

# Two-Way Cluster-Robust Standard Errors—A Methodological Note on What Has Been Done and What Has Not Been Done in Accounting and Finance Research

Lan Sun<sup>1</sup>, Yueh-Hsia Huang<sup>2</sup>, Tyng-Bin Ger<sup>3</sup>

<sup>1</sup>UNE Business School and School of Business & Law, Central Queensland University, Sydney, Australia

<sup>2</sup>Department of International Trade, Chinese Culture University, Taiwan

<sup>3</sup>Department of Information Management, Minghsin University of Science and Technology, Taiwan

Email: l.sun@cqu.edu.au, hyx3@faculty.pccu.edu.tw, tbger@must.edu.tw

**How to cite this paper:** Sun, L., Huang, Y.-H. and Ger, T.-B. (2018) Two-Way Cluster-Robust Standard Errors—A Methodological Note on What Has Been Done and What Has Not Been Done in Accounting and Finance Research. *Theoretical Economics Letters*, 8, 1639-1655.

<https://doi.org/10.4236/tel.2018.89105>

**Received:** March 9, 2018

**Accepted:** June 12, 2018

**Published:** June 15, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

There is a widely application of panel data estimation in accounting and finance research. The approach is well accepted, because the pooled panel data provide rich information as compared to either cross-sections or time series data structure. However, within panel data structure, variables of interest are often cross-sectionally and serially correlated and as a result, OLS standard errors would be biased when panel data are used in the regression analysis. Several techniques, for example firm dummy variables, one-way cluster-robust standard errors, Fama-MacBeth procedure, and Newey-West procedure, are documented as a solution in analyzing panel data. These techniques to some extent correct either cross-sectional correlation or serial correlation. None is designed to deal with correlations in two dimensions (across firms and across time). With panel data structure, correlations are more likely to appear in two dimensions with both firm effects and time effects. This study suggests that two-way cluster-robust standard errors approach can correct both cross-sectional correlation and serial correlation and therefore should be considered as a better alternative in handling panel data. Nonetheless, two-way cluster-robust standard errors approach could be biased when applying to a finite sample. This study uses a real data set and constructs an empirical application of the estimation procedures of two-way cluster-robust regression estimation with and without finite-sample adjustment and the results show that finite-sample adjusted estimates are superior to unadjusted asymptotic estimates.

---

---

## Keywords

Panel Data, Two-Way Cluster-Robust Standard Errors, Cross-Sectional Correlation, Serial Correlation, SAS

---

## 1. Introduction

Panel data are characterized by pooling data that combines cross-sectional data on  $N$  spatial units (firms) and  $T$  time periods (years) to produce a data set of  $N \times T$  observations. The use of panel data becomes popular in research for two reasons. First, the pooled panel data provide a rich amount of information. The panel data set can increase the number of data points and decrease the likelihood of an omitted-variable. The panel design has a higher quality and quantity data than that of either cross-sections or time series design as the latter two research designs only consider one dimension [1]; the panel data capture the variation of two dimensions simultaneously [2]. Second, the pooled time series cross sectional design allows for testing the impact of a large number of predictors of the level and changing in the dependent variable within the framework of a multivariate analysis [3]. Also, panel data allow for both variations of a single industry/firm over time and variations of all sampled industries/firms at a given point of time [4]. This study shows why panel data are ideal for the examination of variations of industries/firm within time series.

This study also describes the estimation procedures of two-way cluster-robust regression used in handling a panel data set that is in alignment with [5] and [6]. Like many other statistical methods, the two-way cluster-robust methods are built on asymptotic foundations. Authors [7] point out that in a finite sample with a limited number of clusters, the asymptotic estimates of two-way cluster-robust standard errors are biased downwards and researchers who use this method will tend to over-reject a null hypothesis when it is true. When applying the two-way cluster-robust regression to a small panel data set caution needs to be exercised and researchers need to be aware that finite-sample adjusted estimates are superior to unadjusted asymptotic estimates. As such, this study shows that corrections are necessary for variance-covariance matrix estimation when using a finite sample. This study also outlines several SAS procedures that researchers can execute two-way cluster-robust regression, particularly with corrections for finite sample estimation. The purpose of this paper is to show what has been done in accounting and finance research—a widely application of panel data estimation; to suggest that two-way cluster-robust standard errors approach is a better alternative to correct both cross-sectional correlation and serial correlation when using a panel data set and more importantly the adjustment to variance-covariance matrix in a finite sample estimation which by large has been ignored by contemporary research.

The paper is organized in seven sections. The following Section 2 reviews the

theoretical background and what has been done in accounting and finance literature that are based on panel data structure, and mathematically describes and outline the estimation procedures of two-way cluster-robust regression estimation; Section 3 discusses what has not been done in application. In order to provide valid empirical findings, it is vital to understand the key assumptions and what need to be adjusted for a finite sample. Section 4 shows how to use SAS statistics tool to estimate two-way cluster-robust standard errors, especially with corrections in handling finite sample. Section 5 is an empirical application of the estimation procedures of two-way cluster-robust regression estimation with and without adjustment and compares their relative performance. Section 6 concludes the paper.

