Mitchell A. Petersen, 2009, "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, vol. 22, no. 1, pp. 436-480

- Many works of Finance often use panel data.
- $^{\exists}$cross- or time- correlation $\Rightarrow$ $^{\exists}$bias in OLS standard errors (hereafter *SE*)
- Historically, researchers used different solutions for above two problems.
- objectives of this paper
  - examines the different methods used in the literature
  - explains when the different methods yield the same (and correct) *SE* and when they diverge
  - provide intuition as to why the different approaches sometimes give different answers and give researchers guidance for their use

- for $\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}$, *iid* $\boldsymbol{\varepsilon} \Rightarrow$ $^{\nexists}$bias in OLS *SE*
- Researchers have addressed possible biases in the *SE* varies widely and in many cases is incorrect.
  - 42% of the papers have used wrong standard errors. (207 published papers of JF, JFE, RFS in 2001-2004)
- Approaches for estimating $\boldsymbol{\beta}$s and *SE*s in the presence of the within-cluster correlation varied among the remaining papers.
  - 34% used Fama-MacBeth procedure for $\boldsymbol{\beta}$, *SE*
  - 29% used within-cluster dummy variables
  - 7% used Newey-West standard errors for *SE*
  - 23% used Rogers (analogous with White) standard errors for *SE*
  - etc.
- how the methods compare & how to select the correct one is important.

- 2 general fomrs of dependence most common in Finance applications
  - time-series dependence=firm effect (ϱ for same firm, different time)
  - cross-sectional dependence=time effect (ϱ for same time, different firm)
- 3 panel data simulation ① only firm effect ② only time effect ③ both effects
- ① $\boldsymbol{\beta}$ & *SE* estimation by each of them ② relative performance comparision

- section 1 $\rightarrow$ $^{\exists}$firm effect
  - $\Rightarrow$ $^{\exists}$downward big bias in OLS *SE*, Fama-MacBeth *SE*
  - $\Rightarrow$ $^{\exists}$downward small bias in Newey-West *SE*
  - $\Rightarrow$ $^{\nexists}$bias in clustered *SE*
  - ex. financial leverage, dividends, investment
- section 2 $\rightarrow$ $^{\exists}$time effect
  - $\Rightarrow$ $^{\nexists}$bias in Fama-MacBeth *SE*
  - ex. equity returns, earnings surprises

- section 3 $\rightarrow$ $^{\exists}$firm effect+time effect
  - estimating standard errors clustered on more than 1 dimension
- section 4 $\rightarrow$ $^{\exists}$general correlation structure
  - comparison of the OLS, clustered, and Fama-MacBeth *SE* in the setting
  - examine accuracy of 3 additional methods for adjusting *SE*
    - ✓ $^{\exists}$fixed firm effect $\Rightarrow$ $^{\nexists}$bias in fixed-effects *SE*, random-effects *SE*
    - ✓ most cases $\Rightarrow$ $^{\exists}$bias in adjusted Fama-MacBeth *SE*
- section 5 $\rightarrow$ example; real data application
  - some published papers $\rightarrow$ may $^{\exists}$biases$^{\downarrow}$ in *SE* and biases$^{\uparrow}$ in t stat.

1. $^{\exists}$firm effect

- standard regression model for panel data

$$\begin{aligned} Y_{it} &= X_{it}\beta+\varepsilon_{it} && i\in\{1,2,\cdots,N\},\ t\in\{1,2,\cdots,T\} \\ X &\sim^{iid} ?(0,{\sigma_X}^2) && {\sigma_X}^2<\infty \\ \varepsilon &\sim^{iid} ?(0,{\sigma_\varepsilon}^2) && {\sigma_\varepsilon}^2<\infty \end{aligned}$$

- then OLS beta estimator

$$\begin{aligned} \hat{\beta}_{OLS} &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}Y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} \\ &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}(X_{it}\beta+\varepsilon_{it})}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} \\ &= \beta+\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2} \end{aligned}$$

- and its asymptotic variance

$$\begin{aligned} AVar\left(\hat{\beta}_{OLS}-\beta\right) &= \operatorname*{plim}_{\substack{N\rightarrow\infty \\ T\ fixed}}\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}\varepsilon_{it}\right)^2\left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right] \\ &= \operatorname*{plim}_{\substack{N\rightarrow\infty \\ T\ fixed}}\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2\right)\left(\frac{\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}^2}{N}\right)^{-2}\right] && \text{if independent} \\ &= \frac{1}{N}\left(T\sigma_X^2\sigma_\varepsilon^2\right)\left(T\sigma_X^2\right)^{-2} && \text{if identical} \\ &= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT} \end{aligned}$$

