(F*orthcoming* in *Handbook of Financial Econometrics and Statistics*, Springer)

Chapter 24

Panel Data Analysis and Bootstrapping: Application to China Mutual Funds

Win Lin Chou, Shou Zhong Ng, and Yating Yang

Abstract

Thompson (2011) argues that double-clustering the standard errors of parameter estimators matters the most when the number of firms and time periods are not too different. Using panel data of similar number in firms and time periods on China's mutual funds, we estimate double- and single-clustered standard errors by wild cluster bootstrap procedure. To obtain the wild bootstrap samples in each cluster, we reuse the regressors (X), but modify the residuals by transforming the OLS residuals with weights which follow the popular two-point distribution suggested by Mammen (1993) and others. We then compare them with other estimates in a set of asset pricing regressions. The comparison indicates that bootstrapped standard errors from double-clustering outperform those from single clustering. Our findings support Thompson's argument. They also suggest that bootstrapped critical values are preferred to standard asymptotic t-test critical values to avoid misleading test results.

Keywords: Asset-pricing regression, Bootstrapped critical values, Cluster standard errors, Double clustering, Firm and time effects, Finance panel data, Single clustering, Wild-cluster bootstrap.

Win Lin Chou
Department of Economics and Finance, City University of Hong Kong, Hong Kong
wlchou@cityu.edu.hk

Shou Zhong Ng
Hong Kong Monetary Authority
bszng@hkma.gov.hk
Yating Yang
Institute of Finance, National Chiao Tung University, Taiwan
yatingyang.iof98g@nctu.edu.tw

24.1. Introduction

Researchers using finance panel data have increasingly realized the need to account for the residual correlation across both firms and/or time in estimating standard errors of regression parameter estimates. Ignoring such clustering can result in biased OLS standard errors. Two forms of residual dependence that are common in finance applications are time-series dependence and cross-sectional dependence. The former is called a firm effect, whereas the latter a time effect. The usual solution to account for the residual dependence is to compute clustered standard errors. The notable examples are Petersen (2009), and Thompson (2011).

Using Monte Carlo simulated panel data, Petersen (2009) compares the performance of many different standard error estimation methods surveyed in the literature. These methods include White's heteroskedasticity-robust standard errors, single clustering (by firm or by time) and double clustering (by both firm and time). His findings suggest that the performance of different methods depends on the forms of residual dependence. For example, in the presence of a firm effect, the clustered standard errors are unbiased and can produce correctly sized confidence intervals while those estimated by OLS, White, or Fama-MacBeth method are biased.

Much of the analysis in Petersen (2009) is based on simulated panel data set whose data structure is certain. With simulated panel data set, it is easier to choose among the estimation methods. This paper chooses an alternative method, namely, bootstrapping, to investigate the performance of standard errors estimated by White's OLS, single- and double-clustering methods with actually observed data. The use of the bootstrap method is motivated by Kayhan and Titman (2007) who show that bootstrapped standard errors are robust to heteroskedasticity, and serial correlation problems in panel finance data applications. Moreover, despite the wide use of the bootstrap in statistical and econometric applications, the survey finding of Petersen (2009) found the bootstrap applications are relatively scarce in the finance literature. Hence, it may be of some interest to investigate the bootstrapping application to a set of panel finance data on Chinese mutual funds.

The bootstrap method is applied to a panel data set on the monthly returns for 54 Chinese mutual funds over the period of September 2002 to August 2006. The data set is applied to a set of asset pricing regressions. Table 1 contains summary statistics such as sample skewness, sample excess kurtosis, and two test statistics for normality for the variables used in the asset pricing regressions. They suggest that normality does not characterize the variables. Additionally, since the time-series and/or cross-sectional independence assumption is most likely to be violated in panel datasets, ignoring these dependence could result in biased estimates of the standard errors. As evidenced in Kayhan and Titman (2007), bootstrapping is a possible alternative to handle this dependence issue.

In this paper, we are particularly interested in the performance of the bootstrapped double-clustered standard error estimates, because Thompson (2011) has argued that double-clustering matters most when the number of firms and time periods are not too different. Given the panel data set we have collected which consists of 54 China mutual fund returns for 48 months with data exhibiting firm and time effects, double-clustering is likely to show a significant difference. Our findings show that the bootstrapped standard errors from double-clustering leads to more significant test results. We also demonstrate the importance of using bootstrapped critical values in hypothesis testing.

