.00	11.24	1.56	8.62	10.62	1.32	8.38	10.10
WB(k)	1.62	9.18	11.60	1.30	8.82	11.30	1.70	8.92	11.32	1.44	8.46	10.12

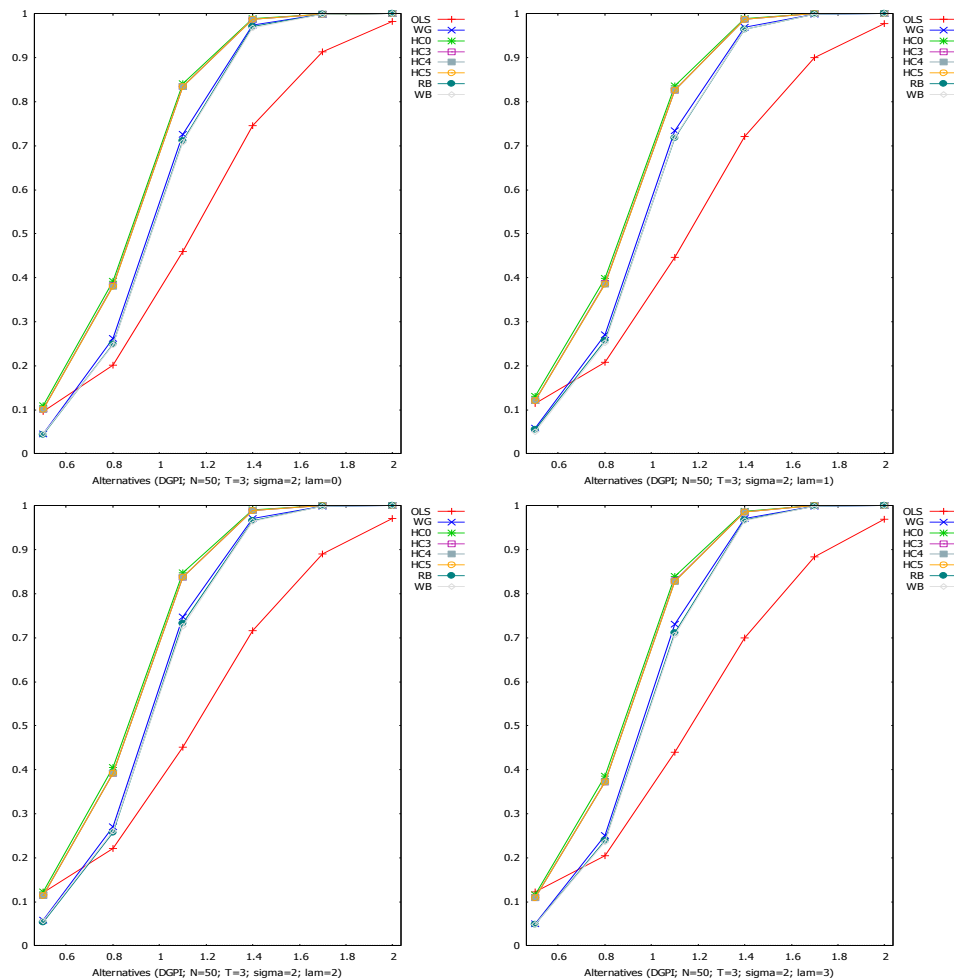


Figure 3 (a) Empirical power of test at 5% LOS (DGP I; $n = 50, T = 3$)