## 2. Literature Review

It is well known that OLS standard errors are correct when the error terms are independent and identically distributed (iid). However, within panel data structure, variables of interest are often cross-sectionally and serially correlated. For example, industry-specific shocks may induce correlation between firms in a given industry. Firm-specific shocks may be persistent and induce correlation across time. Moreover, some shocks maybe persistent and common among firms: business cycles will induce correlations between different firms across different years. If this is the case, errors generated from OLS with panel data are more likely to be correlated across firms, such that errors in firm  $i$  at year  $t$  are correlated with errors in firm  $j$  at year  $t$ . At the same time, errors are more likely to be correlated from one period to the next, in such a way that errors in firm  $i$  at year  $t$  are correlated with errors in firm  $i$  at year  $t + 1$ . Therefore, the OLS assumption of independence in regression error term is generally violated by the presence of both cross-sectional and time-series dependence [8]. Moreover, for OLS to be optimal it is important that all the errors have the same variance (homoschedasticity). However, there is a risk of producing a regression with heteroschedastic in the pooled time-series cross-sectional setting because it is assumed that the level of the dependent variable is homogenous across firms and time periods while in the case of panel data the dependent variable may differ between firms [9]. In fact, errors for individual firms belonging to the same group may be correlated, with heteroskedasticity and correlation.

Therefore, OLS standard errors would be biased when panel data are used in the regression analysis. Econometric researchers have worked out several solutions to this problem. First, we can use fixed effects to take into account the unobserved time-invariant heterogeneity, the fixed effects method is primarily useful for testing the variables that vary within firm. It focuses on the within-firm variation but neglects the between-firm variation. [10] suggests when fixed effects may not fully control for within and between cluster correlations, the standard errors assumed errors to be i.i.d. will be invalid. The cluster-robust standard errors do consider the correlations in all dimensions because the

two-way clustering method obtains three different cluster-robust variance matrices from, the firm dimension, the time dimension, and the intersection of the firm and time, respectively. Second, the simplest way is to include dummy variables for each cluster, for example, use firm dummy variables and year dummy variables to account for cross-sectional dependence and time-series dependence. Third, use one-way cluster-robust standard errors (also known as Rogers or Huber-White standard errors) to adjust possible correlations within a cross-sectional dimension or a time-series dimension depending on which dimension is clustered [11] [12] [13].

The one-way cluster-robust standard errors generalize the heteroscedasticity robust standard errors of [14] with observations grouped into several clusters. Fourth, use Fama-MacBeth procedure to adjust possible correlations between observations on different firms in the same year, but not to account for correlations between observations on the same firm in different years [15]. Finally, the Newey-West procedure traditionally is used to account for serial correlations of unknown form in the residuals of a single time-series [16]. Now it has been modified for use in a pooled time-series cross-sectional data set by estimating correlations between lagged residuals in the same cluster (see [17] and [18]). Although the above procedures to some extent correct either cross-sectional correlation or serial correlation, none is designed to deal with correlations in two dimensions (across firms and across time). This is because those techniques often cluster by firm and assume independence across time; or cluster by time and assume independence across firms. Unfortunately, with panel data structure, correlations are more likely to appear in two dimensions with both firm effects and time effects.

### Two-Way Cluster-Robust Standard Errors

An alternative approach—two-way cluster-robust standard errors, was introduced to panel regressions in an attempt to fill this gap. Cameron *et al.* (2011) and Thompson (2011) proposed an extension of one-way cluster-robust standard errors to allow for clustering along two dimensions. In this case, the variance estimate for an OLS estimator is expressed as:

$$V(\hat{\beta}) = V(\hat{\beta})_{\text{firm}} + V(\hat{\beta})_{\text{year}} - V(\hat{\beta})_{\text{white}} \quad (1)$$

where  $V(\hat{\beta})_{\text{firm}}$  and  $V(\hat{\beta})_{\text{year}}$  are the estimate variances that cluster by firm and year (Huber, 1967; Rogers, 1983; and Williams, 2000), respectively, and  $V(\hat{\beta})_{\text{white}}$  is the estimate variance for the “intersection” clusters—the within firm variance. Essentially, the two-way clustering method first obtains three different cluster-robust variance matrices for the OLS estimator from one-way clustering in, the firm dimension, the time dimension, and the intersection of the firm and time, respectively. Then, the first two variance matrices, clustering by firm and year are added together and the third intersection matrix is subtracted in order to correct for double-counting the within-firm variance. In this

manner, two-way clustering is robust to both cross-sectional and time-series dependence.

Consider a typical panel regression is expressed as:

$$y_{it} = X_{it}\beta + \varepsilon_{it} \quad \text{for } i=1, \dots, N; t=1, \dots, T \quad (2)$$

where  $y_{it}$  is a  $T \times 1$  vector of observations on the dependent variable in the  $i$ th group,  $X_{it}$  denotes a  $T \times k$  matrix of observations on the explanatory variables;  $\beta$  is the unknown  $K \times 1$  vector of regression parameters and  $\varepsilon_{it}$  is a  $T \times 1$  vector of error terms; and  $\varepsilon \sim N(0, \sigma^2)$ . So the OLS estimator is:

$$\hat{\beta}_{\text{OLS}} = (X'X)^{-1} X'y \quad (3)$$

And the variance of the OLS estimator is:

$$V(\hat{\beta})_{\text{OLS}} = (X'X)^{-1} (X'\Omega X) (X'X)^{-1} \quad (4)$$

where  $\Omega$  is the unknown error variance matrices, which can be written as:

$$\Omega = E[\varepsilon_{it}\varepsilon_{it}' | X_{it}] \quad \text{or} \quad \Omega = \begin{bmatrix} \sigma_{\varepsilon,11}^2 & \sigma_{\varepsilon,12}^2 & \cdots & \sigma_{\varepsilon,1N}^2 \\ \sigma_{\varepsilon,21}^2 & \sigma_{\varepsilon,22}^2 & \cdots & \sigma_{\varepsilon,2N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\varepsilon,N1}^2 & \sigma_{\varepsilon,N2}^2 & \cdots & \sigma_{\varepsilon,NN}^2 \end{bmatrix} \quad (5)$$