- $X,\varepsilon=iid$ $\Rightarrow$ OLS=correct
- however, $^{\exists}$fixed firm effect

$$
\begin{aligned}
X_{it} &= \mu_i + v_{it} \\
\varepsilon_{it} &= \gamma_i + \eta_{it} \\
\mu, v, \gamma, \eta &\sim^{iid} ?[0, \sigma^2(\cdot)] \\
Corr(X_{it}, X_{js}) &= 1 && \text{for } i=j \ \& \ t=s \\
&= \varrho_X (= \sigma_\mu{}^2 / \sigma_X{}^2) && \text{for } i=j \ \& \ t \neq s \\
&= 0 && \text{for } i \neq j \\
Corr(\varepsilon_{it}, \varepsilon_{js}) &= 1 && \text{for } i=j \ \& \ t=s \\
&= \varrho_\varepsilon (= \sigma_\gamma{}^2 / \sigma_\varepsilon{}^2) && \text{for } i=j \ \& \ t \neq s \\
&= 0 && \text{for } i \neq j
\end{aligned}
$$

- then its asymptotic variance

$$
\begin{aligned}
A\,Var\left(\hat{\beta}_{OLS} - \beta\right) &= \underset{\substack{N \to \infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it} \varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \underset{\substack{N \to \infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it} \varepsilon_{it} \right)^2 \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \underset{\substack{N \to \infty \\ T\ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}^2 \varepsilon_{it}^2 + 2 \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} X_{it} X_{is} \varepsilon_{it} \varepsilon_{is} \right) \left(\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2}{N} \right)^{-2} \right] \\
&= \frac{1}{N} \left[T \sigma_X^2 \sigma_\varepsilon^2 + T(T-1) \rho_X \sigma_X^2 \rho_\varepsilon \sigma_\varepsilon^2 \right] \left(T \sigma_X^2 \right)^{-2} \qquad \text{if identical} \\
&= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT} \left[1 + (T-1) \rho_X \rho_\varepsilon \right]
\end{aligned}
$$

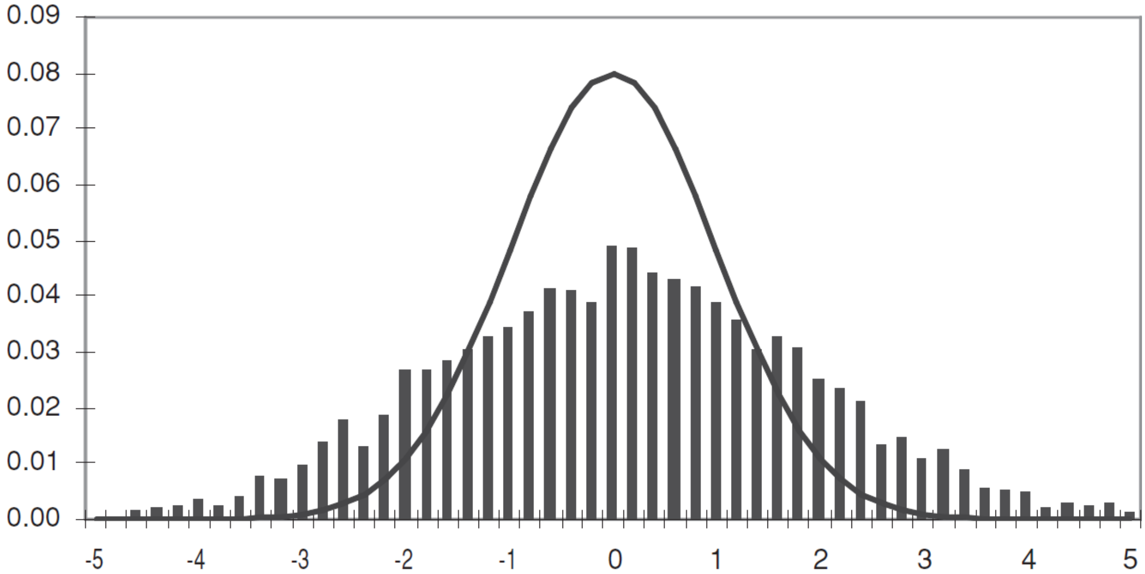
- $\therefore$ $^{\exists}$firm effect $\Leftrightarrow$ $\varrho_X, \varrho_\varepsilon \neq 0$ $\Rightarrow$ $^{\exists}$bias$^{\downarrow}$ in OLS SE
  - extreme case $\rightarrow$ $\varrho_X = \varrho_\varepsilon = 1$ $\Rightarrow$ $N^{-1} \sigma_X{}^{-2} \sigma_\varepsilon^2$ i.e. additional obs. in time=no information in reality, but OLS SE will regard these obs. as information then bias$^{\downarrow}$ in SE
- clustered SE $\rightarrow$ designed to correct the correlation of the residuals within a cluster

