FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 4 No. 4, December, 2020, pp 561 – 569
DOI: https://doi.org/10.33003/fjs-2020-0404-689

ROBUST PARAMETER ESTIMATION FOR RANDOM EFFECT PANEL DATA MODEL IN THE PRESENCE OF HETEROSCEDASTICITY AND INFLUENTIAL OBSERVATIONS

***Muhammad Sani, Shamsuddeen Suleiman, Baoku Ismail G**

Department of Mathematical Sciences, Federal University Dutsin-Ma, Katsina State, Nigeria

*Corresponding Author's email: **sanimksoro@gmail.com**

ABSTRACT

Panel data estimators can strongly be biased and inconsistent in the presence of heteroscedasticity and anomalous observations called influential observations (IOs) in Random effect (RE) panel data model. The existing methods (LWS, WLSF, WLSDRGP) address only the problem of IO but fail to remedy the combine problem of heteroscedasticity and IOs. Therefore, in this research we develop a method that will remedy the combine problem of heteroscedasticity and IOs based on robust heteroscedasticity consistent covariance matrix (RHCCM) estimator and fast improvised influential distance (FIID) weighting method denoted by WLSFIID. The simulation and numerical evidences show that our proposed estimation method is more efficient than the existing methods by providing smallest bias, and smallest standard error of HC4 and HC5.

Keywords: Heteroscedasticity; Influential observations; Panel Data; Random Effect Model; RHCCM; Weighted Least Square

INTRODUCTION

Panel data is a data that has two dimensionalities (time series and cross-section dimensions). It is usually used in the field of economics and finance to analyze a number of questions that cannot possibly be analyze using cross-sectional or time series analysis (Baltagi, 2005). The panel data model given by Muhammad et al. (2019) is;

$$y_{it} = u_i + x'_{it}\beta + \varepsilon_{it,} \quad i = 1,2,\dots,n \text{ and } t = 1,2,\dots,T \quad (1)$$

where, y_{it} are the response variables, x_{it} is the kth explanatory variables, u_i is the unobserved time-invariant effects and ε_{it} is the error term (idiosyncratic error) that is assumed to be normal, uncorrelated across individual units and time.

Panel data estimators can strongly be biased and inconsistent in the presence of heteroscedasticity and influential observations (Bramati and Croux, 2007). Rousseeuw and Zomeren (1990) pointed out that in panel data influential observations usually occur in y-axis (vertical outlier) or in x-axis (high leverage point). The most dangerous type of influential observation (IO) is the high leverage points (HLPs). There are two major models for analyzing a panel data, which are random effect (RE) and fixed effect (FE). The major difference between these two models is the assumption of the time invariant effect.

The random effect (RE) model estimation technique is the same as FE model estimation except in the data transformation. In RE model there is no correlation between unobserved time invariant effect and regressor i.e. $cov(x_{it}, u_i) = 0$. The RE model uses partially demeaned transformation instead of demeaned transformation used in FE model. The estimation technique in RE model is to apply OLS to the partially demeaned transformed data (Crowder and Hand, 1990).

Indeed, the OLS approach is known to be very sensitive to HLPs which causes bias in the parameter estimates. The problem of heteroscedasticity was addressed by many researchers in linear regression (Habshah et al., 2017; Furno, 1996; Rana et al., 2012). In recent years, researchers developed robust estimators in panel data regression models in order to provide more consistent and efficient estimator (Muhammad et al., 2019; Maronna et al., 2006; Bramati and Croux, 2007; Baltagi, 2008; Baltagi et al., 2009; Verardi and Wagner, 2011; Mazlina and Habshah, 2015, Habshah and Sani, 2018). Nevertheless, their techniques do not take into consideration the combined problem of heteroscedasticity and IO.

Recently, Visek (2015) used the least weighted squares (LWS) to estimate the parameters of the FE and RE models in panel data by employing classical centering method (mean centering) to transform the data and apply LWS, where the weight is defined by the residual order statistic. The limitation of this method is that, when there exist heteroscedasticity of unknown structure it is inefficient and produces large variances which lead to inconsistency of the Variance-covariance matrix. Moreover, the mean centering used by Visek (2015) is easily affected by the presence of IO. These shortcomings motivated us to propose a new estimation technique for RE panel data model based on fast improvised influential distance (FIID) of Habshah et al., (2021) and robust heteroscedasticity consistent covariance matrix (RHCCM).

In this paper, we used a numerical data and Monte Carlo simulation studies to assess the performance of the proposed estimation technique (WLS_{FIID}) and the existing methods;

Least weighted squares (LWS), Furno's weighting method (F) and Diagnostic Robust Generalize Potential's weighting method (DRGP) for RE panel data model.

