International Journal of Innovative Mathematics, Statistics & Energy Policies
12(2):22-40, April-June, 2024

© SEAHI PUBLICATIONS, 2024 **www.seahipaj.org** **ISSN: 2467-852X**

Robust Location-Scale Estimator In Static Linear Finite Sample Fixed Effect Panel Model With Outliers

[1]E. M. Eke, [2]I. D. Essi, [3]C. A. Nse

[1]Department of Statistics, Covenant Polytechnic, Owerrinta, Abia State, Nigeria
[2-3]Department of Mathematics, Rivers State University, Port-Harcourt, Nigeria
[1]ekeemenikemong@gmail.com; [2]Essi.isaac@ust.edu.ng; [3]akwumbuom.nse@ust.edu.ng

ABSTRACT
The special feature of the static linear fixed effect panel data model estimation is that it has two phases of estimation: location and within group estimation. The first phase (location estimation) also called centering is considered in this paper. Eight centering estimators (Mean, Median, HuberM, BisquareM, optM, MoptM, MM and the proposed DCML centering) are compared at different sample sizes using the mean square error of the estimated location parameter and the scale estimate. From the results of the Monte Carlo simulations and the real-life economic datasets analysis, the following findings are made from this study: (1) For the simulated finite sample panel data having 0% outlying observations, the Mean centering estimator outperformed the other seven estimators for small samples sizes of less than or equals to 2000 (2) The optM and the MoptM estimators yield the best performing location estimators in the simulated finite sample panel data linear model having 5% random vertical outlier (3) The proposed DCML location estimator is stable as outliers increases and varies in nature. Its performance is quite similar to that MoptM, optM and BisquareM location estimators as well as that of MM location estimator. The proposed DCML location estimator is recommended in finite panel data with 10% or more outliers, given its capability to give highly robust results.
Keywords: Panel data, robust location-scale estimator, Outliers, DCML centering.

I. INTRODUCTION

The special feature of the fixed effect panel data model estimation is that it has two phases of estimation: location estimation and within group estimation. Location estimation first obtains the location estimate, a numerical index of the statistical data that generally indicates the centre, middle or representative of the dataset in a concise manner that the mind can get a quick understanding of the general size of the individual values in the dataset. The true value of the location parameter given as μ is usually unknown in a process. Therefore, the need arises to obtain the estimate of that location parameter generally represented as $\hat{\mu}$. In the classical estimation models, the mean represents the measure of the location. The sample mean as the classical location estimator is highly sensitive to outlier contamination and deviation from normality as a single outlying observation can make the entire data to fail to provide a good fit for the bulk of the data, Schoonhoven et al (2011) and Maronna et al (2019).

In fixed effect within group panel model, a common method used to analyze panel data where the focus is on the variation within each individual unit over time, the use of location estimates to transform data is imperative. This transformation is the within-group deviations from the time-invariant individual effects of the explanatory variable(s) and the dependent variable is often called demeaning of the data which simple entails the use of the location estimator to centre the observations in the panel data model. And this has come to be known in the panel data modelling as centering (Baltagi, 2021; Wooldridge, 2010; Midi and Mohammad, 2018).

Centering the panel data as a common pre-processing step in the fixed effect panel model involves subtracting the time-average of each variable from its individual observations. The importance of centering in panel data modelling according to Bramati and Croux (2007) includes (1) reduction of number of parameters. (2) making computational time for robust regression and algorithm remains feasible. (3) reduction of the multicollinearity in the explanatory variables by reducing the correlation between the individual explanatory variables. (4) removal of the time-invariant individual effects, hence making the model more efficient. However, the classical centering based on the mean also called Mean centering can be sensitive to outliers, which can lead to biased estimates.

One approach to mitigate this problem of the classical centering being sensitive to outliers and leading to biased estimates is to use robust centering methods, which are less sensitive to outliers. Among the robust location estimators are the Median, Trimmed Mean, Winsorized Mean, M-estimators of location (Huber, 1964) and the MM-estimator of location, Bakar and Midi (2015). The Median is a simple and widely used robust estimator. It is defined as the middle value of a sorted data set. The trimmed mean is a more efficient estimator than the Mean as it removes a fixed percentage of the extreme values from the data set before computing the mean. The Winsorized mean is similar to the trimmed mean, but instead of removing the extreme values, it replaces them with the nearest non-extreme value.

Martin et al (1989) and Marcela et al (2002) showed that for the M-estimator of location to fulfil the three equivariance properties with respects to scale, regression and affine transformation in the linearity transformation, the simultaneous M-estimator of location and scale (dispersion) is recommended. Such location estimators are called M-estimator of location and scale. The family of M-estimators of location and scale are Huber M-estimator of location and scale (Huber, 1964), simply referred here as HuberM; Tukey Bisquare M-estimator of location/scale (Beaton and Tukey, 1974) referred here as BisquareM; Optimal M-estimator of location/scale (Martin and Zamar, 1993), referred here as optM; Modified Optimal M-estimator of location/scale (Svarc, Yohai and Zamar, 2002), referred here as MoptM and the MM-estimator of location and scale (Bakar and Midi (2015) are more sophisticated estimators that minimizes a robust loss function that balances the influence of the extreme values with the rest of the data. But each of these M-estimators of location and scale are built on the asymptotic properties. The search for a location and scale estimators in the finite samples is still dearly desired. The objective of this paper is to develop a new robust panel data centering estimator suitable for finite-sample panel data model and determine the robustness of the new centering method for finite-sample panel data having certain percentages of block concentrated vertical outliers and random vertical outliers using simulated data and real-life economic datasets.

Huber (1964) developed the M-estimator of location parameter as a robust estimator in the presence of outliers. This M-estimator was compared with the sample mean, the sample median and the trimmed mean estimators. He also showed that the M-estimators have pleasant asymptotic properties such as sufficient conditions for asymptotic normality and an explicit expression for their asymptotic variance. Chaudhuri and SenGupta (1993) proposed the modified version of the M-estimation problem and discovered that the sample median is shown to be a solution of this problem for a wide range of choices of the score function. But their model has not been used in recent studies as the usefulness of the Huber M-estimator of location has dominated the Statistics and Econometrics literature. Schoonhoven et al (2011) applied the trimmed trimean location estimator in the areas of statistical quality control study and their simulation study showed that the trimmed trimean outperformed the $\bar{\bar{X}}$ which is very sensitive to outliers so that the phase I limits are biased in terms of their mean-squared errors and their effect on the $\bar{X}$ control charts used for real-time process monitoring. Martin (2018) showed that for the univariate location estimation, M-estimators are better than trimmed means or Winsorized means. The M-estimator is adjudged the most popular robust estimator as it has many variants based on different choices of the rho/psi/weight functions including Andrews, Bisquare, Cauchy, Fair, Hampel, Huber, Logistic, Median, Talworth, Welsch according to Eke, 2016 and Martin, 2021). Among these M-estimators, the ones that have been applied in estimation of the location and scale estimator are the Huber M-estimator of location (Huber, 1964) which has 95% efficiency; Bisquare M-estimator of location (Beaton and Tukey, 1974) with 95% efficiency; optimal M-estimator of location (Martin and Zamar, 1993) with 99% efficiency; Modified Optimal M-estimator of location (Svarc, Yohai and Zamar, 2022) with 99% efficiency. These location

estimators are considered in the cross-sectional data framework. Bakar and Midi (2015) proposed the MM Centering for panel data with 95% efficiency. Each of these M-estimators has 0.5 breakdown point, Zhang and Li (1998).

In the panel data model location estimation with outliers, Bramati and Croux (2007) and Aquaro and Cizek (2010) proposed the use of Median centering. But the Median location estimate used in Median centering has its challenges of efficiency (having efficiency of $2/\pi$ = 64%) and a shortcoming of the non-linearity of its loss function, To overcoming these problems of the Median centering of the within groups estimators, Bakar and Midi (2015) proposed the use of MM-centering after the MM-estimator of Yohai (1987) to provide robust solution to the classical within group Mean location parameter estimate when the data are contaminated with outliers. And to lend more empirical evidential support on the superiority of the MM-centering over the Median-centering and Mean-centering of data-contaminated panel data estimation procedures, Midi and Mohammed (2018) compare these three-centering estimators (Mean, Median and MM) and the Mean centering was found to be very sensitive to outliers, while the Median centering was non-linear with low efficiency. The MM-centering outperformed the Median-centering and Mean-centering. They aligned their simulation study with that of Bramati and Croux (2007) having four data contamination conditions (random vertical outliers, random leverage, block concentrated vertical outlier and block concentrated leverage) but with 15% and 25% contamination levels for linear panel models.

The rest of the sections of this paper are organised as follows: Section II considers the methodologies applied in the centering estimators applied in the classical and robust within group panel fixed effect panel model which include the Mean centering, Median centering, HuberM Centering, BisquareM Centering, OptM Centering, MoptM Centering and the MM centering. And then proceeded to develop the proposed robust estimator of location and scale based on the Distance Constrained Maximum Likelihood (DCML) estimator which we called the DCML Centering and explain the Monte Carlo simulation study carried out to evaluate the performance of the proposed DCML estimators of location and scale against the existing estimators of the location and scale. In section III, the results of the various analyses are presented in tables and charts to address the objective of the study and the discussion of the results. The conclusion drawn from the discussions of the results and the recommendations on the areas for future research are presented in the section IV.