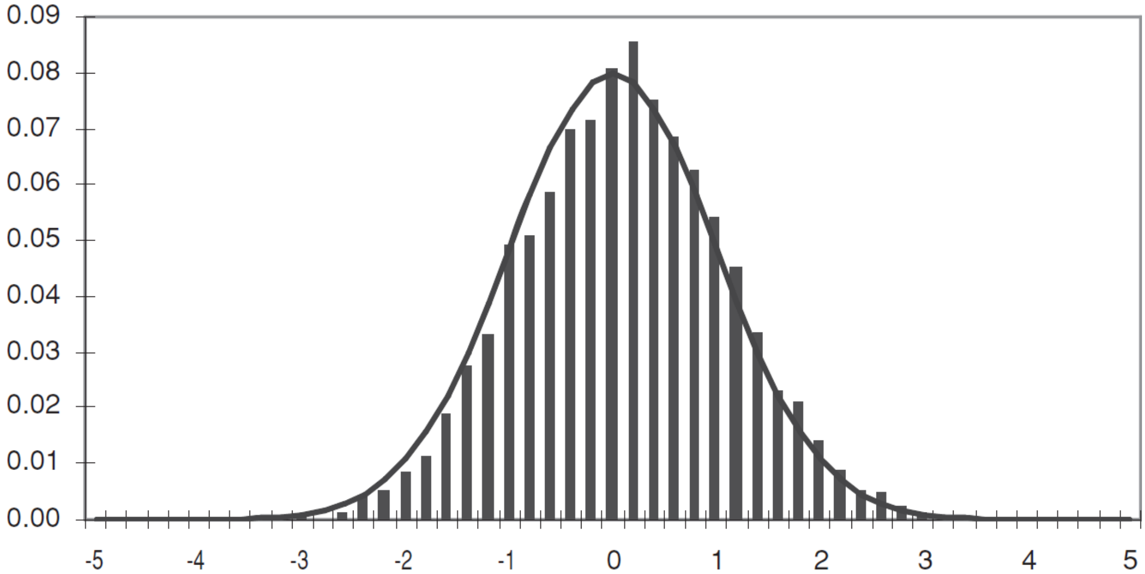
$$
A\,Var(\beta) = \frac{N(NT-1) \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it} \varepsilon_{it} \right)^2}{(NT-k)(N-1) \left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}^2 \right)^2}
$$

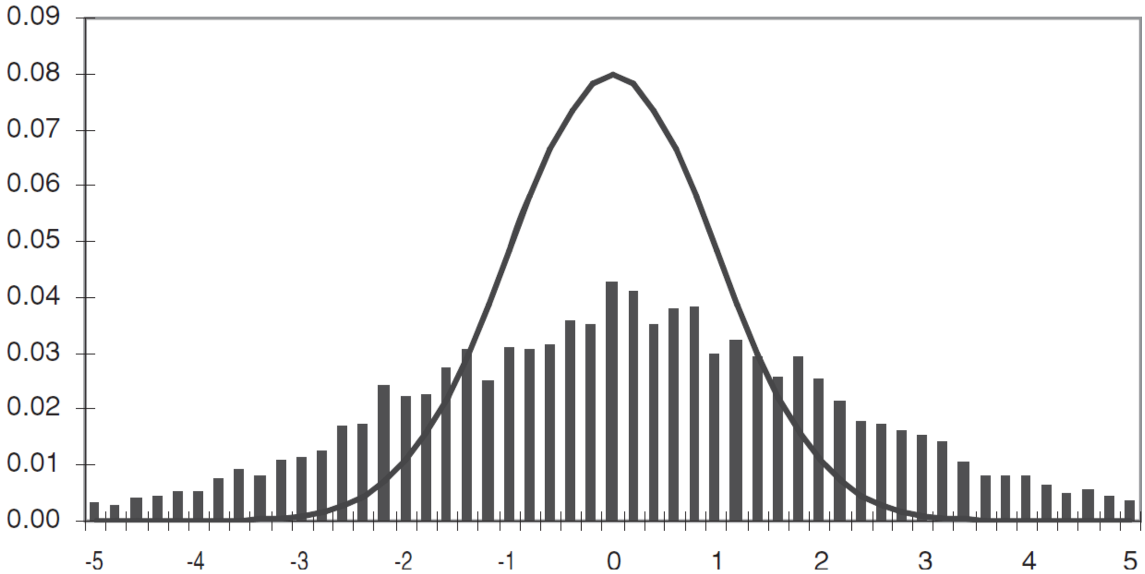
	firm 1			firm 2			firm 3		
firm 1	$\varepsilon_{11}{}^2$	$\varepsilon_{11}\varepsilon_{12}$	$\varepsilon_{11}\varepsilon_{13}$	0	0	0	0	0	0
	$\varepsilon_{12}\varepsilon_{11}$	$\varepsilon_{12}{}^2$	$\varepsilon_{12}\varepsilon_{13}$	0	0	0	0	0	0
	$\varepsilon_{11}\varepsilon_{13}$	$\varepsilon_{13}\varepsilon_{12}$	$\varepsilon_{13}{}^2$	0	0	0	0	0	0
firm 2	0	0	0	$\varepsilon_{21}{}^2$	$\varepsilon_{21}\varepsilon_{22}$	$\varepsilon_{21}\varepsilon_{23}$	0	0	0
	0	0	0	$\varepsilon_{22}\varepsilon_{21}$	$\varepsilon_{22}{}^2$	$\varepsilon_{22}\varepsilon_{23}$	0	0	0
	0	0	0	$\varepsilon_{21}\varepsilon_{23}$	$\varepsilon_{23}\varepsilon_{22}$	$\varepsilon_{23}{}^2$	0	0	0
firm 3	0	0	0	0	0	0	$\varepsilon_{31}{}^2$	$\varepsilon_{31}\varepsilon_{32}$	$\varepsilon_{31}\varepsilon_{33}$
	0	0	0	0	0	0	$\varepsilon_{32}\varepsilon_{31}$	$\varepsilon_{32}{}^2$	$\varepsilon_{32}\varepsilon_{33}$
	0	0	0	0	0	0	$\varepsilon_{31}\varepsilon_{33}$	$\varepsilon_{33}\varepsilon_{32}$	$\varepsilon_{33}{}^2$