MATERIALS AND METHODS

The partially demeaned centering transforms a panel data within each time series by subtracting some component of the average in each time series. As mentioned earlier, RE model assumed that $cov(x_{it}, u_i) = 0$ and u_i is part of the error term for all i = 1, 2, ...,n, t = 1, 2, ...,T. i.e $\tau_{it} = u_i + \varepsilon_{it}$, $E(\tau_{it}) = 0$, $E(\tau_{it}^2) = \sigma_u^2 + \sigma_\varepsilon^2$ and $E(\tau_{it}, \tau_{is}) = var(u_i) = \sigma_u^2$ for all $t \neq s$ (see Judge et al., 1985; Baltagi, 2001).

Proposed Demeaned Centering based on MM estimator

The MM centering for RE model based on partially demeaned transformation has the same procedure as that of partially demeaned transformation based on OLS method. The only difference is that, the demeaned transformed data within each time series by MM centering which is now given as:

$$y_{it} - \hat{\theta}.\hat{\mu}_{mm}\{y_{it}\} = \left(x_{it}^{(j)} - \hat{\theta}.\hat{\mu}_{mm}\left\{x_{it}^{(j)}\right\}\right)\beta + \varepsilon_{it} \quad (2)$$

for $1 \leq t \leq T$, $1 \leq i \leq n$ and $1 \leq j \leq k$, where $x^{(j)}$ is the j^{th} explanatory variables.

1) Proposed Robust RE Estimation Method

The new estimation method was design to remedy the effect of heteroscedasticity of unknown structure and IO based on robust HCCM estimator and detection measure (FIID). The algorithm of the proposed technique is summarized as follows:

Step 1. Use partially demeaned centering based on MM-centering method to transform the data.

Step 2. Compute the weight function w_i based on FIID method.

Step 3. Fit a weighted least square (WLS) to the transformed data in Step1 using weight w_i obtained in Step2, calculate the residuals (r_i) and coefficient of estimates.

Step 4. Compute the RHCCM estimator using the residuals(r_i) obtained in Step3.

a) Monte Carlo Simulation Study

We employed a simulation technique of Visek (2015) and Lima et. al. (2009) to assess the performances of the new proposed weighting method (WLS_{FIID}) in RE panel data model. Let consider the following RE panel data model,

$$y_{it} = u_i + x_{it}'\beta + \varepsilon_{it}, \; i = 1,2, \dots, n \quad \text{and} \quad t = 1,2, \dots, T \quad (3)$$

Three explanatory variables ($x_{it1}, x_{it2}, x_{it3}$) and u_i were generated from normal distribution. We set the true parameters $b_0 = b_1 = b_2 = b_3 = 1$, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$. Three sample sizes $n =$ 5, 10 and 15 with the corresponding $t =$ 10, 15 and 20 were replicated twice to form $n =$ 10, 20, 30 and $t =$ 20, 30, 40 respectively, in order to create heteroscedasticity. The degree of heteroscedasticity is assess by $\lambda = \max(\sigma_\varepsilon^2)/\min(\sigma_\varepsilon^2)$. The skedastic function was set based on Lima et al., (2009) as $\sigma_\varepsilon^2 = \exp\{c_1 x_{it1}\}$, where $c_1 = 0.75$ which gives the value of $\lambda \approx 42.8$ and will be constant for all the sample sizes considered. The value of λ indicate the level of heteroscedasticity present in the data, whereby for homoscedasticity, $\lambda = 1$. Regular data points in both response and explanatory variables were replaced with data points generated from k-variate normal distribution $N(10,1)$ at 0%, 5% and 10% contamination level for all the sample sizes considered at an average of 1000 replications.

Artificial heteroscedastic RE panel data

An artificial heteroscedastic RE panel data set with n=6 and t=20 number of observations was generated. The independent and response variable were generated from $N(10,1)$ and $y_{it} = 1 + x_{it1} + x_{it2} + x_{it3} + \varepsilon_{it}$, respectively. The heteroscedasticity was created as in the Monte Carlo Simulation.

RESULTS AND DISCUSSION

Tables 1-3 shows the performance of the proposed method (WLS_{FIID}) and the existing methods (LWS, WLS_F and WLS_{DRGP}), in a simulated heteroscedastic random effect panel data with different sample sizes and IO contamination level. The results show that the new proposed method WLS_{FIID} is more efficient than the existing methods, by providing less standard error of the estimates, less variances of HC4 and HC5, and also produce the coefficient of estimates that is closed to the true parameter coefficient. Justification using standard error of the estimates here is inappropriate and inefficient, as the form of heteroscedasticity is unknown. Therefore, the estimation will be based on HC4 and HC5 method employed. Figure 1 clearly shows the performance of all the methods at different sample sizes, where WLS_{FIID} is the best followed by WLS_{DRGP}, WLS_F, and finally LWS.