II. METHODOLOGY

The methodology used for the classical and robust centering estimation procedures are considered in this section.

The Mean centering is the classical methods of centering panel data without outliers. The mean centered data also called time demeaned data, $\ddot{y}_{it}$ is given as

$$\ddot{y}_{it} = y_{it} - \bar{y}_{i.} \qquad (1)$$

$where\ \bar{y}_{i.} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$ is group mean for the panel data.

The standard deviation (sd) which is the estimator of scale for the panel data, it is given as

$$sd = \sqrt{\frac{1}{n}\sum_{it=1}^{n}(y_{it} - \ddot{y}_{it})^2} \qquad (2)$$

where $n = N \times T$ and $\ddot{y}_{it}$ is the time demeaned centered value.

When the panel data contain outliers, the Median is the first considered centering methods. The median centered data, $\tilde{y}_{it}$ is given as

$$\tilde{y}_{it} = y_{it} - \underset{t}{median}(y_{it}) \quad (3)$$

Median absolute deviation (MAD), the estimator of scale for outliers ridden data is given as

$$MAD = Median\left|y_{it} - \underset{t}{median}(y_{it})\right| \quad (4)$$

For the normalized and unbiased estimator of scale, $\hat{\sigma}$ is given as $\hat{\sigma} = \frac{MAD}{0.6745}$.

The Huber M-estimator of location and scale, referred here as $\hat{\mu}_{HuberM}$ is a minimizer of

$$y_i = \mu + \sigma\varepsilon_i \quad (5)$$

and is given as

$$\hat{\mu}_{HuberM} = argmax\ L(y_1, y_2, \dots, y_n) = arg\min_{\mu} \sum_{i=1}^{n} \rho\left(\frac{y_i - \mu}{\sigma}\right) \quad (6)$$

where $\rho(x) = -\log(f_0)$

Huber (1964) showed that $\hat{\mu}_{HuberM}$ has 95% efficiency with monotone and bounded ψ function. The ρ and ψ functions for $\hat{\mu}_{HuberM}$ are given in (7) and (8) below:

$$\rho(\varepsilon) = \begin{cases} \frac{\varepsilon^2}{2} & if\ |\varepsilon/c| \le 1 \\ c|\varepsilon| - \frac{\varepsilon^2}{2} & if\ |\varepsilon/c| > 1 \end{cases} \quad (7)$$

$$\psi(\varepsilon) = \begin{cases} \varepsilon & if\ |\varepsilon/c| \le 1 \\ c \times sign(\varepsilon) & if\ |\varepsilon/c| > 1 \end{cases} \quad (8)$$

where $c = 1.345$ for the $\hat{\mu}_{HuberM}$

The Bisquare M-estimator of location and scale, referred here as $\hat{\mu}_{BisquareM}$ has the ρ and ψ functions given below:

$$\rho(\varepsilon) = \begin{cases} \frac{c^2}{6}\{1 - [1 - (\varepsilon/c)^2]^3\} & if\ |\varepsilon/c| \le 1 \\ \frac{c^2}{6} & if\ |\varepsilon/c| > 1 \end{cases} \quad (9)$$

$$\psi(\varepsilon) = \begin{cases} c^2[1-(\varepsilon/c)^2]^2 & if \ \ |\varepsilon/c| \le 1 \\ 0 & if \ \ |\varepsilon/c| > 1 \end{cases} \qquad (10)$$

where c is a tuning constant that determines the sensitivity to outliers, given to be equal to 4.685 and c can be tuned for breakdown point efficiency of 95%.

Because the constant c in the Huber and Tukey psi weights can be tuned to obtain a higher breakdown point efficiency. Martin (2021) confirmed this that "Among all the robust regression estimators available today, only two estimators have the property that they minimize an asymptotic bias measure over a Tukey-Huber family of distributions, namely the Yohai and Zamar (1997) estimator, which we refer in this thesis as the Optimal M-estimator, and the slightly different Svarc et al. (2002) estimator" referred to as Modified Optimal M-estimator. Martin and Zamar (1993) derived the optimal M-estimate of Location and scale, which we refer here as $\hat{\mu}_{OptM}$ with optimal psi function given as

$$\psi(\varepsilon) = sign(y)\left(|y| - \frac{a}{\phi(y)}\right)^{+} = \begin{cases} y - sign(y)\dfrac{a}{\phi(y)} & if \ \ |y| \in (lower, upper) \\ 0 & otherwise \end{cases} \qquad (11)$$

The psi function of the Modified Optimal M-estimator of location $\hat{\mu}_{MoptM}$ is given as

$$\psi_a^{(mod)}(\varepsilon) = \left(\frac{\phi(1)}{\phi(1)-a}\right)\psi_a^{(0)}(\varepsilon) = \begin{cases} \varepsilon & -1 \le \varepsilon \le 1 \\ \left(\dfrac{\phi(1)}{\phi(1)-a}\right)y - sign(y)\dfrac{a}{\phi(y)} & if \ \ |y| < upper \\ 0 & otherwise \end{cases} \qquad (12)$$

where

$$\psi_a^{(0)}(\varepsilon) = \begin{cases} \left(1-\dfrac{a}{\phi(1)}\right)\varepsilon & -1 \le \varepsilon \le 1 \\ y - sign(y)\dfrac{a}{\phi(y)} & if \ \ |y| < upper \\ 0 & otherwise \end{cases} \qquad (13)$$

The optimal M-estimator of Location and scale and the Modified Optimal M-estimator of location and scale have 99% efficiency. HuberM, BisquareM, optM and MoptM location estimators are implemented in the locScaleM function of RobStatTM package in R software.

The MM-location and scale estimator for fixed effect panel data proposed by Bakar and Midi (2015) follow the MM-Estimator of linear regression of Yohai (1987) that showed that for a normally distributed error term, ε_{it} with initial scale estimate of $\hat{\sigma}_0^2 = 1$, the MM estimate depends on a chosen tuning constant of 4.685 such that that the MM-estimate can be highly efficient without affecting its breakdown point of 0.5 and efficiency of 0.95 for normal error. The MM-estimate is based on the Bisquare rho function. The location model used by Bakar and Midi (2015) is the same with the location model of used by Huber (1964) given as $x_i = \mu + \sigma e_i$. The derivations of the estimators of μ and σ are in line with the Huber-Tukey model approach. MM-estimator of Location, which we refer in this paper as $\hat{\mu}_{MM}$ is the solution of

$$\sum \psi\left(\frac{e_i}{\hat{\sigma}}\right) x_i = 0 \qquad (14)$$

following the MM-estimator of regression given by Yohai (1987). And as given in Bakar and Midi (2015) and also in Ismaeel and Midi (2018), the summary of the algorithm for obtaining the MM-estimates of location and scale is:

Step 1: Compute the initial estimate of location ($\hat{\mu}_0$) and the initial estimates of scale ($\hat{\sigma}_0$) obtained from the S-estimator of Rousseeuw (1984).

Step 2: Calculate the residuals (e_i) by subtracting the observed response variable values (x_i) from the predicted values based on the current estimate of step 1.

Step 3: Use the residuals obtained in step 2 to compute the M estimate of scale $\hat{\sigma}$, where $\hat{\sigma}$ is the solution to $\frac{1}{n}\sum \rho_0\left(\frac{e_i}{\hat{\sigma}}\right) = 0.5$ and $\psi = \rho'$ is the Tukey's Bisquare Psi function.

Step 4: Compute M estimate of $\hat{\mu}$ using ρ_1 according to Yohai (1987) highlight of using Tukey's Bisquare weight function $c_1 = 4.685$, result in high efficiency such that $\hat{\mu}$ is a solution to $\sum \psi\left(\frac{e_i}{\hat{\sigma}}\right) x_i = 0$ where $\psi = \rho'\left(\frac{t}{c_1}\right)$, and $t = \frac{e_i}{\hat{\sigma}}$ using the Iteratively Weighted Least Squares algorithm (IWLS).

Step 5: Repeat steps 2-4 iteratively until convergence is achieved. Convergence is typically determined by checking whether the estimates of the coefficients stabilize. Upon convergence, $\hat{\mu}$ and $\hat{\sigma}$ is the MM location and MM scale estimates.

The MM location estimate and MM scale estimate denoted in this paper as $\hat{\mu}_{MM}$ and $\hat{\sigma}_{MM}$ respectively are obtained by implementing the above algorithm in the R software.