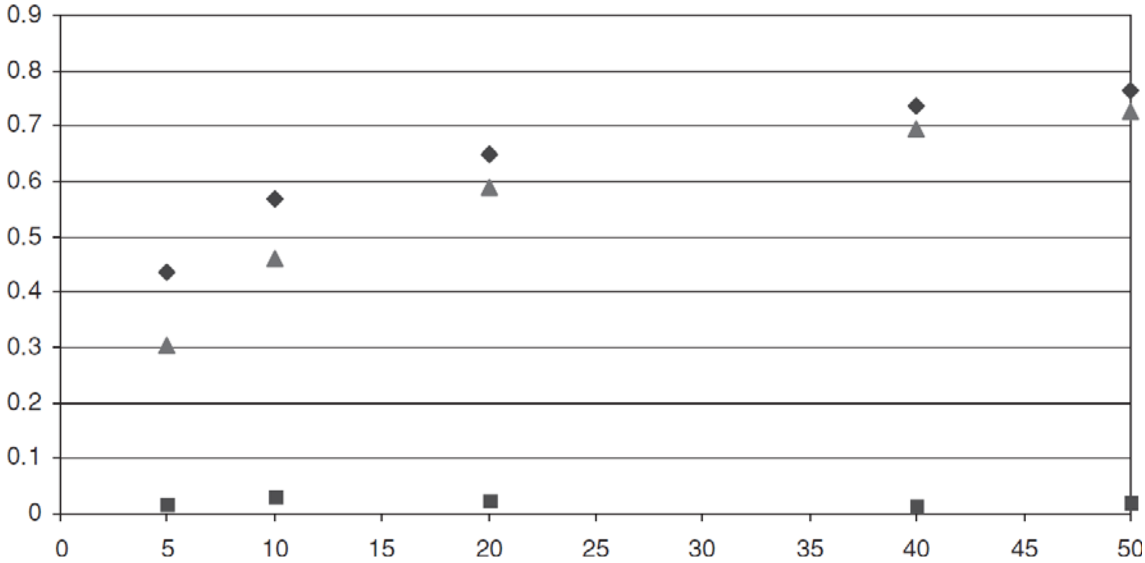
Table 1
Estimating standard errors with a firm effect OLS and clustered standard errors

		Source of independent variable volatility			
$Avg(\beta_{OLS})$ $Std(\beta_{OLS})$ $Avg(SE_{OLS})$ % $Sig(T_{OLS})$ $Avg(SE_C)$ % $Sig(T_C)$		0%	25%	50%	75%
Source of residual volatility	0%	1.0004	1.0006	1.0002	1.0001
		0.0286	0.0288	0.0279	0.0283
		0.0283	0.0283	0.0283	0.0283
		[0.0098]	[0.0088]	[0.0094]	[0.0094]
		0.0283	0.0282	0.0282	0.0282
		[0.0108]	[0.0092]	[0.0096]	[0.0098]
	25%	1.0004	0.9997	0.9999	0.9997
		0.0287	0.0353	0.0403	0.0468
		0.0283	0.0283	0.0283	0.0283
		[0.0116]	[0.0348]	[0.0678]	[0.1174]
		0.0283	0.0353	0.0411	0.0463
		[0.0120]	[0.0064]	[0.0112]	[0.0118]
	50%	1.0001	1.0002	1.0007	0.9993
		0.0289	0.0414	0.0508	0.0577
		0.0283	0.0283	0.0283	0.0283
		[0.0124]	[0.0770]	[0.1534]	[0.2076]
		0.0282	0.0411	0.0508	0.0590
		[0.0128]	[0.0114]	[0.0088]	[0.0102]
	75%	1.0000	1.0004	0.9995	1.0016
		0.0285	0.0459	0.0594	0.0698
		0.0283	0.0283	0.0283	0.0283
		[0.0128]	[0.1090]	[0.2230]	[0.2906]
		0.0282	0.0462	0.0589	0.0693
		[0.0128]	[0.0114]	[0.0094]	[0.0112]