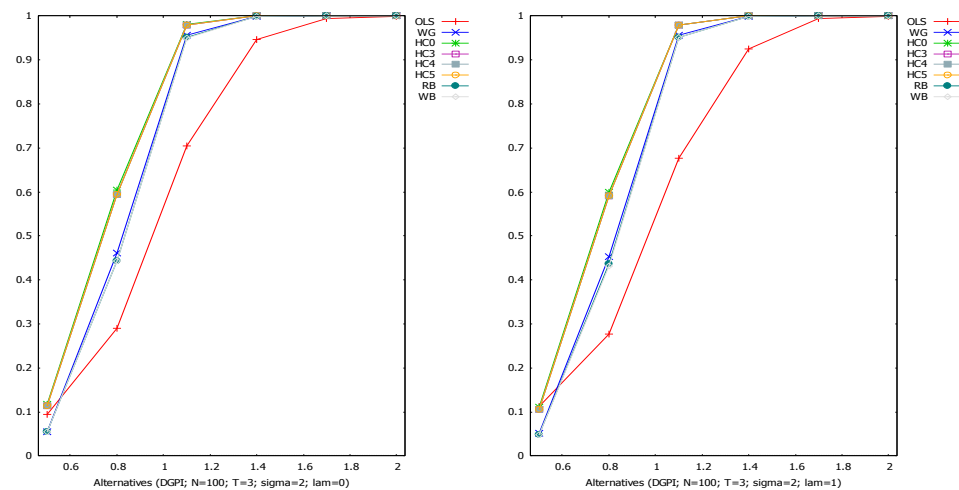


Figure 3 (b) Empirical power of test at 5% LOS (DGP I; $n = 100$, $T = 3$)
Continued.....

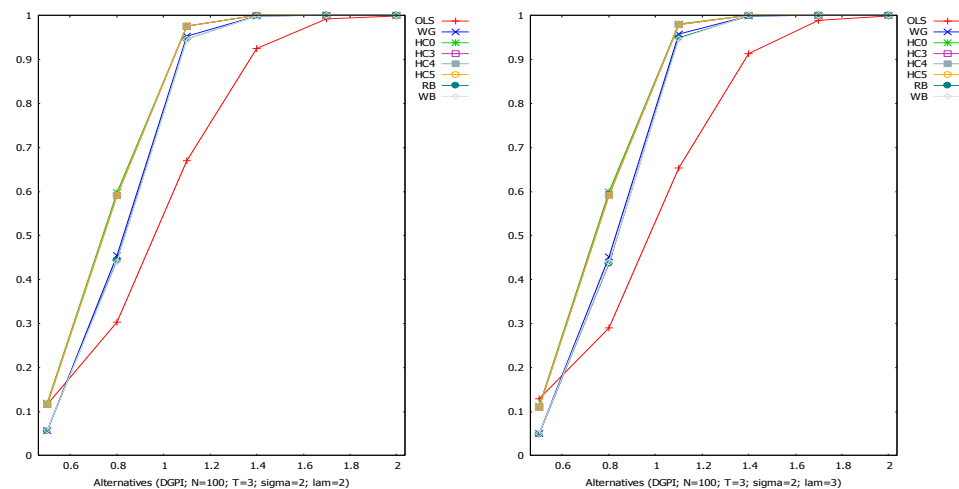


Figure 3 (b) Empirical power of test at 5% LOS (DGP I; $n = 100$, $T = 3$)