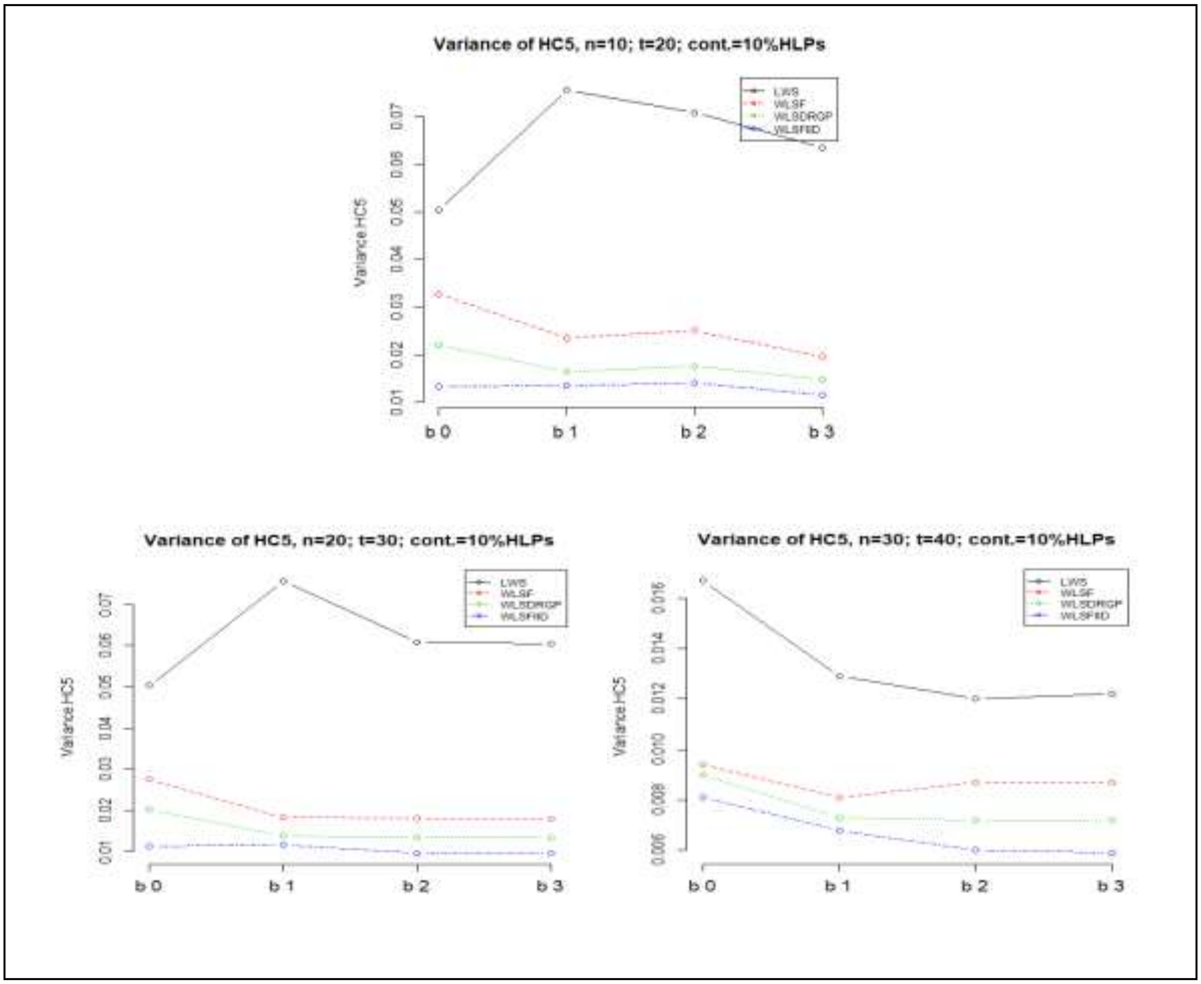

Figure 1: Plot of variance of HC5 for 10% HLPs contamination level with different sample sizes