0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
-5
-4
-3
-2
-1
0
1
2
3
4
5


0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
-5
-4
-3
-2
-1
0
1
2
3
4
5


0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
-5
-4
-3
-2
-1
0
1
2
3
4
5



- Fama-MacBeth beta estimator

$$\begin{aligned}
\hat{\beta}_{FM} &= \sum_{t=1}^{T} \frac{\hat{\beta}_{FM}}{T} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N} X_{it} Y_{it}}{\sum_{i=1}^{N} X_{it}^2}\right) \\
&= \beta + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{\sum_{i=1}^{N} X_{it}^2}\right)
\end{aligned}$$

- and Fama-MacBeth SE estimator

$$S(\hat{\beta}_{FM}) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\frac{(\hat{\beta}_t - \hat{\beta}_{FM})^2}{T-1}}$$

- in reality, asymptotic variance of FM SE

$$\begin{aligned}
A\,Var(\hat{\beta}_{FM}) &= \frac{1}{T^2} A\,Var\left(\sum_{t=1}^{T}\hat{\beta}_t\right) \\
&= \frac{A\,Var(\hat{\beta}_t)}{T} + \frac{1}{T^2} 2 \sum_{t=1}^{T-1}\sum_{s=t+1}^{T} A\,Cov(\hat{\beta}_t, \hat{\beta}_s) \\
&= \frac{A\,Var(\hat{\beta}_t)}{T} + \frac{T(T-1)}{T^2} A\,Cov(\hat{\beta}_t, \hat{\beta}_s) \\
\Rightarrow A\,Cov(\hat{\beta}_t, \hat{\beta}_s) &= \underset{\substack{N\to\infty \\ T\ fixed}}{\text{plim}} \left[\left(\frac{\sum_{i=1}^{N} X_{it}^2}{N}\right)^{-1}\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{N}\right)\left(\frac{\sum_{i=1}^{N} X_{is}\varepsilon_{is}}{N}\right)\left(\frac{\sum_{i=1}^{N} X_{is}^2}{N}\right)^{-1}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \left(\sigma_X^2\right)^{-2} \operatorname*{plim}_{\substack{N\to\infty \\ T\ fixed}} \left[\left(\frac{\sum_{i=1}^{N} X_{it}\varepsilon_{it}}{N}\right)\left(\frac{\sum_{i=1}^{N} X_{is}\varepsilon_{is}}{N}\right)\right] \\
&= \left(\sigma_X^2\right)^{-2} \operatorname*{plim}_{\substack{N\to\infty \\ T\ fixed}} \left[\frac{\sum_{i=1}^{N} X_{it}X_{is}\varepsilon_{it}\varepsilon_{is}}{N^2}\right] \\
&= \left(\sigma_X^2\right)^{-2} \frac{N\rho_X\sigma_X^2\rho_\varepsilon\sigma_\varepsilon^2}{N^2} \\
&= \frac{\rho_X\rho_\varepsilon\sigma_\varepsilon}{N\sigma_X^2}
\end{aligned}
$$

$$
\begin{aligned}
\therefore\ \ A\,Var\left(\hat{\beta}_{FM}\right) &= \frac{A\,Var\left(\hat{\beta}_t\right)}{T} + \frac{T(T-1)}{T^2} A\,Cov\left(\hat{\beta}_t, \hat{\beta}_s\right) \\
&= \frac{1}{T}\left(\frac{\sigma_\varepsilon^2}{N\sigma_X^2}\right) + \frac{T(T-1)}{T^2}\left(\frac{\rho_X\rho_\varepsilon\sigma_\varepsilon^2}{N\sigma_X^2}\right) \\
&= \frac{\sigma_\varepsilon^2}{\sigma_X^2 NT}\left[1 + (T-1)\rho_X\rho_\varepsilon\right]
\end{aligned}
$$