A number of bootstrap procedures are available in the literature. The bootstrap procedure we consider in this paper is the wild cluster bootstrap procedure, which is an extended version of the wild bootstrap proposed by Cameron, et al. (2008) in a cluster setting. This procedure has been shown by Cameron, et al (2008) to perform very well in practice, despite the fact that the pairs cluster bootstrap works well in principle. In this paper, our comparison of the finite-sample size of the bootstrapped t-statistics resulting from the pairs cluster bootstrap and wild cluster bootstrap also indicates the wild cluster bootstrap performs better.

The rest of the paper is organized as follows. Section 24.2 presents the wild cluster bootstrap procedure. Section 24.3 discusses the empirical results, and the last section gives conclusions.

24.2. Wild Cluster Bootstrap

The bootstrap we use in this paper is known as the “wild cluster bootstrap” which is based on the nonclustered wild bootstrap proposed by Wu (1986). Proofs of the ability of the wild bootstrap to provide refinements in the linear regression model for linear regression models with heteroskedastic errors can be found in Liu (1988) and Mammen (1993). Cameron, et al. (2008) extended Wu’s (1986) wild bootstrap to the clustered setting.

The wild cluster bootstrap procedure involves two stages. In the first stage, we consider the asset pricing model with G clusters (subscripted by g), and with N_g observations within each cluster, namely, $y_g = X_g^{'}\beta + e_g$, $g = 1,...,G$, β is $k \times 1$, X_g is $N_g \times k$, and y_g and e_g are $N_g \times 1$ vectors. We fit the model to the actually observed panel data set by OLS, and estimate White’s heteroskedasticity-robust standard errors, as well as standard errors clustered by firm, by time, and by both. We then save residuals, and denote them as $\hat{e}_g$.

The second stage is the resampling procedure which creates bootstrap-samples for each cluster, $\{(\hat{y}_1^*, X_1),...,(\hat{y}_G^*, X_g)\}$ where $\hat{y}_g^* = X_g'\hat{\beta} + e_g^*$. For each bootstrap-sample in a cluster, the explanatory variables are reused unchanged. The residuals e_g^* are constructed according to $e_g^* = a_g \hat{e}_g$, where the weight a_g serves as a transformation of the OLS residuals $\hat{e}_g$. A variety of constructions of weights a_g have proposed in the literature[1]. We use the two-point distribution of the weight variable a_g suggested in Mammen (1993), Brownstone and Valletta (2001), and Davidson and Flachaire (2008), namely, a_g takes on one of the following values: (i) $(1-\sqrt{5})/2 \approx$ -0.6180 with probability $(1+\sqrt{5})/(2\sqrt{5}) \approx$0.7236; or (ii) $(1+\sqrt{5})/2 \approx$1.6180 with probability $1-(1+\sqrt{5})/(2\sqrt{5}) \approx$0.2764. Note that this random variable a_g has a mean zero with variance equal to one and the constraint $E(a_g^3) = 1$. We perform 1000 replications. On each replication, a new set of e_g^* is generated and a new set of bootstrap-data is created based on $\hat{y}_g^* = X_g'\hat{\beta} + e_g^*$, and therefore a new set of parameter estimates, denoted as $\hat{\beta}^*$ is obtained.

For the 1000 starred estimates (e_g^*), we calculate their bootstrapped standard errors using different estimation methods. The bootstrapped test statistics are calculated by dividing the 1000 parameter estimates by the corresponding

---

[1] For example, in Cameron, et al. (2008), a_g takes the value +1 with probability 0.5, or the value -1 with probability 1-0.5.

bootstrapped standard errors. The bootstrapped critical values can be obtained from the bootstrapped distribution of these test statistics. A detailed explanation of the procedure we follow is documented in Appendix 24A.

24.3. Empirical Results

24.3.1 Data and Definitions of the Variables

The Data. The sample consists of the returns on 54 publicly traded closed-end mutual funds that are gathered for 48 months from September 2002 to August 2006, a total of 2592 observations. The mutual fund data set is purchased from the GTA Information Technology Company, Shenzhen, China. For simplicity, we divide the mutual funds investment objectives into equity growth and non-growth funds. Our sample consists of 37 (68.5%) growth funds, and 17 (31.5%) non-growth funds. Although the first closed-end fund in China was sold to the public in April 1998[2], complete data for all 54 mutual funds are collected from September 2002.