The classical OLS specifies that:

$$E[\varepsilon_{it}] = 0,$$

$$\text{Var}[\varepsilon_{it}] = \sigma^2,$$

$$\text{Cov}[\varepsilon_{it}, \varepsilon_{js}] = 0 \quad \text{if } t \neq s \text{ or } i \neq j.$$

Then the error variance is

$$\Omega = \sigma^2 I_{NT} \quad \text{or} \quad \Omega = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} \quad (6)$$

The above is the inference of an OLS<sup>1</sup> estimator for a classical linear model. Now consider if errors for individuals belonging to the same group may be correlated, with general heteroscedasticity and correlation across firms or across times. If errors for individuals belonging to the same group are correlated across firms and times, then the method of two-way robust cluster estimation is robust. The estimation procedure of two-way robust cluster regression can be described in three steps:

Step 1. OLS regression of  $y$  on  $X$  with variance matrix estimate computed using clustering by firms  $i$ , with  $i$  in  $\{1, \dots, N\}$ , assigning each observation to firm cluster yields the White (1980) heteroscedasticity consistent estimator which is robust to correlation across firms at a moment in time.

<sup>1</sup>It is assumed that OLS standard errors are unbiased when the residuals are independent and identically distributed.

$$V(\hat{\beta})_{\text{firm}} = (X'X)^{-1} (X'\hat{\Omega}X) (X'X)^{-1} \quad (7)$$

$$V(\hat{\beta})_{\text{firm}} = (X'X)^{-1} \left( \sum_{i=1}^N (e_i x_i)' (e_i x_i) \right) (X'X)^{-1} \quad (8)$$

Step 2. OLS regression of  $y$  on  $X$  with variance matrix estimate computed using clustering on years  $t$ , with  $t$  in  $\{1, \dots, T\}$ , assigning each observation to year cluster yields the White (1980) heteroscedasticity consistent estimator which is robust to correlation within a firm across time.

$$V(\hat{\beta})_{\text{year}} = (X'X)^{-1} (X'\hat{\Omega}X) (X'X)^{-1} \quad (9)$$

$$V(\hat{\beta})_{\text{year}} = (X'X)^{-1} \left( \sum_{t=1}^T (e_t x_t)' (e_t x_t) \right) (X'X)^{-1} \quad (10)$$

Step 3. OLS regression of  $y$  on  $X$  with variance matrix estimate computed using clustering on both firms and years  $(i, t)$ , with  $(i, t)$  in  $\{(1,1), \dots, (N,T)\}$ . This is the usual White OLS standard error:

$$V(\hat{\beta})_{\text{white}} = (X'X)^{-1} (X'\hat{\Omega}X) (X'X)^{-1} \quad (11)$$

$\Omega$  is estimated by White's heteroscedastic-consistent covariance matrix by squaring OLS residuals of  $\varepsilon_{it}$

$$\hat{\Omega} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{NT}^2 \end{bmatrix}$$

So that,

$$V(\hat{\beta})_{\text{white}} = (X'X)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T (e_{it} x_{it})' (e_{it} x_{it}) \right) (X'X)^{-1} \quad (12)$$

Thus, a two-way cluster-robust variance matrix by firm and by year is estimated as:

$$\begin{aligned} V(\hat{\beta}) &= V(\hat{\beta})_{\text{firm}} + V(\hat{\beta})_{\text{year}} - V(\hat{\beta})_{\text{white}} \\ &= (X'X)^{-1} \left( \sum_{i=1}^N (e_i x_i)' (e_i x_i) \right) (X'X)^{-1} \\ &\quad + (X'X)^{-1} \left( \sum_{t=1}^T (e_t x_t)' (e_t x_t) \right) (X'X)^{-1} \\ &\quad - (X'X)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T (e_{it} x_{it})' (e_{it} x_{it}) \right) (X'X)^{-1} \end{aligned} \quad (13)$$

### 3. What Has Not Been Done in Application

Since two-way cluster-robust standard regression was introduced, researchers in the fields of accounting and finance have been applied constantly in analyzing panel data. However, there are several application issues that researchers rarely explain in their empirical analyses. One concern is whether two-way clus-

ter-robust standard errors are still valid for a finite sample. A finite sample will rely on the normal distribution in making inferences, for example, using the standard 1.64, 1.96 and 2.58 as critical values. However, the two-way cluster-robust method like many statistical methods are built on asymptotic foundations. A key assumption before implementing two-way cluster-robust standard errors is that the number of clusters goes to infinity, that is,  $\min(G_1, G_2) \rightarrow \infty$ , where there are  $G_1$  clusters in the first dimension of firm and  $G_2$  clusters in the second dimension of time.