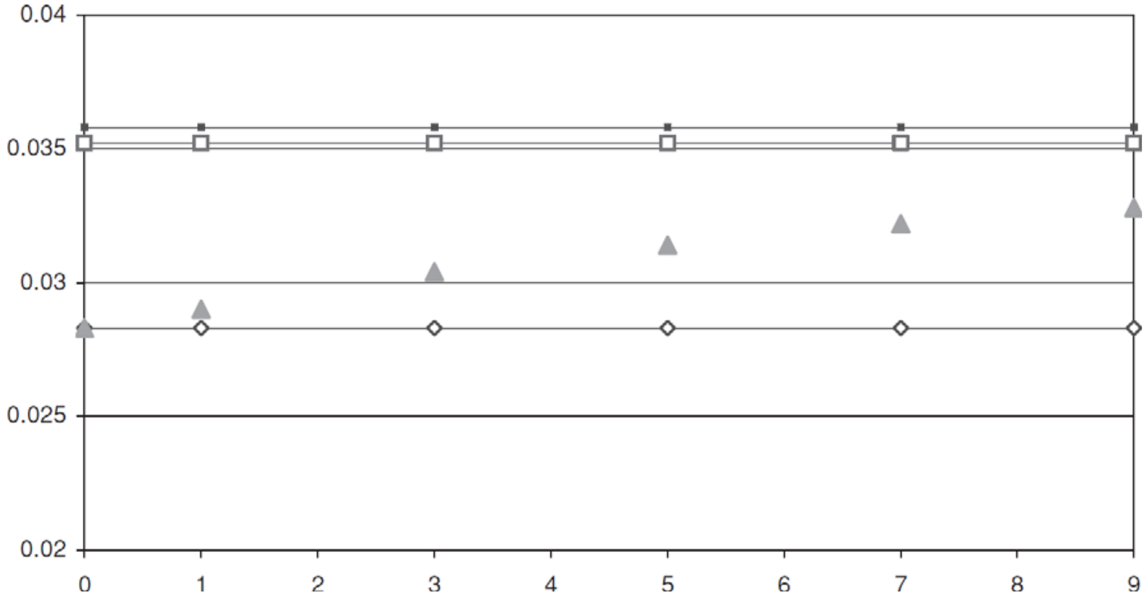
- thus, Fama-MacBeth *SE* are biased in the same way as the OLS estimates.

Avg(β_{FM}) Std(β_{FM}) Avg(SE_{FM}) % Sig(T_{FM})		0%	25%	50%	75%
Source of residual volatility	0%	1.0004 0.0287 0.0276 [0.0288]	1.0006 0.0288 0.0276 [0.0304]	1.0002 0.0280 0.0277 [0.0236]	1.0001 0.0283 0.0275 [0.0294]
	25%	1.0004 0.0288 0.0275 [0.0336]	0.9997 0.0354 0.0268 [0.0758]	0.9998 0.0403 0.0259 [0.1202]	0.9997 0.0469 0.0250 [0.1918]
	50%	1.0000 0.0289 0.0276 [0.0330]	1.0002 0.0415 0.0259 [0.1264]	1.0007 0.0509 0.0238 [0.2460]	0.9993 0.0578 0.0219 [0.3388]
	75%	1.0000 0.0286 0.0277 [0.0310]	1.0004 0.0460 0.0248 [0.1778]	0.9995 0.0595 0.0218 [0.3654]	1.0016 0.0699 0.0183 [0.4994]

- examples; persistent dependent and independent variables
  - $\text{dividend}_{it}=f(\text{firm characteristics}_{it})$? → Fama and French (2001, JFE)
  - $\text{M/B ratio}_{it}=f(\text{firm characteristics}_{it})$? → Pastor and Veronesi (2003, JF), Kemsley and Nissim (2002, JF)
  - capital structure literature; firm $\text{leverage}_{it}=f(\text{firm characteristics}_{it})$? → Baker and Wurgler (2002, JF), Fama and French (2002, RFS), Johnson (2003, RFS)
  - $\varrho_Y{}^{\uparrow}$ $\varrho_X{}^{\uparrow}$ $\Rightarrow$ $^{\exists}$bias$^{\downarrow}$ Fama-MacBeth *SE*
- Many authors believe (incorrectly) that these approaches are correct.
- The problem is actually worse; the literatures has gone on to provide incorrect advice that states that the Fama-MacBeth approach corrects the *SE* for the residual correlation in the presence of a firm effect. → Wu (2004, JFE), Denis, Denis, and Yost (2002, JF), Choe, Kho, and Stulz (2005, RFS)
- Fama-MacBeth *SE*
  - $^{\exists}$cross-correlation $\Rightarrow$ $^{\nexists}$bias in *SE* (time effect)
  - $^{\exists}$serial-correlation $\Rightarrow$ $^{\exists}$bias in *SE* (firm effect)
- Newey-West *SE* by Newey and West (1987)
  - account for a serial correlation of unknown form in the residuals of a single time-series
  - assumption → correlation decay $\propto$ distance between observation
  - covariance weight → $[1-j/(M+1)]$ M=maximum lag $\Rightarrow$ distance$^{\uparrow}$=weight$^{\downarrow}$
  - initially developed for a single time-series → weighting function should have made the weighting matrix positive semidefinite
  - modified for use in a panel data set; estimating only correlations between lagged residuals in the same cluster [Brockman and chung (2001), MacKay (2003), Bertrand, Duflo, and Mullainathan (2004), Doidge (2004)]