The summary statistics for the variables used in China's mutual fund return regressions are displayed in Table 24.1, and the test statistics for normality for these variables suggest that the variables are non-normal.

Definitions of the variables. The following are the definitions of the variables used in the estimation procedures: $R_{i,t}$ = the return of a mutual fund i in excess of the risk-free rate in month t. Savings deposit rate of the People's Bank of China[3] is used as the proxy for the risk-free rate. $R_{m,t}$ = is the market return in excess of the risk-free rate in month t, and is calculated[4] as follows:

---

2 Chen and Lin (2006), p.384.
3 Data are taken from the website of the People's Bank of China: http://www.pbc.gov.cn/publish/zhengcehuobisi/627/index.html
4 The calculation follows that of Shen and Huang (2001), p.24.

$$R_{m,t} = 0.4 \times R_{1,t} + 0.4 \times R_{2,t} + 0.2 \times 0.0006, \qquad (24.1)$$

where R_1 is the monthly return on Shanghai Stock Exchange index, R_2 the monthly return on Shenzhen Stock Exchange index, and 0.06% is the monthly return on savings deposits.

24.3.2 Results

Table 24.2 presents the results from a set of asset pricing regressions of China mutual fund returns on its market returns. The firm and time effect in OLS residuals and data can be seen graphically in Figure 24.1. Figure 24.1, Panel A, shows the within-firm autocorrelations in OLS residuals and independent variable, respectively, for lags 1 to 12. Panel B of Figure 24.1 displays the within-month autocorrelations for residuals for lags 1 to 12. As the independent variable is a monthly series without cross-sectional units, we cannot calculate its within-month autocorrelations, thus no within-month plot is available for the independent variable. Figure 24.1 suggests that the residuals exhibit both firm and time effects, whereas independent variable shows firm effects. Given Thompson's (2011) argument that double-clustering matters most when the number of firms and time periods are not too different, our data set which has the number of mutual funds (54) similar to the number of months (48) is expected to imply that double-clustering is important in our analysis.

In Table 24.2, the second column presents the OLS parameter estimates, whereas the remaining columns report their corresponding t-statistics by dividing each parameter estimate by its corresponding standard error. These t-statistics indicate all beta coefficients are statistically significant at the 1% level, whereas the intercept is only significant in one case (under column II) when the standard error computed by single-clustering by firm is used. More importantly, the t-statistics in Table 24.2 enable us to compare the clustered standard errors constructed from double and single clustering with the OLS White estimate. Notice that the t-statistic for beta coefficient obtained from double-clustering ($\hat{SE}_{both}$) is 9.10 (column IV) which is much smaller than 46.9 calculated by White method (column I), indicating the presence of firm and time effects. It also means that the double-clustering standard errors are much larger. A comparison of t-statistics in columns III and IV implies the $\hat{SE}_{both}$ of the beta coefficient (9.10) is similar to the standard error clustered by time which is 9.01. This means that the firm effects do not matter much. The comparison reveals that OLS White standard errors are underestimated when residuals exhibit both firm and time effects.

Turning to the results obtained by the wild-cluster bootstrapping, the story changes. The bootstrapped t-statistics of the beta coefficient estimates displayed in columns V to VIII of Table 24.2 differ quite significantly from those in columns I to IV. The t-statistic is 47.2 when the bootstrapped White standard error is used and 66.8 if the bootstrapped standard error clustered by firm is used. This means the firm effect is significant in the data. By a similar comparison, the bootstrapped t-statistic is 73.7 when the double-clustered standard error is used meaning both the time and firm effects are strong in the residuals. The bootstrapped t-statistic is 50.9 when the bootstrapped standard error clustered by time is used implying the time effect exists in the data. These comparisons suggest both the firm and time effects matter in the computation of the bootstrapped standard errors using double as well as single clustering. Their implication is that we might follow what Kayhan and Titman (2007) have done in their study to simply compute the bootstrapped standard errors with our panel data set.