Proposed DCML location and Scale Estimators in Fixed Effect Panel Data Model

In a panel data random process, the unknown location parameter μ, the individual specific unobserved effects α_i, the scale parameter σ and the error term ε_{it} combine simultaneously and additively in the form

$$y_{it} = \mu + \alpha_i + \sigma\varepsilon_{it} \qquad (15)$$

where y_{it} is observed value for the i[th] cross-sectional unit at t[th] time period, μ is location parameter, α_i are the individual specific unobserved effects, σ is the unknown scale parameter and ε_{it} is the normally distributed error term for the i[th] unit and the t[th] time period.

And in this, (15) is called DCML location model which is different from the Huber location model of (5). The choice of the proposed location model (15) which is different from Bakar and Midi MM-location model is informed by the argument of Hsiao (2022, p. 19) "*if y is also affected by other variables that we are not able to control and standardize within classes, the simple within-class sum of squares will be an overestimate of the stochastic component in* y*, and the differences between class means will reflect not only any class effect but also the effects of any differences in the values assumed by the uncontrolled variables in different classes*" and for the stochastic component in y not to be overestimated, we propose the introduction of the individual specific unobserved effects in the location model.

Following the approach similar to that of the DCML regression estimator of Maronna and Yohai (2014), the DCML location estimator $\hat{\mu}_{DCML}$, and that of scale $\hat{\sigma}_{DCML}$ are derived below:

Using the maximum likelihood estimation, we assume the error term ε_{it} follows the normal distribution with mean zero and variance one, $i.e.\ \varepsilon_{it} \sim N(0,1)$. The likelihood function for y_{it} given by the probability density function of the normal distribution is given as:

$$f(y_{it};\mu,\alpha_i,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y_{it}-\mu-\alpha_i)^2}{2\sigma^2}\right) \quad (16)$$

To estimate the parameters $\mu, \alpha_i, and\ \sigma$, we maximize the log-likelihood function:

$$logL(\mu,\alpha_1,\alpha_2,\ldots,\alpha_N,\sigma) = \sum_{i=1}^{N}\sum_{t=1}^{T}\log f(y_{it};\mu,\alpha_i,\sigma) \quad (17)$$

To incorporate the Kullback-Leibler (KL) statistical distance constraint, as given in Maronna and Yohai (2014), we add a penalty term to the log-likelihood function and obtained a KL distance between two probability distributions P and Q defined as:

$$D_{KL}(P||Q) = \int P(x)\log\left(\frac{P(x)}{Q(x)}\right) \quad (18)$$

In this case of location and scale estimator, we use the Kullback-Leibler distance between the estimated distribution of α_i and a specified reference distribution. The reference distribution chosen is the uniform distribution based on prior knowledge or assumptions about α_i that follows a continuous uniform distribution over the parameters 1 and 12. Adding the penalty term to the log-likelihood function, the constrained log-likelihood function becomes:

$$logL_c(\mu,\alpha_1,\alpha_2,\ldots,\alpha_N,\sigma) = logL(\mu,\alpha_1,\alpha_2,\ldots,\alpha_N,\sigma) - \lambda.D_{KL}(P||Q) \quad (19)$$

λ is a tuning parameter that controls the trade-off between maximizing the likelihood and minimizing the KL distance, P_α is estimated distribution of α_i and Q is reference distribution.
The constrained log-likelihood function maximized using optimization techniques to obtain the estimators of μ and σ. The estimation procedure involves iteratively updating the estimators until convergence, adjusting the penalty term based on the Kullback-Leibler distance, and maximizing the constrained log-likelihood. Applying the Lagrange multipliers optimization technique to obtain the estimators of μ and σ, we have the objective function to be maximized as the log-likelihood function given as:

$$L(\mu,\sigma) = f(y_{it};\mu,\alpha_i,\sigma) = \sum_{i=1}^{N}\sum_{t=1}^{T}\log f(y_{it};\mu,\alpha_i,\sigma) \quad (20)$$

subject to the constraint:

$$g(\mu,\sigma) = \sum_i \alpha_i = 0 \quad (21)$$

Introducing the Lagrange multiplier λ to the objective function to incorporate the constraint:

$$L(\mu,\sigma,\lambda) = f(y_{it};\mu,\alpha_i,\sigma) = \sum_{i=1}^{N}\sum_{t=1}^{T}\log f(y_{it};\mu,\alpha_i,\sigma) + \lambda\sum_i \alpha_i \quad (22)$$

To find the estimators, we differentiate the Lagrangian with respect to μ, σ, and λ, and set the derivatives equal to zero: $\partial L/\partial\mu = 0$; $\partial L/\partial\sigma = 0$; $\partial L/\partial\lambda = 0$. Solving these equations gives the following normal equations that yield the solutions of μ and σ :

$$\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{y_{it}-\hat{\mu}-\alpha_i}{\hat{\sigma}^2}\right)=0 \qquad (23)$$

$$\sum_{i=1}^{N}\left(\frac{(y_{it}-\hat{\mu}-\alpha_i)^2}{\hat{\sigma}^3}\right)-\lambda\sum_{i}\sum_{t}\left(\frac{1}{\hat{\sigma}}\right)=0 \qquad (24)$$

$$\sum_{i}\alpha_i=0 \qquad (25)$$

To derive the estimators of μ and σ, we solve the normal equations (23), (24) and (25):

From (25), this implies that the average of the α_i is zero. That is, estimator of $\alpha_i=\bar{\alpha}$ is

$$\bar{\alpha}=\frac{1}{N}\sum_{i}\alpha_i=0 \qquad (26)$$

We substitute the estimator $\bar{\alpha}$ = 0 into (23) and solve out to have the final estimator for μ as:

$$\hat{\mu}=\frac{\sum y}{NT} \qquad (27)$$

Solving for $\hat{\sigma}$ using (23) and (24) together, we obtain.

$$\hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{it}-\frac{\sum y}{NT}\right)^2}{\lambda NT}} \qquad (28)$$

Maronna et al (2014) showed that $\lambda=1.547$. So, the classical DCML estimators of σ when the y_{it} has no outlier are:

$$\hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\hat{\mu})^2}{1.547NT}} \qquad (29)$$

For a robust estimation of μ and σ, the MM-estimator optimization technique proposed by Yohai (1987) and used by Bakar and Midi (2015) is followed which minimize the objective function by iteratively updating the estimators in line with Maronna and Yohai (2014) approach. And if we choose the initial estimates for μ and σ and denote them as $\hat{\mu}_0, \hat{\sigma}_0$ respectively, then, for each observation y_{it}, the standardized residual r_{it} becomes

$$r_{it}=\left(\frac{y_{it}-\hat{\mu}_0}{1.547\hat{\sigma}_0}\right) \qquad (30)$$

The robust estimator for the residuals will be solution of the MM-location and scale estimator

$$\frac{1}{NT}\sum_{it=1}^{NT}\rho_B\left(\frac{y_{it}-\hat{\mu}_0}{1.547\hat{\sigma}_0}\right)=0.5 \qquad (31)$$

The Kullback_Leibler (KL) distance, a measure of the statistical distance between the initial estimate of μ and its estimated parameter becomes

$$d_{KL}(\hat{\mu}_0,\hat{\mu})=\frac{1}{\sigma^2}(\hat{\mu}_0-\hat{\mu})'\boldsymbol{C}(\hat{\mu}_0-\hat{\mu}) \qquad (32)$$

with $\boldsymbol{C}=\boldsymbol{E}(\alpha_i\alpha_i')=\frac{1}{N}\sum_{i=1}^{N}\alpha_i\alpha_i'$

Maronna and Yohai (2014) showed that the tuning constant c_0 = 1.547 based on the MM-estimators with a desired breakdown point of 50%, efficiency of 99% and optimal ρ yields the weight function w_{it} given as

$$w_{it}=\begin{cases}1 & \text{if } |r_{it}|\le 1.547\\ 1.547/|r_{it}| & \text{if } |r_{it}|>1.547\end{cases} \qquad (33)$$

This weight function assigns a weight of 1 to observations within a certain standardized residual distance from the estimated location and a decreasing weight to observations outside this distance. $\hat{\mu}_{DCML}$ and $\hat{\sigma}_{DCML}$ are given as

$$\hat{\mu}_{DCML}=arg\min_{\mu}\sum_{it=1}^{NT}r_{it}^2(\mu) \qquad Subject\ to\ d_{KL}(\mu_0,\mu) \qquad (34)$$

$$\hat{\sigma}_{DCML}=arg\min_{\mu}\sum_{it=1}^{NT}r_{it}^2(\mu) \quad Subject\ to\ d_{KL}(\sigma_0,\sigma)=\frac{1}{\hat{\sigma}^2}(\hat{\sigma}_0-\hat{\sigma})'\boldsymbol{C}(\hat{\sigma}_0-\hat{\sigma}) \qquad (35)$$

These are implemented in R Software to obtain the $\hat{\mu}_{DCML}$ and $\hat{\sigma}_{DCML}$.