Table 1: Simulation result of RE panel data estimates for n = 10, t=20

Con. Level	Estimator		Coeff. of Estimates	Standard error of Estimates	Variance HC4	Variance HC5
0 % HLPs	LWS	$\hat{\beta}_0$	1.0004	0.1180	0.0148	0.0148
		$\hat{\beta}_1$	0.9999	0.1075	0.0214	0.0214
		$\hat{\beta}_2$	0.9918	0.1081	0.0115	0.0115
		$\hat{\beta}_3$	1.0000	0.1082	0.0122	0.0122
	WLS_F	$\hat{\beta}_0$	1.0003	0.1039	0.0111	0.0111
		$\hat{\beta}_1$	0.9809	0.0979	0.0126	0.0126
		$\hat{\beta}_2$	0.9888	0.0976	0.0090	0.0090
		$\hat{\beta}_3$	0.9920	0.0981	0.0091	0.0091
	WLS_{DRGP}	$\hat{\beta}_0$	1.0002	0.1041	0.0112	0.0112
		$\hat{\beta}_1$	0.9814	0.0960	0.0130	0.0130
		$\hat{\beta}_2$	0.9896	0.0956	0.0091	0.0091
		$\hat{\beta}_3$	0.9920	0.0959	0.0092	0.0092
	WLS_{FIID}	$\hat{\beta}_0$	1.0000	0.0641	0.0102	0.0102
		$\hat{\beta}_1$	0.9917	0.0960	0.0141	0.0141
		$\hat{\beta}_2$	0.9895	0.0890	0.0090	0.0090
		$\hat{\beta}_3$	0.9890	0.0956	0.0091	0.0091
5% HLPs	LWS	$\hat{\beta}_0$	0.8989	0.2859	0.0524	0.0524
		$\hat{\beta}_1$	0.9864	0.2519	0.0403	0.0403
		$\hat{\beta}_2$	0.9548	0.2537	0.0383	0.0383
		$\hat{\beta}_3$	0.9320	0.2529	0.0385	0.0385
	WLS_F	$\hat{\beta}_0$	0.9748	0.2593	0.0218	0.0216
		$\hat{\beta}_1$	0.9629	0.2395	0.0219	0.0216
		$\hat{\beta}_2$	0.9453	0.2426	0.0226	0.0226
		$\hat{\beta}_3$	0.9453	0.2362	0.0219	0.0219
	WLS_{DRGP}	$\hat{\beta}_0$	0.9982	0.2213	0.0178	0.0183
		$\hat{\beta}_1$	0.9888	0.2145	0.0174	0.0177
		$\hat{\beta}_2$	0.9966	0.2165	0.0171	0.0174
		$\hat{\beta}_3$	0.9835	0.2134	0.0171	0.0174
	WLS_{FIID}	$\hat{\beta}_0$	**0.9993**	**0.1906**	**0.0122**	**0.0122**
		$\hat{\beta}_1$	**0.9896**	**0.1901**	**0.0166**	**0.0166**
		$\hat{\beta}_2$	**0.9950**	**0.1897**	**0.0165**	**0.0165**
		$\hat{\beta}_3$	**0.9958**	**0.1891**	**0.0171**	**0.0171**
10% HLPs	LWS	$\hat{\beta}_0$	0.8365	0.2913	0.0503	0.0503
		$\hat{\beta}_1$	0.9312	0.2523	0.0756	0.0756
		$\hat{\beta}_2$	0.9335	0.2531	0.0709	0.0709
		$\hat{\beta}_3$	0.9375	0.2523	0.0634	0.0634
	WLS_F	$\hat{\beta}_0$	0.9566	0.2832	0.0327	0.0327
		$\hat{\beta}_1$	0.9184	0.2534	0.0235	0.0235
		$\hat{\beta}_2$	0.9292	0.2542	0.0250	0.0250
		$\hat{\beta}_3$	09374	0.2510	0.0195	0.0195
	WLS_{DRGP}	$\hat{\beta}_0$	0.9901	0.2479	0.0220	0.0220
		$\hat{\beta}_1$	0.9776	0.2248	0.0164	0.0164
		$\hat{\beta}_2$	0.9836	0.2267	0.0176	0.0176
		$\hat{\beta}_3$	0.9871	0.2212	0.0147	0.0147
	WLS_{FIID}	$\hat{\beta}_0$	**0.9982**	**0.1806**	**0.0132**	**0.0132**
		$\hat{\beta}_1$	**0.9876**	**0.1900**	**0.0134**	**0.0134**
		$\hat{\beta}_2$	**0.9960**	**0.1797**	**0.0141**	**0.0141**
		$\hat{\beta}_3$	**0.9952**	**0.1792**	**0.0115**	**0.0115**