The statistical significance of the bootstrapped t-statistics of the beta coefficient estimates is determined by using the bootstrapped critical values reported in Table 24.3. Compared with the bootstrapped critical values presented in Table 24.3, we notice that all bootstrapped t-statistics constructed from bootstrapped standard errors are statistically significant at the 1% level. The intercept is now significant in three cases (columns V to VII) when the standard errors computed by White, and by single-clustering. It is noteworthy that in Table 24.3 the bootstrapped critical values on beta coefficient estimates when double clustering is used are numerically larger than the corresponding asymptotic t-test critical values of 2.58 (1%), 1.96 (5%) and 1.65 (10%), indicating the use of the large-sample (normal approximation) critical values can lead to misleading test results when both firm and time effects exist in the residuals. On the other hand, intercept coefficient in column VIII was not significant when bootstrapped double clustering is used. Interestingly, we observe from Table 24.3 that bootstrapped critical values differ considerably depending on different standard error estimation methods.

We now examine the finite-sample size of the bootstrapped t-statistics assuming the beta coefficient takes the OLS estimate under the null hypothesis. Table 24.4 shows that for the bootstrapped t-statistics, no serious size distortion is found as reflected by the close to 5% size values. However, for the tests based on the standard asymptotic t-test critical values, all tests suffer from serious size distortion. For example, those based on the OLS White and clustered by time are undersized, while others oversized.

24.4. Conclusion

In this paper, we examine the performance of single- and double-clustered standard errors using the wild cluster bootstrap method. The panel data set on the

Chinese mutual funds used in the analysis has similar number of firms (54) and time periods (48), we are particularly interested in the performance of the bootstrapped double-clustered standard errors. This is mainly due to the conclusion made in Thompson (2011) that double-clustering the standard errors matters the most when the number of firms and time periods are not too different.

In the presence of firm and time effects, the standard OLS White standard errors are found to be underestimated when compared to standard errors computed from double-clustering (columns I to IV, Table 24.2). Further, the wild-cluster bootstrapped standard errors are found to account for the firm and time effects in residuals, as evidenced in column VIII of Table 24.2. The bootstrapped t-statistic computed by OLS White method is found to be much smaller than that calculated from the double-clustering, suggesting that the firm and time effects in residuals are strong.

The size values for the test statistics of the beta coefficient estimates in Table 24.4 suggest that the bootstrapped double-clustering outperforms the single-clustering either by firm or by time. They support the Thompson’s (2011) argument that double-clustering the standard errors of parameter estimators matters the most when the number of firms and time periods are not too different. Size distortions reported in Table 24.5 imply that it may not be appropriate to compare the bootstrapped t-statistics with standard t-test critical values. These findings also suggest that to avoid obtaining misleading test results with the presence of either firm or time effects or both, the bootstrapped critical values are preferred to conventional critical values. Additionally, a comparison of the sizes displayed in Table 24.4 with those calculated using the pairs cluster bootstrap method shown in Table 24.6 also suggests the wild cluster bootstrap approach performs better.

References

Brownstone, D. and R. Valletta. 2001. "The bootstrap and multiple imputations: Harnessing increased computing power for improved statistical tests." *Journal of Economic Perspectives* 15, 129-141.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics* 90, 414-427.

Chen, Z, and R. Lin. 2006. "Mutual fund performance evaluation using data envelopment analysis with new risk measures." *OR Spectrum* 28, 375-398.

Davidson, R. and E. Flachaire. 2008. "The wild bootstrap, tamed at last." *Journal of Econometrics* 146, 162-169.

Kayhan, A., and S. Titman. 2007. "Firms' histories and their capital structures." *Journal of Financial Economics* 83, 1-32.

Liu, R.Y. 1988. "Bootstrap procedures under some non-IID models." *Annals of Statistics* 16, 1696-1708.

Mammen, E. 1993. "Bootstrap and wild bootstrap for high dimensional linear models." *Annals of Statistics* 21, 255-285.

Petersen, M.A. 2009. "Estimating standard errors in finance panel data sets: comparing approaches." *The Review of Financial Studies* 22, 435-480.

Shen, W.T. and S.L. Huang. 2001. "An empirical analysis of China's mutual funds: performance and evaluation." *Economic Research* 9, 22-30 (in Chinese).

Stephens, M.A. 1974. "EDF statistics for goodness of fit and some Comparisons. *Journal of the American Statistical Association* 69, 730-737.

Thompson, S. B. 2011. "Simple formulas for standard errors that cluster by both firm and time." *Journal of Financial Economics* 99, 1-10.

Wu, C.F.G. 1986. "Jackknife, bootstrap and other resampling methods in regression analysis." *Annals of Statistics* 14, 1261-1295.