So what if clusters in both dimensions are small? [7] points out that in a finite sample with a limited number of clusters, the asymptotic estimates of two-way cluster-robust standard errors are biased downwards and researchers use this method will tend to over-reject a null hypothesis when it is true. Though researchers may call their standard errors as “two-way cluster-robust standard errors”, when applying to a small panel data set caution needs to be exercised. Accordingly, finite-sample adjusted estimates are superior than unadjusted asymptotic estimates and the simplest modification is the following:

$V(\beta)_{\text{firm}}$  is multiplied by  $\frac{G_1}{G_1 - 1} \frac{N - 1}{N - k}$ , where  $G_1$  is the number of firm-clusters,  $N$  is the sample size, and  $k$  is the number of independent variables. When  $N$  becomes large (relative to  $k$ ), this modification is approximately  $\frac{G_1}{G_1 - 1}$ . In a similar vein,  $V(\beta)_{\text{year}}$  is multiplied by  $\frac{G_2}{G_2 - 1} \frac{N - 1}{N - k}$ , where  $G_2$  is the number of time-clusters,  $N$  is the sample size, and  $k$  is the number of independent variables. When  $N$  becomes large (relative to  $k$ ), this modification is approximately  $\frac{G_2}{G_2 - 1}$ .

The above modification<sup>2</sup> could yield very different finite-sample estimates than those asymptotic estimates without modification. For instance, it is very common for accounting and finance studies to have a panel data set with 100 firms for a period of 5 years (500 firm-year observations). If the researcher runs two-way cluster-robust regression by both firm and year, there will be 100 groups in firm-cluster ( $G_1 = 100$ ) and 5 groups in time-cluster ( $G_2 = 5$ ). The modification for  $V(\beta)_{\text{firm}}$  is  $100/(100 - 1) = 1.01$  and for  $V(\beta)_{\text{year}}$  is  $5/(5 - 1) = 1.25$ , thus the overall adjustment will be 2.26.

#### 4. Two-Way Cluster-Robust Regression Using SAS

This section reviews SAS 9.4 (<http://support.asa.com>) and outlines several SAS procedures that researchers can use in estimating two-way cluster-robust standard errors. The first SAS procedure is the GENMOD procedure which does not adjust estimates for a finite sample. The GENMOD procedure fits a generalized

<sup>2</sup>Considering a two-way cluster-robust regression without the adjustment,  $V(\beta)$  is inflated by 2.26 and as a result the  $t$ -statistics (which is the square root of variance) is over-estimated by 1.5, leading to a very different significance level. As such, the researcher will tend to over-reject a null hypothesis when in fact it is true.

linear model and co-variances and standard errors are computed based on the asymptotic normality of maximum likelihood estimators. So, for a finite sample the variance-covariance matrix obtained from the below “PROC GENMOD” procedure needs to have a manual adjustment by multiplying by  $G/(G - 1)$ .

```
PROC GENMOD DATA = MYDATA;  
CLASS IDENTIFIER;  
MODEL DEPENDENT VARIABLE = INDEPEENDENT VARIABLES;  
REPEATED SUBJECT = IDENTIFIER/TYPE = IND;  
RUN;
```

The CLASS statement identifies subjects (clusters) in the input data set. Response from different subjects (clusters) is assumed to be statistically independent, and responses within subjects are assumed to be correlated. If modelling firm-effect and time-effect variables such as firm and year must be listed as IDENTIFIERS. The REPEATED statement invokes the generalized linear estimation method, the option SUBJECT = IDENTIFIER specifies that individual subjects (clusters) are identified by the CLASS statement. The TYPE = IND option specifies that the structure of the correlation matrix used to model the correlation of the response from subjects (clusters) and IND means responses from different subjects (cluster) are statistically independent.

The second SAS procedure is the SURVEYREG procedure which does adjust estimates for a finite sample and this procedure is designed to analyze survey data.

```
PROC SURVEYREG DATA = MYDATA;  
CLUSTER VARIABLE;  
MODEL DEPENDENT VARIABLE = INDEPENDENT VARIABLE;  
RUN;  
PROC SURVEYREG DATA = MYDATA TOTAL = OPTION (RATE =  
OPTION);  
CLUSTER VARIABLE;  
MODEL DEPENDENT VARIABLE = INDEPENDENT VARIABLE;  
RUN;
```

The CLUSTER statement identifies clusters in a panel data sample for example researchers can cluster by firms and years. In handling a finite sample, PROC SURVEYREG procedure can be followed by either TOTOAL = option or the RATE = option. The first TOTAL = option is to input population totals and RATE = option is to input sampling rates and as a result the correction for a finite population is incorporated when computing variance covariance estimates. For example, TOTAL = 1000 option specifies the total in the population is 1000 from which the sample is drawn. The value in the RATE = option must be positive numbers. For example, a sampling rate can be a number between 0 to 1, or it can be a percentage between 0.01% to 100%. PROC SURVEYREG uses the Taylor series expansion theory to estimate the covariance-variance matrix of the estimated regression coefficients [19]. According to SAS 9.2 User’s Guide page 206, the matrix is as follow:

$$r = y - X\hat{\beta} \quad (14)$$

where  $y$  denotes the dependent variable,  $X$  denotes the design matrix, and the  $(h, i, j)$ th element is  $r_{hij}$ . Now compute the covariance-variance matrix:

$$\hat{V} = (X'WX)^{-1} G (X'WX)^{-1} \quad (15)$$

In the above covariance-variance matrix,  $G$  is expressed as:

$$G = \frac{n-1}{n-p} \sum_{h=1}^H \frac{n_h(1-f_h)}{n_h-1} \sum_{i=1}^{n_h} (e_{hi} - \bar{e}_{i.})' (e_{hi} - \bar{e}_{i.}) \quad (16)$$

where  $H$  is the stratum number,  $n_h$  is the number of clusters,  $f_h$  is the sampling rate for stratum  $h$ . The number of  $f_h$  is negligible, unless a unique sample rate is specified. Therefore,  $f_h$  is generally negligible when using the code above.  $n$  is the total number of observation in the sample and  $p$  is the total number of parameters. When input stratum totals, PROC SURVEYREG computes  $f_h$  as the ratio of the stratum sample size to the stratum total; when input stratum sampling rates, PROC SURVEYREG will use values directly from  $f_h$ . Considering  $G$  has a component  $(n-1)/(n-p)$  this can be viewed as an adjustment to finite sample.