The DCML Centered data, $\breve{y}_{it}$ will be given as

$$\breve{y}_{it}=y_{it}-\hat{\mu}_{DCML}(y_{it})\ for\ i=1,2,\dots,N\ and\ t=1,2,\dots,T \qquad (36)$$

The algorithm for obtaining the DCML estimates of μ and σ is:

1. Compute the Pena-Yohai Estimator of μ and σ using y_{it}.
2. Compute the MM-location-Scale Estimator:
   - Use the Pena-Yohai estimates as starting values for the MM-Location-Scale estimator.
   - Solve for the M estimate of scale using the equation $\frac{1}{n}\sum_{it=1}^{n}\rho_B\left(\frac{e_{it}}{1.547*\hat{\sigma}_0}\right)=0.5$ where ρ_B is the Tukey's Bisquare rho function implementing the Iteratively Weighted Least Squares (IWLS) iterative algorithm to obtain MM-location-Scale estimates of μ and σ.
3. DCML estimation of μ and σ:
   - Calculate $\boldsymbol{C}=\sum_{i=1}^{N}\frac{\alpha_i*\alpha_i'}{N}$.

- Compute $d_{KL\mu} = \frac{1}{\hat{\sigma}^2}(\hat{\mu}_0 - \hat{\mu})' \boldsymbol{C}(\hat{\mu}_0 - \hat{\mu})$, where $\hat{\mu}_0$ is the estimate of $\hat{\mu}$ obtained from MM-Location-Scale estimator of μ.
- Compute $d_{KL\sigma} = \frac{1}{\hat{\sigma}^2}(\hat{\sigma}_0 - \hat{\sigma})' \boldsymbol{C}(\hat{\sigma}_0 - \hat{\sigma})$, where $\hat{\sigma}_0$ is the estimate of σ obtained from MM-Location-Scale estimator of σ.
- The DCML estimates of μ = $\hat{\mu}_{DCML}$ is the minimizer of $e_{it}^2 \; Subject\ to\ d_{KL\mu} \leq \frac{0.3}{n}$. That is $\hat{\mu}_{DCML} = \arg min \sum_{it=1}^{NT} e_{it}^2 \; Subject\ to\ d_{KL\mu} \leq \frac{0.3}{n}$.
- The DCML estimates of σ = $\hat{\sigma}_{DCML}$ is the minimizer of $e_{it}^2 \; Subject\ to\ d_{KL\sigma} \leq \frac{0.3}{n}$. That is $\hat{\sigma}_{DCML} = \arg min \sum_{it=1}^{NT} e_{it}^2 \; Subject\ to\ d_{KL\sigma} \leq \frac{0.3}{n}$
- Implement the Iteratively Weighted Least Squares (IWLS) algorithm to iteratively update the estimates of $\hat{\mu}_{DCML}$ and $\hat{\sigma}_{DCML}$ until convergence.

Simulation study and real-life economic datasets

The simulation study compares the proposed DCML centering with the seven existing centering methods in the static fixed effect panel data model. The benchmark estimator is the classical Mean centering. The data generating process of the location/scale model for the linear static fixed effect panel data with additive effect model is given by (37) below:

$$y_{it} = \mu + \alpha_i + \sigma\varepsilon_{it} \quad ; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \qquad (37)$$

where y_{it}, μ, α_i and σ are as defined in (15).

For (37), the location parameter $\mu = 5$ and the unobservable individual effects $\alpha_i \sim U(0,12)$ in line with Bramati and Croux (2007), Bakar and Midi (2015) and Ismaeel and Midi (2018). The simulated panel data using the combinations of $N = 5, 10, 40$ and $T = 10, 30, 50$ are obtained and their scatter diagrams plotted to confirm that they have the features of the expected panel data. The performance of each estimator is evaluated using $R = 1000$ simulated samples, measured by the mean squared error of the estimated location parameter, $MSE(\hat{\mu})$ and the estimate of the standard deviation, $\hat{\sigma}$. The $MSE(\hat{\mu})$ is given as

$$MSE(\hat{\mu}) = \frac{1}{R}\sum_{r=1}^{R} \|\hat{\mu}^r - \mu\|^2 \qquad (38)$$

where $\hat{\mu}^r$, $r = 1, \ldots, 1000$, are the estimates for R simulated samples.

Outlying Contaminations

To ascertain the robustness performance of the location estimators, nine outlying contaminated panel data are considered: (1) uncontaminated data i.e., 0% outliers (2) 5% random vertical outliers (RVO5%) (3) 5% block concentrated vertical outlier (BCVO5%) (4) 5% random vertical outliers and bad leverage point (RVO+BLP5%) (5) 5% block concentrated vertical outliers and bad leverage point (BCVO+BLP5%) (6) 10% random vertical outlier (RVO10%) (7) 10% block concentrated vertical outlier (BCVO10%) (8) 10% random vertical outlier and bad leverage point (RVO+BLP10%) (9) 10% block concentrated vertical outlier and bad leverage point (BCVO+BLP10%). The proportions of contamination chosen are 5% and 10% in line with the recommendation of Shu (1978, p. 6). The considered contamination schemes depending on the types of outliers are described as: (1) To generate random vertical outliers (RVO) for the dependent variable (y_{it}^r), randomly selected original values of the dependent variable are replaced by the observations from a normal distribution $y_{it}^r \sim \mathrm{N}(0, 20)$. (2) To generate random vertical outliers for the independent variable (x_{it}^r), randomly selected original values of the independent variable are replaced by the observations from a

normal distribution $x_{it}^{r} \sim N(0,10)$. (3) Block concentrated vertical outliers (BCVO) are inserted into the data by substituting the random observations from a normal distribution $y_{it}^{c} \sim$ N(30,2) for the randomly selected blocks of the original values of dependent variable. (4) The concentrated BLP are inserted into the original blocks of the independent variables corresponding to the blocks already contaminated in dependent variable by replacing them by the random values from a normal distribution $N(\mu = 0, \sigma^2 = 25)$.

Real-life economic datasets

Dataset 1: Dataset 1 contains gross domestic product (gdp) and imports variables obtained from WDI 2022 edition. The finite panel dataset is formed by taking seven countries (cross-sectional units) and twenty-six years (1995-2020). Each of the seven countries are the economic giant of each economic continental block. The seven countries alphabetically arranged are Argentina, Australia, China, France, Israel, Nigeria and United States. To control the heteroscedasticity effect in regressing gdp on imports, the natural logarithm (ln) of the two variables was taken and denoted as lngdp and lnimports. And to capture for the various outlying contaminations in the panel data, dataset 1 is modified to reflect the outliers in panel data, Kinal and Lahiri (1993) and Croissant and Millo (2019).

Dataset 2: Dataset 2 is GDP Growth and foreign direct investment (FDI). The dependent variable, GDP Growth is subject to outlier contamination. A surge in international investment leads to an outlier in GDP growth. Also, a recovery from a natural disaster results in an outlier in GDP growth. The exogenous independent variable that affects GDP Growth is FDI which is susceptible to leverage point influence. An increase in foreign investment as part of recovery efforts serves as a leverage point. The panel dataset 2 has five countries (cross-sectional units) and twenty years (2001-2020). The five countries are all the anglophone countries in West Africa. The five countries alphabetically arranged are Gambia, Ghana, Liberia, Nigeria and Sierra Leone. Dataset 2 has all the attributes of influential values contamination in panel data and is also sourced from WDI 2022 edition. And because all the five anglophone countries of West Africa is taken, the fixed effect panel data is most suitable to analyse dataset 2.

The R-Codes for the simulation and the real-life economic datasets analyses can be obtained from the authors on request.

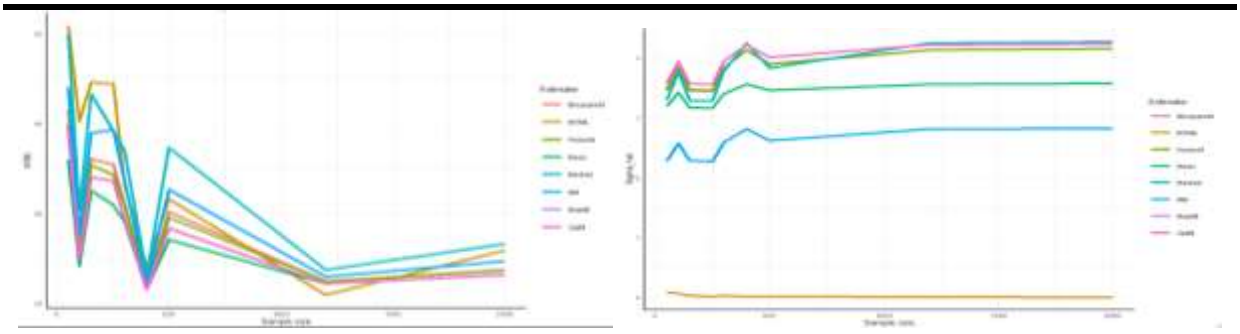
III. RESULTS

Table 1: Mean square error of location and scale estimates for the outlying panel datasets