Table 24.1 Summary statistics for variables used in China's mutual fund regressions (Sept 2002-Aug 2006)

Variable	Mean	Std. deviation	Skewness	Kurtosis	Normality test stat. W-sq.	Normality test stat. A-sq.
Mutual funds returns	0.0006	0.0580	0.1313	0.1489	0.1809***	0.2062***
Market excess returns	0.0025	0.0465	0.1554	-0.7801	2.4993***	18.3650***

Notes: Sample size N=2592. Individual mutual funds return is the dependent variable, while the market excess return is the independent variable. W-sq or W^2=Cramer-von Mises test statistic, and A-sq or A^2 =Anderson-Darling test statistic. Both test statistics are empirical distribution function (EDF) statistics. The computing formulas for W^2 and A^2 statistics are available in Stephens (1974), p.731.
*** denotes statistical significance at the 1% level.

Table 24.2 Application to asset pricing modeling

		t-statistics			
		White	Clustered by Firm	Clustered by Time	Clustered by Firm&Time
Regressor	Estimate	I	II	III	IV
$R_{m,t}$	0.8399	46.935***	65.492***	9.017***	9.099***
1% critical value (CV)		2.576	2.576	2.576	2.576
Intercept	-0.0016	-1.858	-2.222**	-0.326	-0.327
1% critical value (CV)		-2.576	-2.576	-2.576	-2.576
Coefficient estimates	OLS				
R-squared	0.4539				

		t-statistics			
		White	Clustered by Firm	Clustered by Time	Clustered by Firm&Time
Regressor	Estimate	V	VI	VII	VIII
$R_{m,t}$	0.8401	47.215***	66.808***	50.886***	73.672***
Bootstrapped 1% CV		1.951	2.661	2.520	11.898
Intercept	-0.0016	-1.878**	-2.262**	-2.033**	-2.659
Bootstrapped 1% CV		-2.300	-2.717	-2.371	-12.288
Coefficient estimates	Wild cluster bootstrap				
R-squared	0.4579				

Notes: The dependent variable is the monthly mutual fund return in excess of risk-free rate, denoted as $R_{i,t}$, and the independent variable $R_{m,t}$ is the market returns in excess of risk-free rate. Both variables are monthly observations from September 2002 to August 2006.
*** , and ** denote statistical significance at the 1%, and 5% levels, respectively.

Table 24.3 Bootstrapped critical values

Percentiles	0.005	0.025	0.05	0.95	0.975	0.995
			OLS White			
$R_{m,t}$	-1.842	-1.396	-1.174	1.213	1.490	1.951
Intercept	-2.300	-1.656	-1.396	1.385	1.587	2.091
			Clustered by Firm			
$R_{m,t}$	-2.621	-1.984	-1.698	1.766	2.007	2.661
Intercept	-2.717	-2.035	-1.700	1.778	2.014	2.672
			Clustered by Time			
$R_{m,t}$	-2.513	-1.795	-1.510	1.517	1.821	2.520
Intercept	-2.371	-1.956	-1.608	1.629	1.917	2.364
			Clustered by Firm & Time			
$R_{m,t}$	-6.517	-4.376	-3.491	3.081	4.416	11.898
Intercept	-12.288	-4.053	-2.792	3.136	4.384	8.168

Note: Critical values are obtained by implementing the bootstrap procedure presented in the Appendix.

Table 24.4 Rejection rates using wild cluster bootstrapped critical values

	OLS White	Clustered by Firm	Clustered by Time	Clustered by Firm & Time
$R_{m,t}$	0.050	0.067	0.053	0.052
Intercept	0.042	0.034	0.037	0.036

Note: The number of replications is 1000.

Table 24.5 Rejection rates using conventional critical values

	OLS White	Clustered by Firm	Clustered by Time	Clustered by Firm & Time
$R_{m,t}$	0.004	0.069	0.037	0.206
Intercept	0.012	0.046	0.036	0.203

Note: The number of replications is 1000.

Table 24.6 Rejection rates using pairs cluster bootstrapped critical values

	OLS White	Clustered by Firm	Clustered by Time	Clustered by Firm & Time
$R_{m,t}$	0.043	0.040	0.040	0.040
Intercept	0.062	0.048	0.059	0.060

Note: The number of replications is 1000.