## 5. Empirical Application

For the demonstration of how two-way cluster-robust standard errors approach could be biased when applying to a finite sample, this section uses a real data set and constructs an empirical application of the estimation procedures of two-way cluster-robust regression estimation with and without finite-sample adjustment and the results show that finite-sample adjusted estimates are superior to unadjusted asymptotic estimates. The relationship between earnings management and executive compensations has been chosen as an empirical application since the topic is widely tested in accounting and finance research. The objective is to investigate to what extent the aggregate level of earnings management is driven by the executive compensation incentive; and, whether different forms of executive pay will play different roles in shaping earnings management behavior. The analysis is often decomposition structured, with executive compensations decomposed into three tiers: total compensation; fixed remuneration versus at-risk remuneration; and salary, bonus, options, shares and other forms of pay such as long term incentive payments. The association between the magnitude of earnings management and each tier of compensation is examined respectively. The starting point for the sample is the population of all ASX listed firms in the DataStream database including active file, suspended file and dead file with necessary annual accounting and market data from the period of 1999 to 2006. Executive compensation data are obtained from the Connect4 databases with an initial executive compensation data set of 7672 firm-year observations. In order to obtain financial data needed to compute discretionary accruals, executive compensation data (from Connect4) was merged with the accrual estimation sample (from DataStream) by company code and by year. The intersection of these two databases and the selection process yielded a testing sample of 3326 firm-year

observations covering the period of 2000 to 2006.

The results of the association between the magnitude of earnings management and executive compensation incentives using two-way cluster-robust regression without a finite sample correction are presented in **Table 1**. The first tier regression reports the association between the magnitude of earnings management and total executive compensations. Results show the coefficient for total compensation (TCOMP) is negative but insignificant, suggesting there is no association between the magnitude of earnings management and total executive compensation. The coefficients on control variables show some significance and the

**Table 1.** Two-way cluster robust regression results for the association between earnings management and executive compensation (without finite sample correction).

Variable	Total Compensation		Fixed v. At-risk		Individual Components	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Intercept	0.2646	(9.8818)***	0.2481	(9.0049)***	0.2467	(9.1625)***
TCOMP	-0.0004	(-0.0707)				
FIX			-0.0467	(-4.5719)***		
ATRISK			0.0246	(3.4720)**		
SALARY					-0.0477	(-3.7910)***
BONUS					0.0307	(3.3612)**
OPTION					0.0165	(2.1249)*
SHARE					0.0534	(1.7515)
LTIP					0.0044	(0.6541)
Control Variables:						
SIZE	-0.0123	(-5.2758)***	-0.0099	(-3.9231)***	-0.0098	(-4.1686)***
GROWTH	0.0005	(5.9328)***	0.0005	(6.2206)***	0.0005	(6.2749)***
ROE	-0.0342	(-3.2096)**	-0.0358	(-3.4598)**	-0.0355	(-3.4906)**
LEV	0.0031	(0.2927)	0.0046	(0.4149)	0.0048	(0.4459)
BM	-0.0071	(-2.1710)*	-0.0076	(-2.5920)**	-0.0076	(-2.5707)**
CIR	0.0173	(1.0823)	0.0159	(1.0111)	0.0167	(1.0735)
LAGTA	-0.0001	(-0.8082)	-0.0002	(-1.4490)	-0.0001	(-1.2885)
Adj. R-square	0.1034		0.1092		0.1111	
N	3326		3326		3326	

This table reports two-way cluster-robust regression results in testing the magnitude of earnings management and its association with executive compensations. The dependent variable is the magnitude of earnings management which is measured as absolute values of discretionary accruals. Explanatory variables are executive compensations which decomposed into three tiers: executive total compensation (TOMP); executive fixed remuneration (FIX) versus at-risk compensation (ATRISK); and, individual components including fixed salary (SALARY), bonuses (BONUS), options (OPTION), shares (SHARE), and long-term incentive plans (LTIP). Firm characteristics and industry effects are controlled. All variables are defined in Appendix. The estimated coefficients and t statistics are two-way cluster-robust without finite sample correction. T-statistics are given in parentheses, one-tailed tests when we have explicit predictions and two-tailed otherwise. \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5% and 1% respectively.

regression has an adjusted R-square of 10.34%. The second tier reports results from the regression of earnings management on fixed compensation and at-risk compensation components. The coefficient on fixed compensation (FIX) is negative and significant at less than 1% level. In contrast, the coefficient on at-risk compensation, including bonus, options, shares and long-term incentive plans (ATRISK), is positive and significant at 5% level after controlling firm characteristics. The third tier reports results from the regression of earnings management on each compensation component. In this stage, compensation is further decomposed into to salary, bonus, options, shares and long-term incentive plans. Now the results show some dynamic relations between the aggregate level of earnings management and individual compensation components. First, the coefficient on salary (SALARY) is negative and significant at less than 1% level. Second, the coefficients on bonus (BONUS) and options (OPTION) are positive, significant at 5% and 10% level, respectively. Based on the above results from two-way cluster-robust regression, the study may claim that executive compensation creates incentives for earnings management behavior. Moreover, the findings indicate a variety of compensation-related incentive effects, with some features encouraging earnings management, and others, discouraging it. Particularly, fixed compensation and salary are more likely to constrain earnings management. However, at-risk compensation and bonuses induce managers to engage in earnings management because at-risk compensation is usually based on earnings performance and managers would opportunistically use discretionary accruals to exploit the nonlinearity in the payoffs on compensation, which is tied to earnings performance.