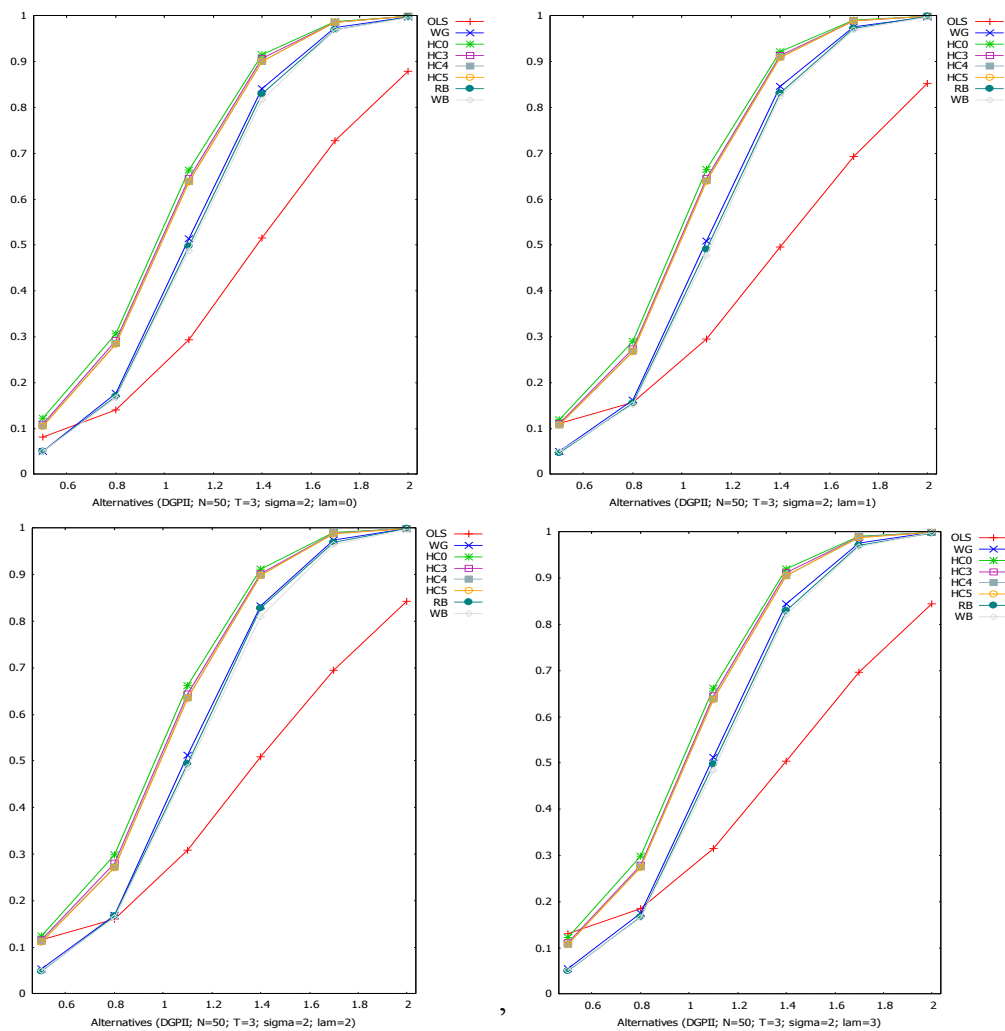


Figure 4 (a) Empirical power of test at 5% LOS (DGP II; $n = 50$, $T = 3$)

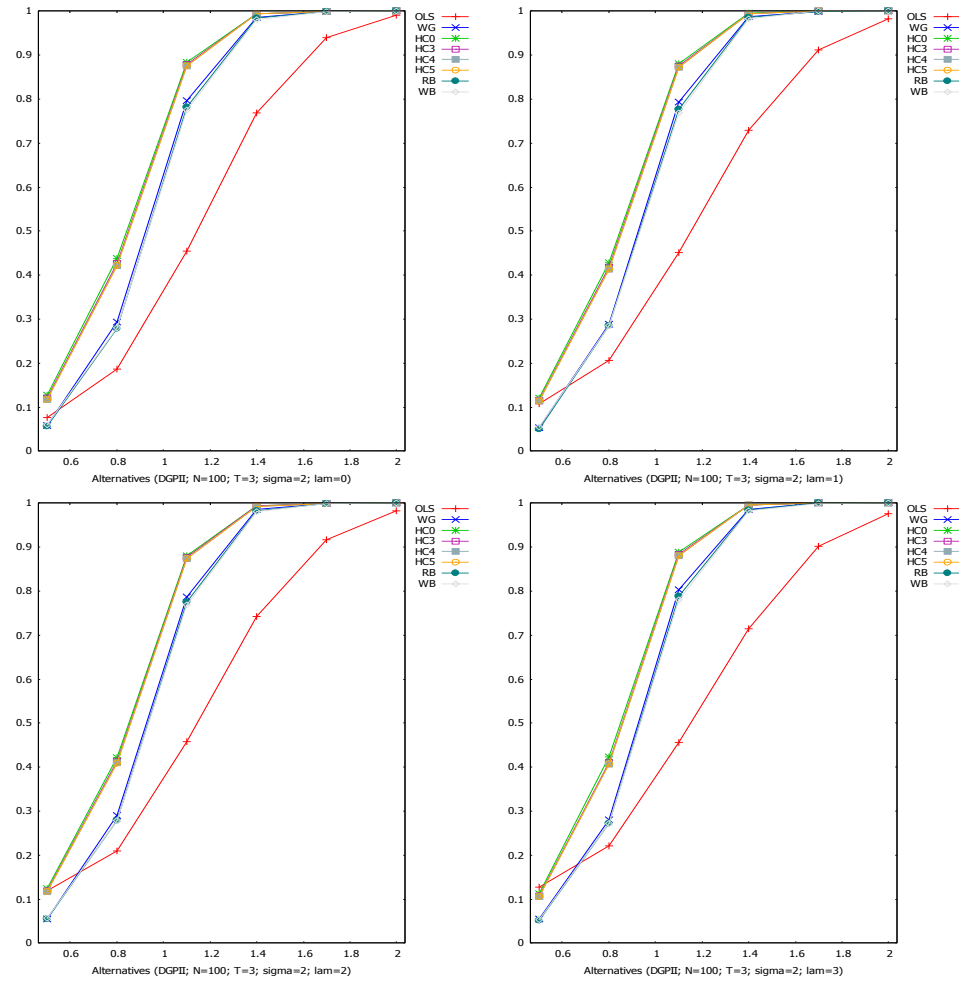


Figure 4 (b) Empirical power of test at 5% LOS (DGP II; $n = 100$, $T = 3$)