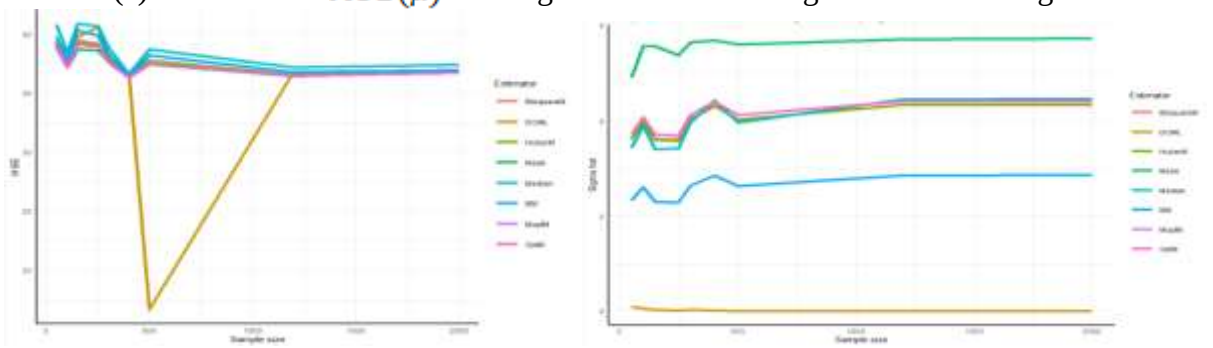
Centering Estimator	Outlier %	N	5	5	5	10	10	10	40	40	40
		T	10	30	50	10	30	50	10	30	50
Mean	0% outlier (No contamination)	$MSE(\hat{\mu})$	**39.19**	**38.51**	**38.18**	**36.82**	**37.82**	**37.41**	**36.31**	**36.43**	**36.63**
		$\hat{\sigma}$	3.18	3.16	3.15	3.42	3.40	3.45	3.56	3.56	3.57
Median		$MSE(\hat{\mu})$	41.97	40.65	39.84	38.07	39.40	39.46	36.62	36.74	37.32
		$\hat{\sigma}$	3.29	3.28	3.28	3.76	3.77	3.82	4.24	4.24	4.27
HuberM		$MSE(\hat{\mu})$	39.98	39.07	38.86	37.11	38.20	37.91	36.37	36.49	36.73
		$\hat{\sigma}$	3.46	3.46	3.45	3.83	3.83	3.89	4.13	4.13	4.15
BisquareM		$MSE(\hat{\mu})$	40.31	39.21	39.08	37.16	38.22	38.03	36.36	36.49	36.74
		$\hat{\sigma}$	3.45	3.44	3.44	3.82	3.83	3.89	4.12	4.13	4.14
OptM		$MSE(\hat{\mu})$	39.86	38.80	38.72	36.99	37.88	37.67	36.31	36.44	36.64
		$\hat{\sigma}$	3.57	3.56	3.56	3.93	3.94	4.00	4.21	4.21	4.23
MoptM		$MSE(\hat{\mu})$	39.86	38.81	38.72	36.99	37.88	37.67	36.31	36.44	36.64
		$\hat{\sigma}$	3.57	3.56	3.56	3.93	3.94	4.00	4.21	4.21	4.23

MM		$MSE(\hat{\mu})$	40.81	39.79	39.86	37.53	38.74	38.53	36.46	36.59	36.93
		$\hat{\sigma}$	2.27	2.28	2.27	2.57	2.58	2.61	2.81	2.81	2.82
DCML		$MSE(\hat{\mu})$	42.17	40.91	40.88	40.06	38.66	38.31	36.80	36.19	37.18
		$\hat{\sigma}$	**0.08**	**0.03**	**0.02**	**0.07**	**0.03**	**0.02**	**0.02**	**0.01**	**0.003**
Mean	5% Random Vertical Outliers	$MSE(\hat{\mu})$	38.88	38.69	38.61	37.13	37.63	37.43	36.27	36.43	36.74
		$\hat{\sigma}$	4.91	5.57	5.37	5.57	5.64	5.60	5.68	5.70	5.72
Median		$MSE(\hat{\mu})$	40.79	40.84	40.65	38.40	38.87	38.70	36.61	37.16	37.36
		$\hat{\sigma}$	3.43	3.41	3.41	3.89	3.96	3.95	4.43	4.44	4.46
HuberM		$MSE(\hat{\mu})$	39.15	39.24	39.02	37.31	37.81	37.71	36.34	36.51	36.76
		$\hat{\sigma}$	3.62	3.61	3.59	3.98	4.03	4.02	4.31	4.33	4.34
BisquareM		$MSE(\hat{\mu})$	39.34	39.44	39.18	37.35	37.82	37.74	36.34	36.54	36.76
		$\hat{\sigma}$	3.60	3.60	3.58	3.97	4.02	4.02	4.31	4.33	4.33
OptM		$\boldsymbol{MSE(\hat{\mu})}$	**39.08**	**39.06**	**38.89**	**37.14**	**37.60**	**37.45**	**36.30**	**36.45**	**36.68**
		$\hat{\sigma}$	3.71	3.71	3.68	4.08	4.12	4.11	4.39	4.41	4.41
MoptM		$\boldsymbol{MSE(\hat{\mu})}$	**39.08**	**39.06**	**38.89**	**37.14**	**37.60**	**37.45**	**36.30**	**36.45**	**36.68**
		$\hat{\sigma}$	3.71	3.71	3.68	4.08	4.12	4.11	4.39	4.41	4.41
MM		$MSE(\hat{\mu})$	39.74	40.31	39.85	37.73	38.29	38.20	36.44	36.72	36.93
		$\hat{\sigma}$	2.33	2.29	2.29	2.60	2.63	2.62	2.85	2.85	2.86
DCML		$MSE(\hat{\mu})$	39.19	39.72	40.59	38.45	37.84	16.68	36.66	36.80	36.86
		$\hat{\sigma}$	**0.09**	**0.03**	**0.02**	**0.06**	**0.02**	**0.01**	**0.02**	**0.01**	**0.003**
Mean	10% Random Vertical Outliers	$MSE(\hat{\mu})$	38.48	39.71	38.78	38.35	37.33	38.01	36.41	36.21	36.52
		$\hat{\sigma}$	6.90	7.02	7.01	7.06	7.18	7.13	7.21	7.23	7.25
Median		$MSE(\hat{\mu})$	39.28	41.15	40.77	39.23	38.83	39.84	36.95	36.55	36.92
		$\hat{\sigma}$	3.69	3.61	3.59	4.12	4.15	4.10	4.58	4.64	4.61
HuberM		$MSE(\hat{\mu})$	38.34	39.69	39.21	38.24	37.53	38.25	36.43	36.29	36.46
		$\hat{\sigma}$	3.86	3.83	3.78	4.20	4.22	4.18	4.49	4.53	4.51
BisquareM		$MSE(\hat{\mu})$	38.48	39.79	39.30	38.25	37.60	38.31	36.44	36.30	36.45
		$\hat{\sigma}$	3.84	3.81	3.77	4.19	4.21	4.18	4.48	4.53	4.50
OptM		$\boldsymbol{MSE(\hat{\mu})}$	**38.19**	**39.45**	**38.93**	**38.02**	**37.32**	**37.97**	**36.39**	**36.24**	**36.39**
		$\hat{\sigma}$	3.94	3.91	3.87	4.27	4.30	4.26	4.56	4.60	4.57
MoptM		$\boldsymbol{MSE(\hat{\mu})}$	**38.18**	**39.45**	**38.94**	**38.02**	**37.32**	**37.97**	**36.39**	**36.24**	**36.39**
		$\hat{\sigma}$	3.94	3.91	3.87	4.27	4.30	4.26	4.55	4.60	4.57
MM		$MSE(\hat{\mu})$	38.92	40.55	39.87	38.71	38.12	38.91	36.57	36.41	36.59
		$\hat{\sigma}$	2.37	2.34	2.33	2.65	2.67	2.65	2.87	2.90	2.88
DCML		$MSE(\hat{\mu})$	41.46	40.07	41.10	37.88	38.37	37.74	36.77	36.59	36.54
		$\hat{\sigma}$	**0.09**	**0.03**	**0.02**	**0.06**	**0.02**	**0.01**	**0.02**	**0.01**	**0.003**
Mean	5% Block Concentrated	$MSE(\hat{\mu})$	54.59	60.20	57.74	57.35	57.82	57.66	56.52	56.73	57.01
		$\hat{\sigma}$	6.80	7.51	7.20	7.42	7.41	7.41	7.46	7.46	7.47
Median		$MSE(\hat{\mu})$	42.83	43.42	43.08	42.07	42.67	42.53	40.40	41.02	41.23
		$\hat{\sigma}$	3.46	3.44	3.43	3.95	4.00	4.00	4.48	4.49	4.51
HuberM		$MSE(\hat{\mu})$	42.02	43.02	42.41	41.36	41.79	41.69	40.58	40.77	40.98
		$\hat{\sigma}$	3.67	3.67	3.64	4.03	4.08	4.08	4.38	4.39	4.40