**Table 2** shows the results of re-estimating the association between the magnitude of earnings management and executive compensation by using two-way cluster-robust regression with finite sample correction. The modification discussed in Section 3 has been applied and this could yield very different finite-sample estimates as compared to those in **Table 1** without any modification. The first tier, executive total compensation remains insignificant, suggesting there is no association between the magnitude of earnings management and total executive compensation. The second tier, the fixed compensation remains significantly negative while the at-risk compensation becomes insignificant. The third tier, the coefficients of salary and bonus remain significantly negative and positive respectively, while option becomes insignificant. To some extent a two-way cluster-robust regression without adjustment tends to inflate the significance level which is evident in **Table 1** and therefore cautious need to be made when suggesting that at-risk compensation and its individual component such as option are more likely to induce managers to engage in opportunistic behavior. In all, a two-way cluster-robust regression without adjusting for finite sample is more likely to inflate the statistics and therefore in a finite sample estimation the adjustment to variance-covariance matrix is crucial to ensure the validity of findings. Nevertheless, we notice the limitation of using two-way cluster robust

errors, there is no way to know whether the correction is exact, too little, or too much. Indeed, Cameron and Miller are also extending their research from two-way to multi-way clustering.

We also re-estimate the association between earnings management and executive compensation using the fixed effects method. **Table 3** shows the results of re-estimating the association between the magnitude of earnings management and executive compensation using fixed effects. The results are very different from that of two-way cluster-robust regression. The first tier, executive total compensation remains insignificant, suggesting there is no association between

**Table 2.** Re-estimate the association between earnings management and executive compensation using two-way cluster robust regression (with finite sample correction).

Variable	Total Compensation		Fixed v. At-risk		Individual Components	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
Intercept	0.2820	(6.7333)***	0.2739	(6.2093)***	0.2673	(4.2364)***
TCOMP	-0.0040	(-0.5026)				
FIX			-0.0297	(-1.9930)*		
ATRISK			0.0109	(0.9453)		
SALARY					-0.035	(-1.7361)*
BONUS					0.0919	(1.6643)*
OPTION					0.0196	(1.3390)
SHARE					-0.0164	(-0.9638)
LTIP					-0.018	(-1.0776)
Control Variables:						
SIZE	-0.0049	-1.5131	-0.0039	(-1.1645)	-0.0036	(-1.0522)
GROWTH	0.0004	0.9417	0.0004	(0.9509)	0.0004	(1.2290)
ROE	-0.0530	(-2.9541)**	-0.0536	(-2.9950)**	-0.0526	(-4.0267)***
LEV	-0.0056	-0.3763	-0.0047	(-0.3072)	-0.0053	(-0.2422)
BM	-0.0065	(-3.3492)**	-0.0068	(-3.7196)	-0.0071	(-1.6477)*
CIR	0.0340	0.7672	0.0334	(0.7539)	0.0348	(1.6359)
LAGTA	0.0008	(7.1064)***	0.0007	(6.0797)***	0.0007	(2.0582)**
Adj. R-square	0.1034		0.1092		0.1111	
N	3326		3326		3326	

This table reports two-way cluster-robust regression in testing the magnitude of earnings management and its association with executive compensations with finite sample correction. The dependent variable is the magnitude of earnings management which is measured as absolute values of discretionary accruals. Explanatory variables are executive compensations which decomposed into three tiers: executive total compensation (TOMP); executive fixed remuneration (FIX) versus at-risk compensation (ATRISK); and, individual components including fixed salary (SALARY), bonuses (BONUS), options (OPTION), shares (SHARE), and long-term incentive plans (LTIP). Firm characteristics, industry and year effects are controlled. All variables are defined in Appendix. The estimated coefficients and t statistics are two-way cluster-robust regression with finite sample correction. T-statistics are given in parentheses, one-tailed tests when we have explicit predictions and two-tailed otherwise. \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5% and 1% respectively.

**Table 3.** Re-estimate the association between earnings management and executive compensation using fixed effects regression.