Table 2: Simulation result of RE panel data estimates for n = 20, t=30

Con. Level	Estimator		Coeff. of Estimates	Standard error of Estimates	Variance HC4	HC5
0 % HLPs	LWS	$\hat\beta_0$	1.0044	0.1095	0.0200	0.0200
		$\hat\beta_1$	0.9882	0.1099	0.0340	0.0340
		$\hat\beta_2$	1.0178	0.0999	0.0170	0.0170
		$\hat\beta_3$	0.9884	0.1099	0.0221	0.0221
	WLS_F	$\hat\beta_0$	1.0008	0.0972	0.0092	0.0092
		$\hat\beta_1$	1.0090	0.0993	0.0129	0.0129
		$\hat\beta_2$	0.9968	0.0956	0.0091	0.0091
		$\hat\beta_3$	0.9925	0.1016	0.0100	0.0100
	WLS_{DRGP}	$\hat\beta_0$	1.0008	0.0977	0.0094	0.0094
		$\hat\beta_1$	1.0092	0.0974	0.0136	0.0136
		$\hat\beta_2$	0.9915	0.0940	0.0094	0.0094
		$\hat\beta_3$	0.9980	0.0995	0.0100	0.0100
	WLS_{FIID}	$\hat\beta_0$	1.0007	0.0973	0.0090	0.0090
		$\hat\beta_1$	1.0081	0.0971	0.0129	0.0129
		$\hat\beta_2$	0.9960	0.0942	0.0091	0.0091
		$\hat\beta_3$	0.9942	0.0993	0.0103	0.0103
5% HLPs	LWS	$\hat\beta_0$	0.8789	0.2859	0.0524	0.0524
		$\hat\beta_1$	0.8864	0.2519	0.0403	0.0403
		$\hat\beta_2$	0.8648	0.2537	0.0383	0.0383
		$\hat\beta_3$	0.9020	0.2529	0.0385	0.0385
	WLS_F	$\hat\beta_0$	0.9072	0.2519	0.0231	0.0223
		$\hat\beta_1$	0.9251	0.2322	0.0211	0.0210
		$\hat\beta_2$	0.9350	0.2324	0.0211	0.0211
		$\hat\beta_3$	0.9213	0.2320	0.0211	0.0210
	WLS_{DRGP}	$\hat\beta_0$	0.9748	0.1445	0.0187	0.0187
		$\hat\beta_1$	0.9895	0.1351	0.0127	0.0127
		$\hat\beta_2$	0.9989	0.1354	0.0126	0.0126
		$\hat\beta_3$	0.9955	0.1351	0.0126	0.0126
	WLS_{FIID}	$\hat\beta_0$	**0.9927**	**0.1385**	**0.0100**	**0.0100**
		$\hat\beta_1$	**0.9966**	**0.1258**	**0.0101**	**0.0101**
		$\hat\beta_2$	**0.9916**	**0.1260**	**0.0118**	**0.0118**
		$\hat\beta_3$	**0.9902**	**0.1263**	**0.0109**	**0.0109**
10% HLPs	LWS	$\hat\beta_0$	0.8365	0.2913	0.0503	0.0503
		$\hat\beta_1$	0.7712	0.2523	0.0756	0.0756
		$\hat\beta_2$	0.8135	0.2530	0.0709	0.0609
		$\hat\beta_3$	0.8075	0.2522	0.0634	0.0604
	WLS_F	$\hat\beta_0$	0.8901	0.2644	0.0275	0.0275
		$\hat\beta_1$	0.9134	0.2352	0.0183	0.0183
		$\hat\beta_2$	0.9086	0.2349	0.0180	0.0180
		$\hat\beta_3$	0.9047	0.2343	0.0178	0.0178
	WLS_{DRGP}	$\hat\beta_0$	0.9737	0.1532	0.0201	0.0201
		$\hat\beta_1$	0.9855	0.1431	0.0137	0.0137
		$\hat\beta_2$	0.9822	0.1419	0.0133	0.0133
		$\hat\beta_3$	0.9710	0.1425	0.0134	0.0134
	WLS_{FIID}	$\hat\beta_0$	**0.9998**	**0.1347**	**0.0113**	**0.0113**
		$\hat\beta_1$	**0.9948**	**0.1342**	**0.0116**	**0.0116**
		$\hat\beta_2$	**0.9922**	**0.1339**	**0.0096**	**0.0096**
		$\hat\beta_3$	**0.9967**	**0.1240**	**0.0096**	**0.0096**