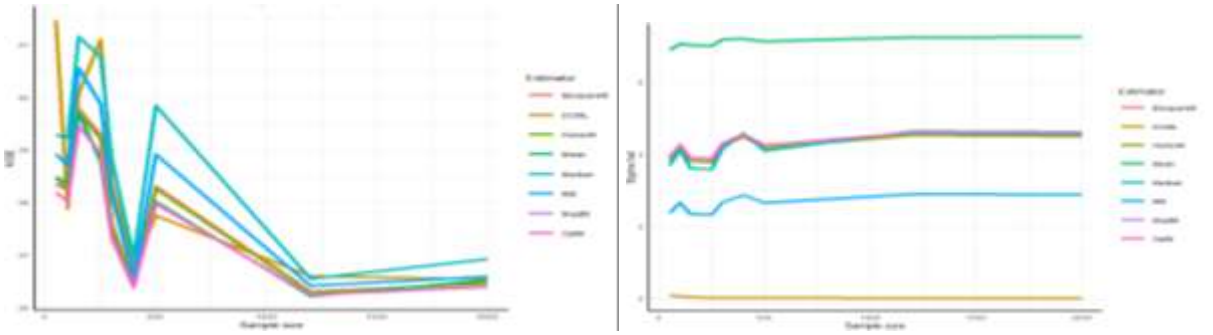
BisquareM		$MSE(\hat{\mu})$	39.51	39.61	39.37	37.56	37.96	37.86	36.39	36.59	36.77
		$\hat{\sigma}$	3.65	3.65	3.62	4.02	4.08	4.07	4.37	4.38	4.39
OptM		$\boldsymbol{MSE(\hat{\mu})}$	**39.20**	**39.15**	**38.99**	**37.22**	**37.65**	**37.49**	**36.31**	**36.46**	**36.65**
		$\hat{\sigma}$	3.75	3.75	3.72	4.13	4.17	4.16	4.45	4.46	4.47
MoptM		$\boldsymbol{MSE(\hat{\mu})}$	**39.20**	**39.15**	**39.00**	**37.22**	**37.65**	**37.49**	**36.31**	**36.47**	**36.65**
		$\hat{\sigma}$	3.75	3.75	3.72	4.12	4.16	4.16	4.45	4.46	4.47
MM		$MSE(\hat{\mu})$	39.84	40.35	39.88	37.83	38.33	38.21	36.45	36.75	36.94
		$\hat{\sigma}$	2.31	2.25	2.26	2.56	2.60	2.59	2.81	2.82	2.82
DCML		$MSE(\hat{\mu})$	41.19	42.34	43.03	42.03	41.52	42.16	16.63	26.74	17.78
		$\hat{\sigma}$	0.09	0.03	0.02	0.06	0.02	0.01	0.01	0.003	0.002
Mean	10% Block Concentrated Vertical Outliers	$MSE(\hat{\mu})$	82.40	84.32	83.98	83.33	82.24	83.20	81.35	81.20	81.46
		$\hat{\sigma}$	9.66	9.63	9.61	9.68	9.67	9.65	9.71	9.70	9.71
Median		$MSE(\hat{\mu})$	45.10	46.54	46.16	46.95	46.55	47.52	45.07	44.77	45.10
		$\hat{\sigma}$	3.80	3.70	3.69	4.26	4.28	4.21	4.71	4.77	4.74
HuberM		$MSE(\hat{\mu})$	46.42	47.68	47.32	47.32	46.53	47.18	45.87	45.88	45.96
		$\hat{\sigma}$	4.01	3.98	3.94	4.36	4.37	4.32	4.66	4.70	4.68
BisquareM		$MSE(\hat{\mu})$	38.85	40.10	39.73	38.60	37.93	38.54	36.49	36.41	36.54
		$\hat{\sigma}$	3.95	3.92	3.88	4.32	4.33	4.30	4.62	4.66	4.64
OptM		$MSE(\hat{\mu})$	38.35	39.54	39.21	38.17	37.47	38.04	36.34	36.26	36.39
		$\hat{\sigma}$	4.03	4.01	3.97	4.38	4.41	4.37	4.68	4.72	4.70
MoptM		$\boldsymbol{MSE(\hat{\mu})}$	**38.35**	**39.54**	**39.20**	**38.17**	**37.47**	**38.04**	**36.34**	**36.26**	**36.40**
		$\hat{\sigma}$	4.03	4.01	3.97	4.38	4.41	4.37	4.68	4.72	4.70
MM		$MSE(\hat{\mu})$	38.97	40.51	39.98	38.82	38.29	39.00	36.55	36.43	36.61
		$\hat{\sigma}$	2.30	2.26	2.26	2.58	2.58	2.57	2.80	2.82	2.81
DCML		$MSE(\hat{\mu})$	47.01	45.73	46.79	45.03	46.02	45.61	45.02	4.57	43.43
		$\boldsymbol{\hat{\sigma}}$	**0.11**	**0.03**	**0.02**	**0.06**	**0.02**	**0.01**	**0.02**	**0.001**	**0.003**



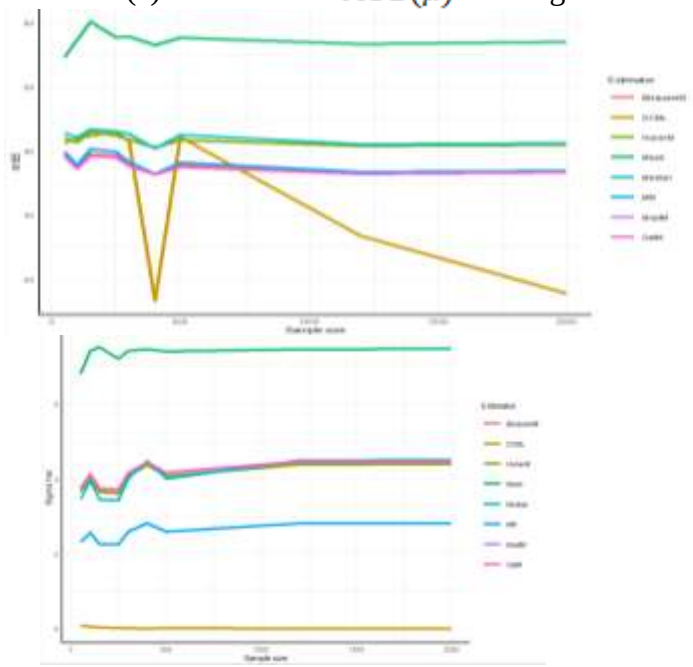
(a) Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having 0% outliers



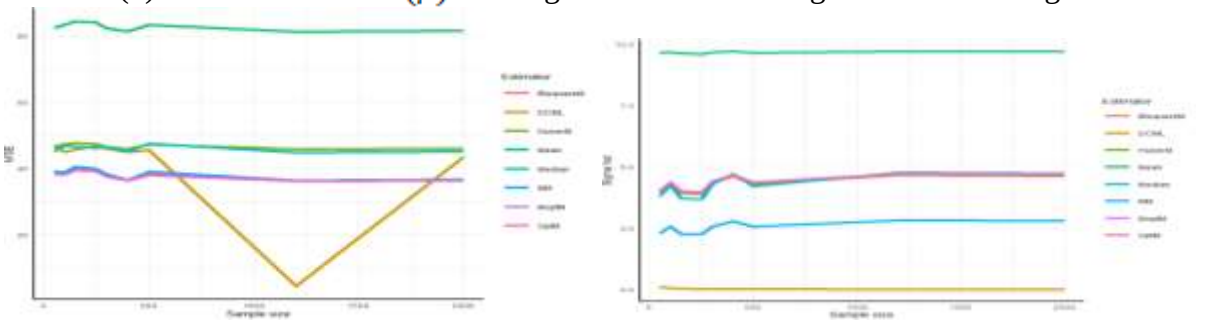
(b) Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having 5% RVO



(c) Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having 10% RVO



(d) Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having 5% BCVO



(e) Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having 10% BCVO

Figure 1: Line chart of $MSE(\hat{\mu})$ and Sigma hat for Centering estimators having various outliers

Table 2: The location and scale estimates for the imports panel-data model having 0%, 5% and 10% Outlier Contamination and GDP growth panel data

Centering Estimator	0% Outliers		5%BCVO		5%RVO		10%BCVO		10%RVO		GDP Growth	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
Mean	26.027	1.607	26.608	3.044	26.740	3.319	27.287	3.872	27.417	4.618	4.352	6.432
Median	25.698	1.749	26.014	2.044	26.106	2.044	26.324	2.348	26.296	2.171	5.291	3.225
HuberM	26.035	1.909	26.182	2.059	26.226	2.069	26.444	2.287	26.401	2.239	4.817	3.411
BisquareM	26.008	1.902	26.009	2.021	26.038	2.025	26.055	2.209	26.090	2.168	4.813	3.413
OptM	26.027	1.908	26.021	2.029	26.051	2.030	26.021	2.218	26.118	2.183	4.687	3.450
MoptM	26.027	1.908	26.021	2.029	26.050	2.030	26.021	2.218	26.118	2.183	4.688	3.450
MM	25.940	1.279	25.940	1.289	25.957	1.278	25.958	1.300	25.941	1.282	5.237	2.061
DCML	**26.009**	**0.074**	**26.014**	**0.062**	**26.106**	**0.030**	**26.324**	**0.016**	**26.296**	**0.012**	**4.993**	**3.562**