- Newey-West SE estimator

$$\begin{aligned}\sum_{i=1}^{N}\left(\sum_{t=1}^{T}X_{it}\varepsilon_{it}\right)^2 &= \sum_{i=1}^{N}\left[\sum_{t=1}^{T}X_{it}^2\varepsilon_{it}^2+2\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}w(t-s)X_{it}X_{is}\varepsilon_{it}\varepsilon_{is}\right]\\ &= \sum_{i=1}^{N}\left[\sum_{t=1}^{T}X_{it}^2\varepsilon_{it}^2+2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t}w(j)X_{it}X_{it-j}\varepsilon_{it}\varepsilon_{it-j}\right]\\ &= \sum_{i=1}^{N}\left[\sum_{t=1}^{T}X_{it}^2\varepsilon_{it}^2+2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t}\left(1-\frac{j}{T}\right)X_{it}X_{it-j}\varepsilon_{it}\varepsilon_{it-j}\right]\end{aligned}$$



- M i.e. maximum lag $(T-1)^{\uparrow} \Rightarrow$ Newey-West SE bias$^{\downarrow}$ ($\because$ restriction in weighting)

2. $^{\exists}$time effect

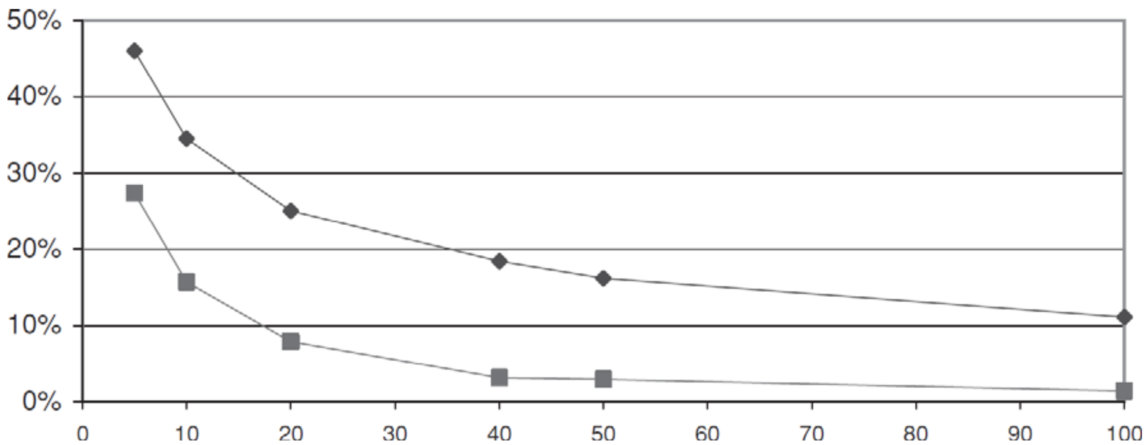
- conversion of correlation structures

$$\begin{aligned}X_{it} &= \zeta_t+v_{it}\\ \varepsilon_{it} &= \delta_t+\eta_{it}\end{aligned}$$

- $Var(\zeta)^{\uparrow}$ & $Var(\delta)^{\uparrow} \Rightarrow$ $^{\exists}$bias$^{\downarrow}$ in OLS SE
- clustered SE; accuracy$^{\uparrow}$ but $^{\exists}$bias$^{\downarrow}$ 13%～19%, $\because$ limited number of clusters $(T^{\downarrow})$
  - N=500=sufficient, T=10=too small [Kezdi (2004), Hansen (2007, *JEconometrics*)]
  - $T^{\uparrow} \Rightarrow$ bias$^{\downarrow}$

$$\begin{aligned}MSE &= E[(\widehat{SE}-SE_{true})^2]\\ &= E[(\widehat{SE}-\overline{SE}+\overline{SE}-SE_{true})^2]\\ &= E[(\widehat{SE}-\overline{SE})^2+(\overline{SE}-SE_{true})^2+2(\widehat{SE}-\overline{SE})(\overline{SE}-SE_{true})]\\ &= E[(\widehat{SE}-\overline{SE})^2+(\overline{SE}-SE_{true})^2]\\ &= Var(\widehat{SE})+[Bias(\widehat{SE})]^2\end{aligned}$$