Variable	Total Compensation		Fixed v. At-risk		Individual Components	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
TCOMP	-0.0002	(-0.5800)				
FIX			-0.0375	(-0.4257)		
ATRISK			0.0481	(0.4720)		
SALARY					-0.0300	(-0.4410)
BONUS					0.0461	(0.3312)
OPTION					-0.0189	(-0.2212)
SHARE					-0.3899	(-1.5171)
LTIP					-0.7415	(-0.7441)
Control Variables:						
SIZE	-0.0123	(-1.2758)	-0.0139	(-1.4031)	-0.0144	(-1.4686)
GROWTH	0.0000	(0.0728)	0.0000	(0.0506)	0.0000	(0.1279)
ROE	0.0156	(1.7796)*	0.0157	(1.7898)*	0.0144	(1.6106)
LEV	0.0165	(0.3927)	0.0164	(0.3907)	0.0170	(0.4059)
BM	-0.0531	(-2.4210)**	-0.0525	(-2.3920)**	-0.0470	(-2.0907)**
CIR	0.0595	(1.9108)*	0.0576	(1.8411)	0.0548	(1.7502)***
LAGTA	0.0027	(0.8882)	0.0027	(0.8990)	0.0023	(0.7528)
Firm Fixed Effects	YES		YES		YES	
Adj. R-square	0.4110		0.4203		0.4212	
N	3326		3326		3326	

This table reports fixed-effects regression results in testing the magnitude of earnings management and its association with executive compensations. The dependent variable is the magnitude of earnings management which is measured as absolute values of discretionary accruals. Explanatory variables are executive compensations which decomposed into three tiers: executive total compensation (TOMP); executive fixed remuneration (FIX) versus at-risk compensation (ATRISK); and, individual components including fixed salary (SALARY), bonuses (BONUS), options (OPTION), shares (SHARE), and long-term incentive plans (LTIP). Firm characteristics and industry effects are controlled. All variables are defined in Appendix. The estimated coefficients and t statistics are estimated using the fixed effects estimation method. T-statistics are given in parentheses, one-tailed tests when we have explicit predictions and two-tailed otherwise. \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5% and 1% respectively.

the magnitude of earnings management and total executive compensation. The second tier, the fixed compensation now becomes insignificantly which is in contrast to the previous results. The third tier, the coefficients for all decomposed variables are insignificant. Again, this is basically inconsistent with previous findings.

The fixed effects method is primarily useful for testing the variables that vary within firm. It focuses on the within-firm variation but neglects the between-firm variation. This is the major concern for using fixed effects. The between-firm variation is very likely to be contaminated by unobserved firm cha-

racteristics that are correlated with managers' decision in exercising discretion over accruals, rather than merely compensation incentives. If we restrict ourselves to the within-firm variation, we are more likely to discard the unobserved firm characteristics, that is, the between-firm variation. As a consequence, the coefficients on the time-invariant variables cannot be estimated and this is the price of the robustness of the specification we need to pay for ignoring unobserved correlation between the common effect and the exogenous variables. Moreover, the choice of estimation method also depends on whether the firm effect is temporary or permanent. If the firm effect dissipates after several years, the effect fixed on firm will no longer fully capture the within-cluster dependence and OLS standard errors are still biased. [20] suggests that the OLS standard errors tend to underestimate the standard errors in the fixed effects regression when the firm effect dies out over time. In this case, it is still necessary to use cluster robust standard errors.

## 6. Concluding Remarks

It is well known that OLS is biased when the residuals are not iid. In accounting and finance literature, many studies are based on panel data samples and the problems with panel data structure variables are often cross-sectionally and serially correlated and thus the residuals are no longer iid. There are various methods for estimating standard errors when the residual are correlated across firms and/or years, for example firm dummy variables, one-way cluster-robust standard errors, Fama-MacBeth procedure, and Newey-West procedure are documented as a solution in handling panel data. These techniques to some extent correct either cross-sectional correlation or time serial correlation. None is designed to deal with correlations in two dimensions, that is, across firms and across time in a panel data structure. In order to provide valid empirical findings, it is vital for researchers to understand the best statistical solution, the appropriate computer procedures, the assumptions and what need to be adjusted for a finite sample. This paper reviews two-way cluster-robust standard errors in panel data studies and mathematically describes the estimation procedures of two-way cluster-robust regression. This paper also discusses the key assumption for two-way cluster-robust standard errors and shows that corrections are necessary for variance-covariance matrix estimation when analyzing a finite sample. Using SAS as a statistical analysis tool, this study also outlines several procedures that researchers can execute two-way cluster-robust regression particularly with corrections for a finite sample estimation. Finally, for the demonstration of how two-way cluster-robust standard errors, approach could be biased when applying to a finite sample. This study uses a real data set and constructs an empirical application of the estimation procedures of two-way cluster-robust regression estimation with and without finite-sample adjustment and the results show that finite-sample adjusted estimates are superior to unadjusted asymptotic estimates. We also compare the two-way cluster-robust estimation with the fixed

effects method. The fixed effects method is primarily useful for testing the variables that vary within firm. It focuses on the within-firm variation but neglects the between-firm variation. This is the major concern for using fixed effects. The between-firm variation is very likely to be contaminated by unobserved firm characteristics that are correlated with managers' decision in exercising discretion over accruals, rather than merely compensation incentives. If we restrict ourselves to the within-firm variation, we are more likely to discard the unobserved firm characteristics, that is, the between-firm variation. As a consequence, the coefficients on the time-invariant variables cannot be estimated and this is the price of the robustness of the specification we need to pay for ignoring unobserved correlation between the common effect and the exogenous variables. Moreover, the choice of estimation method also depends on whether the firm effect is temporary or permanent. If the firm effect dissipates after several years, the effect fixed on firm will no longer fully capture the within-cluster dependence and OLS standard errors are still biased. Nevertheless, we notice the limitation of using two-way cluster robust errors. It is difficult to test whether the correction is exact, too little, or too much and indeed Cameron and Miller are also extending their research from two-way to multi-way clustering.