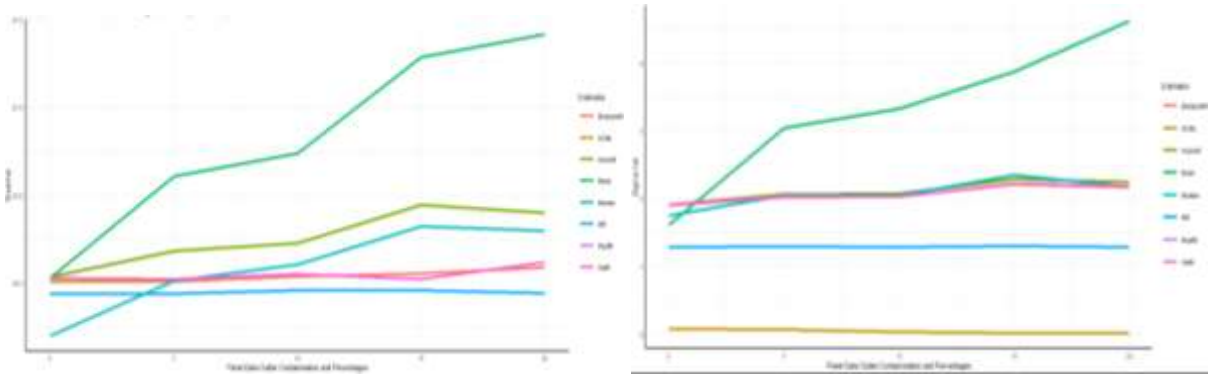


Figure 2: Line chart of Average MSE for Sample Sizes of Panel Data Centering Methods for the lnimports panel-data model having 0%, 5% and 10% Outlier Contamination

Footnote: For the horizonal axis of figure 2, 0 = 0% outliers; 5 = 5% BCVO; 10 = 5% RVO; 15 = 10% BCVO; 20 = 10% RVO.

IV. DISCUSSION OF RESULTS

The results of the analyses are presented in Tables 1 for the simulated finite sample panel data having 0%, 5% and 10% outlying observations. The analysis compares eight centering estimators (Mean, Median, HuberM, BisquareM, optM, MoptM, MM and the proposed centering method DCML) at different sample sizes considered as a product of the number of cross-sectional units N and the time period T (i.e., $n = N \times T$) using the mean square error (MSE) of the estimated location parameter and the scale estimate. When there is 0% outlying observation in the dataset. The results shows that the eight location (centering) estimator give a relative tendency to be consistent at large sample sizes as expected. But for finite samples, the Mean centering estimator yields the lowest average MSE, indicates clearly that the Mean centering estimator out-performed other estimators for small samples sizes of less than or equals to 2000. The superiority of the Mean centering in the uncontaminated panel data is more pronounced at the sample sizes of the less than or equals to 300 but at sample sizes above 300, the others estimators are gradually becoming more consistent. For the scale estimate proxied by the sigma hat, the proposed DCML estimator outstandingly out-performed the other location estimators as its estimated sigma values are far lower than that of the others existing location-scale estimators; taking values of 0.083 for small sample size of 50 and lowering down to 0.003 for sample size of 2000. The MM estimator was the second in the rank of performance with low value of

2.273 for small sample size of 50 to 2.817 for sample size of 2000. It is also observed that the Mean estimator is ranked third in the performance. The Mean estimator yields lower scale value than the four M-estimator (HuberM, BisquareM, optM and MoptM) and the Median. At smaller sample sizes of less than 1000, the Median estimator performed better than the four M-estimators; but is overtaken as the sample size increases more than 1000 as its scale values are higher than those of the four M-estimators.

For panel data linear model having 5% random vertical outlier (RVO), the MSE of the location estimate and scale estimates were provided for the eight-centering estimators. The optM and MoptM location estimators perform better when the sample sizes were less than 300. The DCML estimator yields outstanding lower value for sample sizes between 300 and 1000. And the eight estimators were consistently converging to the parameter for sample size greater than 1000. On the comparison of the scale estimates, the DCML estimator outperformed the other seven estimators. This is followed by the MM-estimator. The Median ranked third and performed better than the M-estimators for the sample sizes less than 1000. But when the sample sizes increase more than 1000, the M-estimators outperformed the Median. The Mean worst-performed in the scale estimate having value lying between 4.905 for n equals 50 to 5.724 for n equals 2000.

The results of the $MSE(\hat{\mu})$ and scale estimate ($\hat{\sigma}$) for the panel data linear model having 10% RVO contaminated data show that the optM and the MoptM estimators outperformed the other estimators for all sample sizes (small and large). It is also observed that the family of M-estimators performed better than the DCML and the MM-estimators in this regard. The Median gave the worst performance in the finite sample and large sizes combined. For the performance of the estimators for the scale, the DCML estimator yielded the best performance, followed by the MM-estimators. Third in the rank of the performance is the Median (though for sample sizes less than 1000). The M-estimator occupy the fourth rank in performance. And the mean performed the least with the scale values of 7 and above values. Worthy to note is that the M-estimators outperformed the median as the sample sizes increased above 1000.

The results of the MSE of the location estimates and scale estimates for the eight estimators for the simulated panel data having 5% block concentrated vertical outlier. shows that for very small sample sizes of less than 300, the optM and MM estimators outperformed other estimators. These were followed by the DCML estimators. The DCML estimator overtook the optM-estimator and MM estimator as the sample size increase between 300 to 500. The result of the performance of the estimator for the scale estimate show that the DCML estimator yield the lowest scale value for 0.09 to 0.002. The MM estimator followed closely in rank with values from 2.254 to 2.819. The M-estimators and the Median are within the same values range with the Median performing relatively better at smaller sample sizes of less than 1000. When the sample size becomes greater than 1000, the four M-estimators did better than the Median. The M-estimator and Median scale estimates lie between 3.462 to 4.506. The Mean values of the scale estimate ranked the last with values between 6.796 to 7.468.

In the results for the 10% BCVO contaminated finite sample panel data, three of the family of M-estimators (optM, MoptM and BisquareM) have the lowest MSE values between 38.347 for sample size of 50 to 36.395 for sample size of 2000. These estimators were followed closely by the MM estimators having values from 38.969 for sample size of 50 to 36.605 for sample size of 2000. The Median and HuberM estimators performed relatively similar with values between 45.098 for sample size of 50 to 45.958 for sample size of 2000. The Median slightly outperformed the HuberM in small sample sizes as well as large sample sizes. The Mean as expected had very large MSE value lying between 82.403 for sample size of 50 to 81.460 for sample size of 2000. The Mean least performed for the 10% BCVO contamination for the scale estimate, the DCML estimator outperformed the other seven estimator with values for lower in the range of 0.109 for sample size of 50 to 0.003 for sample size of 2000. The MM estimator followed in rank with values between 2.299 to 2.806 as the sample size increases from 50 to 2000. The median and the four M-estimators performed relatively the same with value between 3.700 to 4.720 with values as high as 9.708 as the sample size increased from 50 to 2000.

Table 2 and Figures 2 give the results of the location and scale estimates for the lnimports finite sample panel datasets1 having 0%, 5% RVO, 5% BCVO, 10% RV0 and 10% BCVO outliers' contamination; and also, the GDP growth finite sample panel datasets2. It is observed that for the

location estimate, the Mean estimator yields closely similar values the M-estimators, (HuberM, BisquareM, optM and MoptM) and estimator for the uncontaminated data. Though the MM estimate is a bit closer to the M-estimate but the Median estimate is far away from the other estimators. At the 5% BCVO, the mean estimate gets farther away from the other estimators and the Median estimate became closer to the estimates of the BisquareM, optM, MoptM and DCML. The MM estimate is slightly lower than those of the BisquareM, optM, MoptM, MM and DCML. The HuberM estimate move above the estimates of the other family of M-estimators. At the 5% RVO, the behaviour of the Mean estimate continued to soar far higher than others estimate. The median and HuberM estimates move closer to each other. The BisquareM, optM, MoptM and DCML estimates become bunched together. The MM estimates came closer to the group of the optimal M estimate. At the 10% BCVO and 10% RVO, the MM and DCML estimates remain relatively stable, but the M-estimators estimates and the median estimates fluctuated slightly above the 25 values. The MM estimates move slightly lower than the 25 value. But the mean estimate diverges from the value of 25. For the scale estimate, the median, HuberM, BisquareM, optM, MoptM estimators behave relatively similar with the median having the scale value of 1.749 for no outliers, 2.044 for 5% BCVO and 5% RVO, 2.348 for 10% BCVO and 2.171 for10% RVO the M estimators has scale estimate lying between 1.90 to 2.218 for no outliers to 10% BCVO. The MM scale estimator lies between 1.279 and 1.300 for 0% outliers and 10% BCVO. The DCML estimator lie between 0.074 for no outlier to 0.012 10% RVO the DCML estimator outperformed the other estimators followed by the MM-estimators in the second rank, the MoptM is collectively ranked the third, optM is ranked the fourth, BisquareM is ranked the fifth, the HuberM is ranked the sixth, the Median ranked seven and the Mean ranked eighth. Though the M-estimators and the Medians performed relatively similar the mean is far away from the estimate of the M estimator. The result of the GDP growth dataset shows that the MM estimator outperformed the other estimators, followed by DCML estimator. The optM and MoptM that have equal location and scale estimates and they are ranked third.