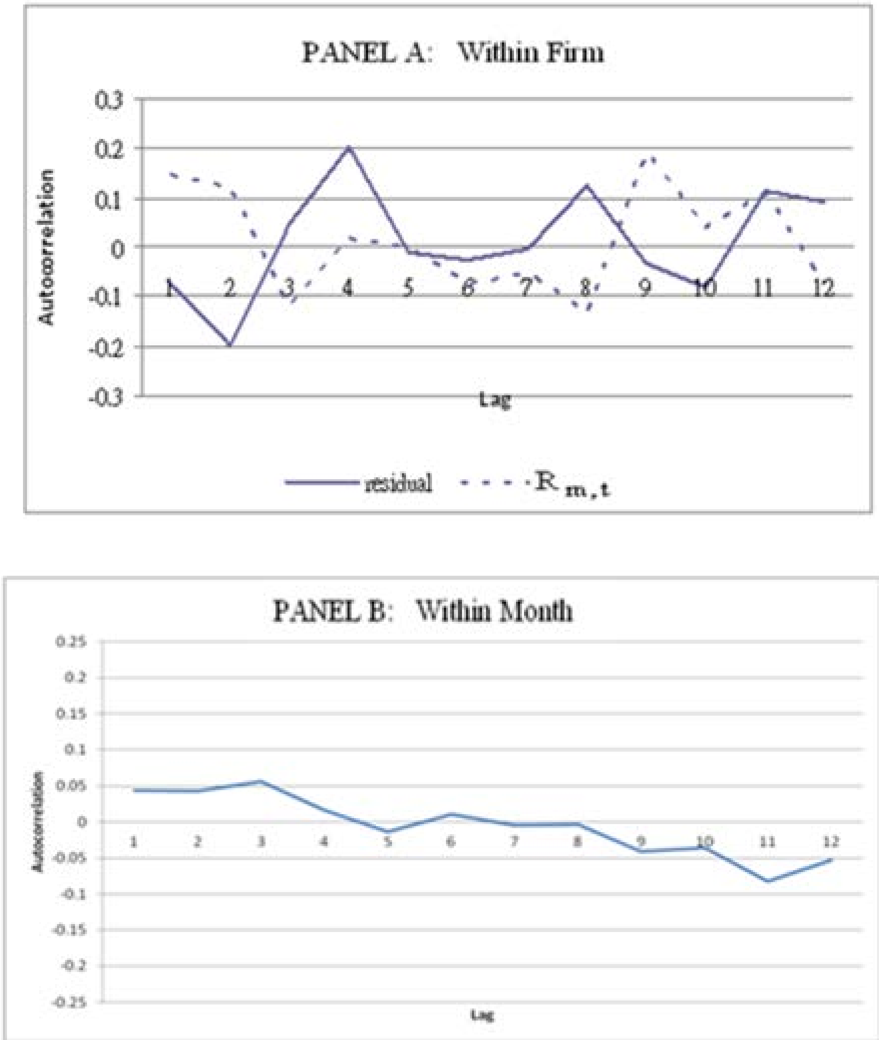

Fig. 24.1 The autocorrelations of residuals and the independent variable are plotted for one to twelve lags. The solid lines in Panel A and B show, respectively, within-firm and within-month autocorrelations in residuals, whereas the dashed line in Panel A shows the within-firm autocorrelations in the independent variable.

Appendix 24A Wild Cluster Bootstrap Procedure[5]

The following steps are used to obtain the Wild-clustered bootstrapped standard errors and critical values.

(i) *Define firm effects, and time effects*. The asset pricing model using individual variables $R_{i,t}$ and $R_{m,t}$ is specified as

$$R_{it} = \beta_0 + R_{m,t}\beta_1 + e_{it}, \quad i = 1,...,N, \quad t = 1,...,T. \tag{24A.1}$$

where the variables $R_{i,t}$ and $R_{m,t}$ are, respectively, the return of mutual fund i and the market return in excess of risk-free rate in month t. β_0 and β_1 are unknown parameters. The construction of the variables are detailed in Section 24.3.1. e_{it} is the error term. It may be heteroskedastic, but is assumed to be independent of the explanatory variable $E(e_{it} \mid R_{m,t}) = 0$.

Following Thompson (2011), we make the following assumptions on the correlations between errors, e_{it}:

(a) *Firm effects*: The errors may be correlated across time for a particular firm, that is, $E(e_{it}, e_{ik} \mid R_{m,t}, R_{m,k}) \neq 0$ for all $t \neq k$.