Table 3: Simulation result of RE panel data estimates for n = 30, t=40

Con. Level	Estimator		Coeff. of Estimates	Standard error of Estimates	Variance	
					HC4	HC5
0 % HLPs	LWS	$\hat{\beta}_0$	1.0010	0.1190	0.0208	0.0208
		$\hat{\beta}_1$	0.9897	0.1085	0.0262	0.0262
		$\hat{\beta}_2$	0.9918	0.1091	0.0165	0.0165
		$\hat{\beta}_3$	1.0020	0.1092	0.0170	0.0170
	WLS_F	$\hat{\beta}_0$	1.0013	0.1049	0.0113	0.0113
		$\hat{\beta}_1$	0.9809	0.0989	0.0116	0.0116
		$\hat{\beta}_2$	0.9988	0.0986	0.0080	0.0080
		$\hat{\beta}_3$	0.9920	0.0991	0.0081	0.0081
	WLS_{DRGP}	$\hat{\beta}_0$	1.0002	0.1052	0.0102	0.0102
		$\hat{\beta}_1$	0.9914	0.0971	0.0127	0.0127
		$\hat{\beta}_2$	0.9996	0.0966	0.0083	0.0083
		$\hat{\beta}_3$	0.9920	0.0959	0.0086	0.0086
	WLS_{FIID}	$\hat{\beta}_0$	1.0000	0.0929	0.0075	0.0075
		$\hat{\beta}_1$	0.9970	0.0858	0.0151	0.0151
		$\hat{\beta}_2$	0.9948	0.0955	0.0070	0.0070
		$\hat{\beta}_3$	0.9970	0.0954	0.0068	0.0068
5% HLPs	LWS	$\hat{\beta}_0$	0.9046	0.2168	0.0182	0.0182
		$\hat{\beta}_1$	0.9144	0.1965	0.0157	0.0157
		$\hat{\beta}_2$	0.9116	0.1965	0.0146	0.0146
		$\hat{\beta}_3$	0.9158	0.1969	0.0146	0.0146
	WLS_F	$\hat{\beta}_0$	0.9381	0.1403	0.0040	0.0040
		$\hat{\beta}_1$	0.9426	0.1583	0.0039	0.0039
		$\hat{\beta}_2$	0.9431	0.1581	0.0038	0.0038
		$\hat{\beta}_3$	0.9415	0.1582	0.0039	0.0039
	WLS_{DRGP}	$\hat{\beta}_0$	0.9632	0.1235	0.0024	0.0023
		$\hat{\beta}_1$	0.9750	0.1295	0.0022	0.0022
		$\hat{\beta}_2$	0.9665	0.1294	0.0020	0.0020
		$\hat{\beta}_3$	0.9604	0.1194	0.0021	0.0021
	WLS_{FIID}	$\hat{\beta}_0$	**0.9846**	**0.1163**	**0.0012**	**0.0012**
		$\hat{\beta}_1$	**0.9950**	**0.1264**	**0.0015**	**0.0015**
		$\hat{\beta}_2$	**0.9947**	**0.1153**	**0.0014**	**0.0014**
		$\hat{\beta}_3$	**0.9942**	**0.1155**	**0.0015**	**0.0015**
10% HLPs	LWS	$\hat{\beta}_0$	0.8584	0.1906	0.0167	0.0167
		$\hat{\beta}_1$	0.7932	0.1735	0.0129	0.0129
		$\hat{\beta}_2$	0.8835	0.1734	0.0120	0.0120
		$\hat{\beta}_3$	0.8840	0.1741	0.0122	0.0122
	WLS_F	$\hat{\beta}_0$	0.9404	0.1800	0.0094	0.0094
		$\hat{\beta}_1$	0.9460	0.1682	0.0081	0.0081
		$\hat{\beta}_2$	0.9464	0.1679	0.0087	0.0087
		$\hat{\beta}_3$	0.9456	0.1686	0.0087	0.0087
	WLS_{DRGP}	$\hat{\beta}_0$	0.9765	0.1776	0.0090	0.0090
		$\hat{\beta}_1$	0.9883	0.1634	0.0073	0.0073
		$\hat{\beta}_2$	0.9897	0.1622	0.0072	0.0072
		$\hat{\beta}_3$	0.9993	0.1614	0.0072	0.0072
	WLS_{FIID}	$\hat{\beta}_0$	**0.9844**	**0.1728**	**0.0081**	**0.0081**
		$\hat{\beta}_1$	**0.9901**	**0.1556**	**0.0068**	**0.0068**
		$\hat{\beta}_2$	**0.9957**	**0.1584**	**0.0060**	**0.0060**
		$\hat{\beta}_3$	**0.9907**	**0.1598**	**0.0059**	**0.0059**

b) Real Data Example (Artificial Data set)

Figure 2 shows that there is presence of heteroscedasticity in the artificial data set by producing a systematic funnel shape in the plot and Figure 3 shows the presence of IO, in which observation 73 is declared as IO, while observations 41, 42, 67, 68 and 74 are declared as GLO.

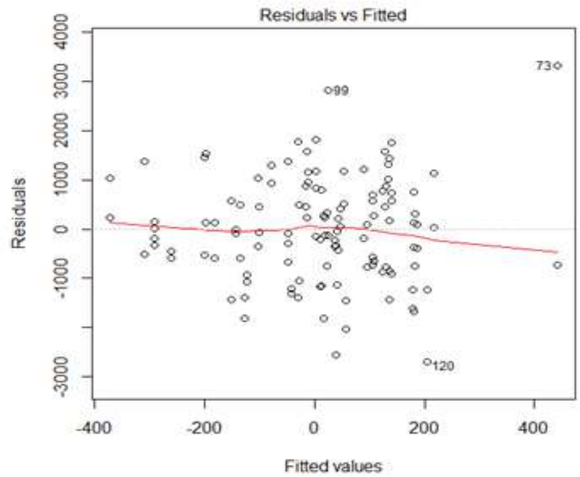

Figure 2: Plot of pooled OLS residuals versus fitted values for the artificial data set