V. CONCLUSION AND RECOMMENDATIONS

The following findings are made from the results of the study:

a) For the simulated finite sample panel data having 0% outlying observations, the Mean centering estimator outperformed the other seven estimators for small samples sizes of less than or equals to 2000 as its average MSE values is the lowest for the various sample sizes considered. The superiority of the mean centering in the uncontaminated panel data is in tandem with results of (Huber (1964); Aquaro and Cizek (2010) and Midi and Mohammad (2018)). For the scale estimate, the proposed DCML estimator outstanding out-performed the other location-scale estimators as its estimated sigma values are far lower than that of the others existing location-scale estimators; The MM estimator was the second in the rank of performance. The mean estimator is ranked third in the performance. The mean yielded lower scale value than the four M-estimators (HuberM, BisquareM, optM and MoptM) and the Median estimator. The result obtained by Midi and Mohammad (2018) in robust estimation of fixed and random effect panel data model analysis comparing three centering estimators (Mean, Median and MM) corroborates with this result.

b) The optM and the MoptM estimators yield the best performing location estimators in the simulated finite sample panel data linear model having 5% random vertical outlier (RVO). These are followed by the proposed DCML location estimator. On the comparison of the scale estimate, the proposed DCML estimator outperformed the other estimator. This is followed by the MM-estimator. The median ranked third and performed better than the M-estimators. The mean worst-performed in the scale estimation of this scenario.

c) For 10% RVO contaminated simulated panel data, the optM-estimator outperformed the other estimators in the location parameter estimation for all sample sizes; as its MSE values of the location estimate is the lowest for all the sample sizes considered. It is also observed that the other family of M-estimators (MoptM, BisquareM and HuberM) performed better than the DCML and the MM-estimators in this regard. This result is as expected as the M-estimators are designed to perform better in y-directional outlying as considers in the location model. For the performance of the estimators for the scale, the DCML estimator yielded the best performance, followed by the MM-estimators. Third in the rank of the performance is the Median (though for

sample sizes less than 1000); the M-estimators outperformed the Median as the sample sizes increased asymptotically. The M-estimator occupy the fourth rank in performance. And the Mean performed the least with very high scale values.

d) For the simulated panel data having 5% block concentrated vertical outlier (BCVO). the optM and MoptM estimators outperformed other estimators in the location MSE estimates. These were followed closely by the MM estimators. For the scale estimate, the DCML estimator yield the lowest scale value; this is followed by the MM estimator. The Median ranked fourth, performed slightly better than the four M-estimators. The Mean values of the scale estimate ranked the last.

e) The result for the 10% BCVO contaminated finite sample panel data. Show that the three of the family of M-estimators (optM, MoptM and BisquareM) have the lowest MSE values. These estimators were followed closely by the MM estimators. The DCML, Median and HuberM estimators performed relatively similar with the Median slightly outperformed the HuberM in small sample sizes as well as large sample sizes. The Mean as expected had very large MSE value and poorly performed and is ranked the last. For the scale estimate, the DCML estimator outperformed the other seven estimator with range of values far lower than the others. The MM estimator followed in rank. The Median and the four M-estimators performed relatively the same with ranks of 3rd to 7th; while mean as expected performed very poorly.

f) The real-life economic variables datasets illustration reveals that the proposed DCML location estimator is stable as outliers increases and varies in nature. Its performance is quite similar to that MoptM, optM and BisquareM location estimators as well as that of MM location estimator. The Mean and HuberM location estimators though they yielded similar location values at 0% contamination, but moved away from the true location value as the percentage of outliers increases and with varying nature of the outlier. The median location value was far lower than the true location value at 0% contamination and behaved similar to the HuberM estimator at increasing percentage of outliers and varying nature of the outliers. In terms of their scale performance, the proposed DCML estimator was more stable than others and is ranked first. The mm estimator followed in performance of low scale and stable response to increase in outliers and varying nature of the outlier. The four M-estimators and the median performed relatively the same. The mean as expected diverged as the percentage of outliers increases and with varying nature of the outlier.

After extensive Monte Carlo simulation studies and the real-life variables analysis of panel data linear regression models, the following conclusions are drawn from the study:

1. In the panel data linear regression analysis, the strength and weakness of any robust estimator depends on the sample size, percentage of outliers present in the data, the nature of the panel data outlier and direction of the outlier.

2. The MoptM, optM, MM and the proposed DCML estimators yield the best robust centered data in the various panel data contaminations.

3. The proposed DCML estimator yields the best robust scale estimator for any type of panel data contamination, irrespective of the percentage of contamination

Based on the results obtained from this research, these recommendations are put forward:

a) The proposed DCML location-scale estimator is highly recommended in finite panel data with ten percent or more outliers, given its capability to give highly efficient results despite the percentage of outliers.

b) Any of these location estimators (MoptM, optM, MM) can also be used to get robust centered data in the various panel data contamination in infinite samples.

REFERENCES

Aquaro, M. and P. Čížek (2010). One-step robust estimation of fixed effects panel data models. Computational Statistics & Data Analysis 57 (1), 536–548.

Bakar N. M. A. and Midi H. (2015). Robust Centering in the fixed effect panel data model. Pak. J. stat., 31:33-48

Baltagi, B. H. (2021). Econometric Analysis of Panel Data, 6th edition. Springer Nature: Switzerland.

Beaton, A. E., & Tukey, J. W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics, 16, 147–185.

Bramati, M. C. and C. Croux (2007). Robust estimators for the fixed effects panel data model. Econometrics Journal 10 (3), 521–540.

Chaudhuri P. and SenGupta D. (1993). A note on robust estimation of location, Statistics & Probability Letters 18 (1993) 241-244

Croissant Y. and Millo G. (2019). Panel Data Econometrics with R. John Wiley & Sons, Inc., New York, USA.

Eke E. M. (2016). Evaluation of Efficiency of M-estimators in Balanced Incomplete Block Designs. An unpublished M.Sc. Thesis, Imo State University, Owerri, Nigeria.

Hsiao Cheng (2022). Analysis of Panel Data, 4th Edition. Cambridge University Press. United Kingdom

Huber P. J. (1964). Robust estimation of a location parameter. Annals of Mathematical Statistics 35, 73–101.

Ismaeel S. S. and Midi H. (2018). Robust Within Group Estimator for Fixed Effect Panel Data. Pakistan Journal of Statistics Vol. 34(4), 297-310

Maronna, R.A. and Yohai, V.J. (2014), High finite-sample efficiency and robustness based on distance-constrained maximum likelihood. Computational Statistics and Data Analysis, 83, 262–274.

Maronna, R. A., Martin, R. D., Yohai, V. J. and Salibi´an-Barrera, M. (2019). Robust Statistics: Theory and Methods (with R). Wiley, New York.

Martin R. D. (2018). Robust Statistics for Quantitative Finance. Online Tutorial and Lecture Note. www.rinfinance.com

Martin R. D. (2021). Optimal Bias Robust Regression Psi and Rho Revisited. Electronic copy available at: https://ssrn.com/abstract=3902862

Martin, R. D. and Zamar, R. H. (1993). Efficient-Constrained Bias-Robust estimation of Location. The Annals of Statistics. Vol 21, No. 1,338-354.

Midi H. and Mohammad S. (2018). Robust estimation for fixed and random effects panel data models with different centering methods. Journal of engineering and applied sciences 13(17):7156-7161.

Rousseeuw, P. J. (1984). Least median of squares regression. Journal of American Statistical Association, 79, 871-880.

Schoonhoven, M; Nazir, H. Z. Riaz, M. and Does R. (2011). Robust Location Estimator for the $\bar{X}$ Control Chart, Journal of Quality Technology, www.asq.org.

Shu, Ven-Shion (1978). Robust estimation of a location parameter in the presence of outliers. Retrospective Theses and Dissertations. 6419.

Svarc, M., Yohai, V. J. and Zamar, R. H. (2002). "Optimal Bias Robust M-Estimates of Regression". In: Statistical Data Analysis Based on the L1-Norm and Related Methods. Springer, pp. 191–200.

Yohai V. J. (1987). High Breakdown Point and High Efficiency Robust Estimates for Regression, Annals of Statistics, vol. 15, pp. 642–656.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data, 2nd Edition, The MIT Press, London

Zhang Jian and Li Guoying (1998). Breakdown Properties of Location M-Estimators. The Annals of Statistics, Vol. 26, No. 3, 1170–1189