(b) *Time effects*: The errors may be correlated across firms within the same time period, that is,

$$E(e_{it}, e_{jt} \mid R_{m,t}) \neq 0 \quad \text{for all } i \neq j.$$

Let G be the number of clusters, and let N_g be the number of observations within each cluster. The errors are assumed to be independent across clusters but correlated within clusters. The asset pricing model can be written as

$$R_{ig} = \beta_0 + R_m\beta_1 + e_{ig}, \quad i = 1,,,N_g, \quad g = 1,...,G,$$

$$\mathbf{R}_g = \delta\,\beta_0 + \mathbf{R}_{mg}\beta_1 + \mathbf{e}_g, \quad g = 1,...,G, \tag{24A.2}$$

where R_{ig}, R_m and e_{ig} are scalars, $\mathbf{R}_g$, $\mathbf{R}_{mg}$, and $\mathbf{e}_g$ are $N_g \times 1$ vectors, and δ is $N_g \times 1$ vector with all elements equal to 1.

(ii) *Fit data to model.* We fit model (24A.2) to the observed data using OLS, and

[5] See also Cameron, et al. (2008) for details.

obtain the parameter estimates $\hat{\beta}_0$ *and* $\hat{\beta}_1$ together with the OLS residuals $\hat{\mathbf{e}}_{\mathbf{g}}$, g = 1,…, G..

(iii) *Construct 1000 bootstrap-samples*

The bootstrap-residuals are obtained according to the following transformation relation: $e_g^* = a_g \hat{e}_g$, where a_g takes on one of the following values: (i) $(1-\sqrt{5})/2 \approx$ -0.6180 with probability $(1+\sqrt{5})/(2\sqrt{5}) \approx 0.7236$; or (ii) $(1+\sqrt{5})/2 \approx 1.6180$ with probability $1-(1+\sqrt{5})/(2\sqrt{5}) \approx 0.2764$.

Hence,

(a) For each cluster g = 1,…, G, set $\mathbf{e}_{\mathbf{g}}^* = 1.618 * \hat{\mathbf{e}}_{\mathbf{g}}$ with probability 0.2764 or $\mathbf{e}_{\mathbf{g}}^* = -0.618 * \hat{\mathbf{e}}_{\mathbf{g}}$ with probability 0.7236.

(b) Repeat (a) 1000 times to obtain $\mathbf{e}_{\mathbf{g}}^*$ and then construct the bootstrap samples $\mathbf{R}_g^*$ as follows:

$$\mathbf{R}_g^* = \delta\, \hat{\beta}_0 + \mathbf{R}_{mg} \hat{\beta}_1 + \boldsymbol{e}_g^* . \tag{24A.3}$$

(iv) With each pseudo sample generated in step (iii), we estimate the parameters $\hat{\beta}_0^*$ *and* $\hat{\beta}_1^*$ by OLS, White standard errors ($\hat{SE}_{white}$) which are OLS standard errors robust to heteroskedasticity, as well as standard errors clustered by firm ($\hat{SE}_{firm}$), by time ($\hat{SE}_{time}$), and by both firm and time ($\hat{SE}_{both}$). The simulations are performed using GAUSS 9.0. The standard error formulas can be found in Thompson (2011).

(v) Construct bootstrapped test statistics by taking ratios of $\hat{\beta}_i^*$ $(i = 0,1)$ obtained by OLS to its corresponding $\hat{SE}_{white}$, $\hat{SE}_{firm}$, $\hat{SE}_{time}$, and $\hat{SE}_{both}$ obtained in step (iv). More specifically, the bootstrapped test statistics are expressed as follows:

$$w_i^* = \frac{\hat{\beta}_i^* - \hat{\beta}_i}{\hat{SE}(\hat{\beta}_i^*)}, \quad i = 0,1, \tag{24A.4}$$

where $\hat{SE}(\hat{\beta}_i^*)$ can either be $\hat{SE}_{white}$, $\hat{SE}_{firm}$, $\hat{SE}_{time}$, or $\hat{SE}_{both}$.

(vi) Obtain the empirical distribution of the individual test statistics by sorting the

1000 test statistics computed in step (v) in an ascending order. Bootstrapped critical values are then obtained from this empirical distribution at the following quantiles: 0.5%, 2.5%, 5%, 95%, 97.5%, and 99.5%, respectively.