## References

- [1] Gujarati, D. (2003) *Basic Econometrics*. 4th Edition, McGraw Hill, New York.
- [2] Pennings, P., Keman, H. and Kleinnijenhuis, J. (1999) *Doing Research in Political Science: An Introduction to Comparative Methods and Statistics*. Sage, London, Oaks, New Delhi.
- [3] Schmidt, M.G. (1997) Determinants of Social Expenditure in Liberal Democracies. The Post World War II Experience. *Acta Politica*, **32**, 153-173.
- [4] Pindyck, R.S. and Rubinfeld, D.L. (1991) *Econometric Models and Economic Forecasts*. 2nd Edition, McGraw-Hill Book Co., New York.
- [5] Cameron, A.C., Gelbach, J.B. and Miller, D.L. (2011) Robust Inference with Multi-Way Clustering. *Journal of Business and Economic Statistics*, **29**, 238-249. <https://doi.org/10.1198/jbes.2010.07136>
- [6] Thompson, S.B. (2011) Simple Formulas for Standard Errors that Cluster by Both Firm and Time. *Journal of Financial Economics*, **99**, 1-10. <https://doi.org/10.1016/j.jfineco.2010.08.016>
- [7] Cameron, A.C. and Miller, D.L. (2011) Robust Inference with Clustered Data. In: Ullah, A. and Giles, D.E., Eds., *Handbook of Empirical Economics and Finance*, CRC Press, 1-28.
- [8] Greene, W.H. (2002) *Econometric Analysis*. 5th Edition, Prentice Hall, New York.
- [9] Beck, N. and Katz, J.N. (1995) What to Do (and Not to Do) with Time-Series Cross-Section Data. *American Political Journal Review*, **89**, 634-647. <https://doi.org/10.2307/2082979>
- [10] Cameron, A.C. and Miller, D.L. (2015) A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resource*, **50**, 317-372. <https://doi.org/10.3368/jhr.50.2.317>
- [11] Huber, P. (1967) The Behaviour of the Maximum Likelihood Estimates under

Non-Standard Conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 221-233.

- [12] Rogers, W. (1993) Regression Standard Errors in Clustered Samples. *Stata Technical Bulletin*, **3**, 19-23.
- [13] Williams, R. (2000) A Note on Robust Variance Estimation for Cluster-Correlated Data. *Biometrics*, **56**, 645-646. <https://doi.org/10.1111/j.0006-341X.2000.00645.x>
- [14] White, H. (1980) A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, **48**, 817-838. <https://doi.org/10.2307/1912934>
- [15] Fama, E.F. and MacBeth, J.D. (1973) Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, **81**, 607. <https://doi.org/10.1086/260061>
- [16] Newey, W. and West, K. (1987) A Simple, Positive Semi-Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**, 703-708. <https://doi.org/10.2307/1913610>
- [17] Bertrand, M., Dufol, E. and Mullainathan, S. (2004) How Much Should We Trust Differences-in-differences Estimates? *Quarterly Journal of Economics*, **119**, 249-275. <https://doi.org/10.1162/003355304772839588>
- [18] Doidge, C. (2004) U.S. Cross-Listings and the Private Benefits of Control: Evidence from Dual Class Firms. *Journal of Financial Economics*, **73**, 519-553. [https://doi.org/10.1016/S0304-405X\(03\)00208-3](https://doi.org/10.1016/S0304-405X(03)00208-3)
- [19] Fuller, W.A. (1975) Regression Analysis for Sample Survey. *Sankhya, Series C*, **37**, 117-132.
- [20] Wooldridge, J. (2003) Cluster-Sample Methods in Applied Econometrics. *American Economic Review*, **93**, 133-138. <https://doi.org/10.1257/000282803321946930>

## Appendix: Data Set Construction

The starting point for the sample is the population of all ASX listed firms in the DataStream database including active file, suspended file and dead file with necessary annual accounting and market data from the period of 1999 to 2006. Executive compensation data are obtained from the Connect4 databases with an initial executive compensation data set of 7672 firm-year observations. In order to obtain financial data needed to compute discretionary accruals, executive compensation data (from Connect4) was merged with the accrual estimation sample (from DataStream) by company code and by year. The intersection of these two databases and the selection process yielded a testing sample of 3326 firm-year observations covering the period of 2000 to 2006.

TCOMP: Dollar value of total compensation earned by CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

FIX: Dollar value of fixed compensation earned by CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

ATRISK: Dollar value of at-risk compensation earned by CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

SALARY: Dollar value of base salary earned by CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

BONUS: Dollar value of bonus earned by CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

OPTION: Dollar value of options granted to CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

SHARE: Dollar value of shares granted to CEOs in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

LTIP: Dollar value paid out to CEOs under the company's long term incentive plan in firm  $i$  at fiscal year  $t$ , measured in millions of dollars

SIZE: Firm size for firm  $i$  for year  $t$ , measured by the logarithm of the total assets at year  $t$

GROWTH: Growth opportunity for firm  $i$  for year  $t$ , measured by the change of sales between year  $t$  and  $t - 1$  divided by total assets at year  $t$

ROE: Profitability, measured by net operating income divided by total equity for firm  $i$  at year  $t$

LEV: Leverage, measured by total debt to total assets for firm  $i$  in year  $t$

BM: Book-to-market effect ratio, measured by book value of common equity to market value of common equity for firm  $i$  in year  $t$

CIR: Capital intensity, measured as gross property, plant and equipment divided by total assets for firm  $i$  in year  $t$

LAGTA: Lagged total accruals, measured as the total accruals for firm  $i$  in year  $t - 1$