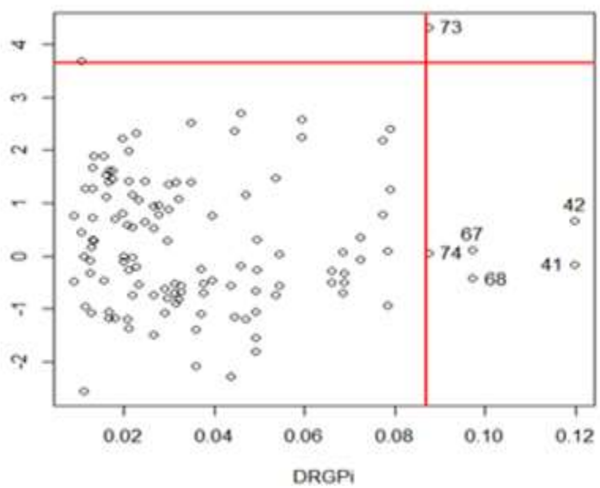

Figure 3: Plot of FIID for artificial data set

Table 4 presents the result of the artificial data set, which clearly shows that with the effect of only one IO the proposed (WLS_{FIID}) outperformed all the other methods by producing the smallest variances of HC4 and HC5, lowest values of standard error of the estimate.

This artificial data set was also modified by introducing more IOs, such that observations 10, 42 and 73 were inflated by 10 for x_1, x_2, x_3. Table 4 presents the result of the modified data, where the WLS_{FIID} was found to be the best method as compared with the existing by providing the lowest values of variances of HC4 and HC5 and lowest values of standard error of the estimate. This is due to the fact that FIID only down weight the bad HLPs whereas, the DRGP down weight both good and bad HLPs. Consequently its efficiency decreases.

Table 4: Regression estimates for the artificial and modified artificial RE panel data set

	Estimator		Coeff. of Estimates	Standard Error of Estimates	Variance	
					HC4	HC5
Artificial data	LWS	$\hat{\beta}_0$	-41.29	2694.44	8577593.75	8577593.75
		$\hat{\beta}_1$	-167.61	106.47	19618.99	19618.99
		$\hat{\beta}_2$	-3.59	110.97	12035.53	12035.53
		$\hat{\beta}_3$	180.55	94.55	49105.40	49105.40
	WLS_F	$\hat{\beta}_0$	-1151.89	2413.67	5259038.35	5259038.35
		$\hat{\beta}_1$	-74.59	96.61	7482.40	7482.40
		$\hat{\beta}_2$	65.29	109.41	9446.21	9446.21
		$\hat{\beta}_3$	93.34	98.24	12335.18	12335.18
	WLS_{DRGP}	$\hat{\beta}_0$	-1517.43	2324.99	4754610.66	4754610.66
		$\hat{\beta}_1$	-69.52	94.66	8002.45	8002.45
		$\hat{\beta}_2$	67.70	104.50	8265.38	8265.38
		$\hat{\beta}_3$	114.19	94.97	14091.77	14091.77
	WLS_{FIID}	$\hat{\beta}_0$	**-1178.49**	**2212.55**	**4749616.14**	**4749616.14**
		$\hat{\beta}_1$	**-50.66**	**91.16**	**6755.11**	**6755.11**
		$\hat{\beta}_2$	**70.96**	**98.65**	**7869.05**	**7869.05**
		$\hat{\beta}_3$	**65.13**	**91.65**	**9171.87**	**9171.87**
Modified artificial data	LWS	$\hat{\beta}_0$	-65.07	2539.97	7222619.91	7222619.91
		$\hat{\beta}_1$	-186.66	105.22	23719.80	23719.80
		$\hat{\beta}_2$	-4.10	109.38	11791.70	11791.70
		$\hat{\beta}_3$	201.89	93.36	59439.83	59439.83
	WLS	$\hat{\beta}_0$	-1272.01	2196.47	3888665.68	3888665.68
		$\hat{\beta}_1$	-78.69	94.44	6999.69	6999.69
		$\hat{\beta}_2$	58.40	104.10	8047.03	8047.03
		$\hat{\beta}_3$	112.61	97.52	11115.67	11115.67
	WLS_{DRGP}	$\hat{\beta}_0$	-1416.04	2204.16	3834424.62	3834424.62
		$\hat{\beta}_1$	-82.85	93.91	7983.33	7983.33
		$\hat{\beta}_2$	57.54	103.61	7781.96	7781.96
		$\hat{\beta}_3$	129.09	94.99	13508.27	13508.27
	WLS_{FIID}	$\hat{\beta}_0$	**-1179.39**	**1812.90**	**1768172.50**	**1768172.50**
		$\hat{\beta}_1$	**-51.75**	**89.51**	**7220.55**	**7220.55**
		$\hat{\beta}_2$	**69.95**	**90.98**	**5456.56**	**5456.56**
		$\hat{\beta}_3$	**65.66**	**92.01**	**10941.64**	**10941.64**