7. Conclusion

We have considered the heteroscedastic FEM and tried to draw robust inference using quasi t -test based on the HCCME and bootstrap estimators. It is clear from the results given above that performance of the HC5 is the best among all the HCCMEs in the presence of heteroscedasticity and leverage data points and the similar results can be seen by Cribari-Neto et al. (2007) for linear regression model. Performance of the kernel bootstrap estimators is better than the bootstrap estimators in respect of coverage, empirical sizes and empirical power. It is concluded that the WB (k) outperforms all the HCCMEs and bootstrap estimators in the presence of severe heteroscedasticity. It is justification of our new formulation for the PDM.

References

1. Arellano, M. (1987). Computing Robust Standard Errors for Within-Group Estimators. *Oxford Bulletin of Economics and Statistics*, 49, 431-434.
2. Aslam, M. and Pasha, G. R. (2007). Evaluation of Error in Rejection Probability for Adaptive and Heteroscedasticity Consistent Covariance Matrix Estimators.

- Proceedings of the 3rd National Conference on Statistical Sciences. Lahore, Pakistan. Islamic Countries Society of Statistical Sciences.
3. Beran, R. (1986). Prepivoting to Reduce Level Error of Confidence Sets. *Biometrika*, 74, 457-468.
 4. Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008). Bootstrap-based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, 90, 414-427.
 5. Cribari-Neto, F. and Lima, M. G. A. (2009). Heteroskedasticity-Consistent Interval Estimators. *Journal of Statistical Computation and Simulation*, 79, 787-803.
 6. Cribari-Neto, F., Ferrari, Silvia L. P. and Vasconcellos, P. L K. (2007). Inference under Heteroscedasticity and Leveraged Data. *Communications in Statistics Theory and Methods*. 36, 1877–1888.
 7. Eicker, F. (1963). Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions. *Annals of Mathematical Statistics*, 34, 447-456.
 8. Greene, W.H. (2003). *Econometric Analysis*. 5th Edition. Prentice Hall, Upper Saddle River, New Jersey.
 9. Halawa, A. M. and El Bassiouni, M. Y. (2000). Tests of Regression Coefficients under Ridge Regression Models. *Journal of Statistical Computation and Simulation*, 65, 341-356.
 10. Li, Q. And Stengos, T. (1994). Adaptive Estimation in the Panel Data Error Component Model with Heteroscedasticity of Unknown Form. *International Economic Review*, 35, 981-1000.
 11. Liu, R. Y. (1988). Bootstrap Procedures under Some Non-I.I.D. Models. *Annals of Statistics*, 16, 1696–1708.
 12. MacKinnon J.G. (2007). Bootstrap Hypothesis Testing. Working Paper 1127.
 13. Mackinnon, J. G. and Webb, M, D. (2013). Wild Bootstrap Inference for Wildly Different Cluster Sizes. Queen’s Economics Department Working Paper 1314.
 14. Mazodier, P. and A. Trognon (1978). Heteroscedasticity and Stratification in Error Components Models. *Annales de l’INSEE*, 30–31, 451–482.
 15. Racine, J. S. and Mackinnon, J. G. (2007). Inference via Kernel Smoothing of Bootstrap P Value. *Computational Statistics and Data Analysis*, 51, 5949– 5957.
 16. Roy, N. (2002). Is adaptive estimation useful for panel models with heteroscedasticity in the individual specific error component? Some Monte Carlo evidence. *Econometric Reviews*, 21, 189-203.
 17. Uchôa, C. F.A., Cribari-Neto, F. and Menezes, T. A. (2014). Testing Inference in Heteroscedastic Fixed Effects Models. *European Journal of Operational Research*, 660-670.
 18. White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroscedasticity. *Econometrica*, 48, 817–838.
 19. Wu, C. F. J. (1986). Jackknife, Bootstrap and other Resampling Methods in Statistics. *Annals of Statistics*, 14, 1261-1350.