CONCLUSION

This paper addresses the combine problem of Influential Observations (IO) and Heteroscedasticity in random effect (RE) panel data model. It is now evident that a very low level of contamination by means of high leverage points and heteroscedasticity in RE panel data set has an effect to the existing robust estimation techniques. The use of hat matrix weighting method in WLS_F suffers tremendous effects due to masking and swamping effect. More efficient robust method for HLPs or IOs detection is needed in order to reduce the effect of swamping and masking. The proposed robust estimation method for RE panel data model used residuals from weighted least square (WLS) based on IO detection measure (FIID) weighting methods in computing RHCCM estimator. The results based on both simulation and numerical examples indicate that the proposed estimation methods outperformed the existing methods by providing smallest bias, smallest standard error of HC4 and HC5. The reason behind, is the good performance of FIID for not only allowing good HLPs to contribute in the estimation but also, the less swamping and masking effect of FIID. We conclude that FIID weighting method was found to be the best among all the method considered in this study.

REFERENCES

Baltagi B.H. (2008). *Econometric Analysis of Panel Data.* Wiley Chichester

Baltagi, B. H., (2001) *Econometric analysis of panel data.* 2^{nd}ed. New York, John Wiley

Bramati, M. C., Croux, C., (2007) Robust estimators for the fixed effects panel data model. *Econometric Journal,* 10, 521 – 540.

Crowder, D., Hand, D. (1990) *Analysis of repeated measures.* Chapman and Hall, London,

Davidson, R. and MacKinnon, J.G. (1993). Estimation and Inference in Econometrics. New York: Oxford University Press.

Furno, M. (1996). Small sample behavior of a robust heteroskedasticity consistent covariance matrix estimator. *Journal of Statistical Computation and Simulation* 54: 115-128.

Habshah M, Norazan MR, Rahmatullah Imon AHM (2009) The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression. *J. Appl. Stat.* 36(5):507–520.

Habshah M., Muhammad S. & Jayanthi A. (2017) Robust Heteroscedasticity Consistent Covariance Matrix Estimator based on Robust Mahalanobis Distance and Diagnostic Robust Generalized Potential Weighting Methods in Linear Regression. *Journal of Modern Applied Statistical Methods* 17(1):17 DOI: 10.22237/jmasm/1530279855.

Habshah Midi & Sani Muhammad (2018) Robust Estimation for Fixed and Random Effect Panel Data Models with Different Centering Methods. *Journal of Engineering and Applied Sciences* 13(17): 7156-7161.

Habshah Midi, Muhammad Sani, Shelan Saied Ismaeel & Jayanthi Arasan (2021) Fast Improvised Influential Distance for the Identification of Influential Observations in Multiple Linear Regression. *Sains Malaysiana,* 50 (7) (2021): 2085-2094 http://doi.org/10.17576/jsm-2021-5007-22

Lim, H. A. and Habshah M. (2016) Diagnostic robust generalized potential based on index set equality (DRGP(ISE)) for the identification of high leverage points in linear models. *Computational statistics*, 31:859-877

Lima, V.M.C., Souza, T,C., Cribari-Neto, F. and Fernandes, G.B. (2009). Heteroskedasticity- robust inference in linear regressions. *Communications in Statistics-Simulation and Computation* 39: 194-206

Maronna, R, A., Martin, R.D and Yohai, V.J., (2006) *Robust statistics: Theory and methods,* John Wiley. New York, ISBN: 10: 0470010924, 436.

Mazlina A., Habshah M. (2015) Robust centering in the fixed effect panel data model. *Pakistan. Journal Statistics,* 31(1) 33 – 48.

Muhammad Sani, Habshah Midi & Jayanthi Arasan (2019) Robust Parameter Estimation for Fixed Effect Panel Data Model in the Presence of Heteroscedasticity and High Leverage Points, *ASM Science. Journal 12, Special Issue 1, 2019 for IQRAC2018, 227-238*

Rana, S., Habshah M, and Imon AHMR (2012) Robust Wild Bootstrap for stabilizing the variance of Parameter Estimates in Heteroscedastic Regression Models in the Presence of Outlier; *Mathematical Problem in Engineering,* Article ID 730328. doi: 10.1155/2012/730328

Rousseeuw, P. J., B. C. Van Zomeren, (1990) Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association,* vol.85, pp. 633–639.

Visek, J.A. (2015) Estimating the model with fixed and random effects by a robust method, *Methodology Computational Applied Probability* 17, 999-1014.

Veradi,V. & Wagner, J. (2011) Robust estimation of linear fixed effects panel data model with an application to the exporter productivity premium. *Journal of Economics and Statistics.* 231(4) 546 – 557.

©2020 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via https://creativecommons.org/licenses/by/4.0/ which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.