Source of residual volatility	Statistic	Source of independent variable volatility 0%	25%	50%	75%
0%	Avg(β_{OLS})	1.0004	1.0002	1.0006	0.9994
	Std(β_{OLS})	0.0286	0.0291	0.0293	0.0314
	Avg(SE_{OLS})	0.0283	0.0288	0.0295	0.0306
	% Sig(T_{OLS})	[0.0098]	[0.0094]	[0.0088]	[0.0114]
	Avg(SE_C)	0.0277	0.0276	0.0275	0.0270
	% Sig(T_C)	[0.0330]	[0.0304]	[0.0348]	[0.0520]
25%	Avg(β_{OLS})	1.0006	1.0043	0.9962	0.9996
	Std(β_{OLS})	0.0284	0.1490	0.2148	0.2874
	Avg(SE_{OLS})	0.0279	0.0284	0.0289	0.0300
	% Sig(T_{OLS})	[0.0114]	[0.6064]	[0.7270]	[0.7874]
	Avg(SE_C)	0.0268	0.1297	0.1812	0.2305
	% Sig(T_C)	[0.0320]	[0.0360]	[0.0506]	[0.0736]
50%	Avg(β_{OLS})	0.9996	0.9997	0.9919	1.0079
	Std(β_{OLS})	0.0276	0.2138	0.3015	0.3986
	Avg(SE_{OLS})	0.0274	0.0278	0.0282	0.0292
	% Sig(T_{OLS})	[0.0100]	[0.7298]	[0.8096]	[0.8536]
	Avg(SE_C)	0.0258	0.1812	0.2546	0.3248
	% Sig(T_C)	[0.0294]	[0.0458]	[0.0596]	[0.0756]
75%	Avg(β_{OLS})	1.0002	0.9963	0.9970	0.9908
	Std(β_{OLS})	0.0273	0.2620	0.3816	0.4927
	Avg(SE_{OLS})	0.0267	0.0271	0.0276	0.0284
	% Sig(T_{OLS})	[0.0110]	[0.7994]	[0.8586]	[0.8790]
	Avg(SE_C)	0.0244	0.2215	0.3141	0.3986
	% Sig(T_C)	[0.0322]	[0.0402]	[0.0588]	[0.0768]



- Fama-MacBeth; $^{\exists}$time effect $\Rightarrow$ $Corr(\hat{\beta}_t, \hat{\beta}_s)=0$ $\Rightarrow$ $^{\nexists}$bias FM SE

Avg(β_{FM}) Std(β_{FM}) Avg(SE_{FM}) % Sig(T_{FM})		Source of independent variable volatility: 0%	25%	50%	75%
Source of residual volatility	0%	1.0004 0.0287 0.0278 [0.0310]	1.0004 0.0331 0.0318 [0.0312]	1.0007 0.0396 0.0390 [0.0252]	0.9991 0.0573 0.0553 [0.0338]
	25%	1.0005 0.0252 0.0239 [0.0376]	1.0003 0.0284 0.0276 [0.0296]	1.0006 0.0343 0.0338 [0.0284]	0.9999 0.0496 0.0480 [0.0294]
	50%	1.0000 0.0200 0.0195 [0.0254]	1.0002 0.0231 0.0227 [0.0304]	1.0006 0.0280 0.0276 [0.0272]	1.0007 0.0394 0.0387 [0.0278]
	75%	1.0001 0.0142 0.0138 [0.0308]	0.9996 0.0161 0.0159 [0.0302]	1.0000 0.0200 0.0196 [0.0284]	0.9999 0.0285 0.0276 [0.0300]

3. $^{\exists}$firm & time effect

- $^{\exists}$firm effect $\Rightarrow$ $^{\nexists}$bias in clustered SE
- $^{\exists}$time effect $\Rightarrow$ FM SE is good for either $T^{\downarrow}$ and $T^{\uparrow}$
  - cost $\rightarrow$ no correlation across clusters
- parametrically estimate one dimension; usually $N > T$ $\Rightarrow$ dummy for t, cluster for i $\rightarrow$ Lamont and Polk (2001, JFE), Anderson and Reeb (2004, JF), Gross and Souleles (2004, RFS), Sapienza (2004, JFE), Faulkender and Petersen (2006, RFS)
  - $^{\exists}$fixed time effect $\Rightarrow$ dummy for t eliminate correlation between observations in the same time period $\Rightarrow$ only $^{\exists}$firm effect $\Rightarrow$ $^{\nexists}$bias in clustered SE
  - only for fixed time effect; otherwise, even $^{\exists}$bias in clustered SE
  - The problem can be solved by simultaneous clustering for both firm and time $\rightarrow$ Cameron, Gelbach, and Miller (2006), Thompson (2006) proposed below.

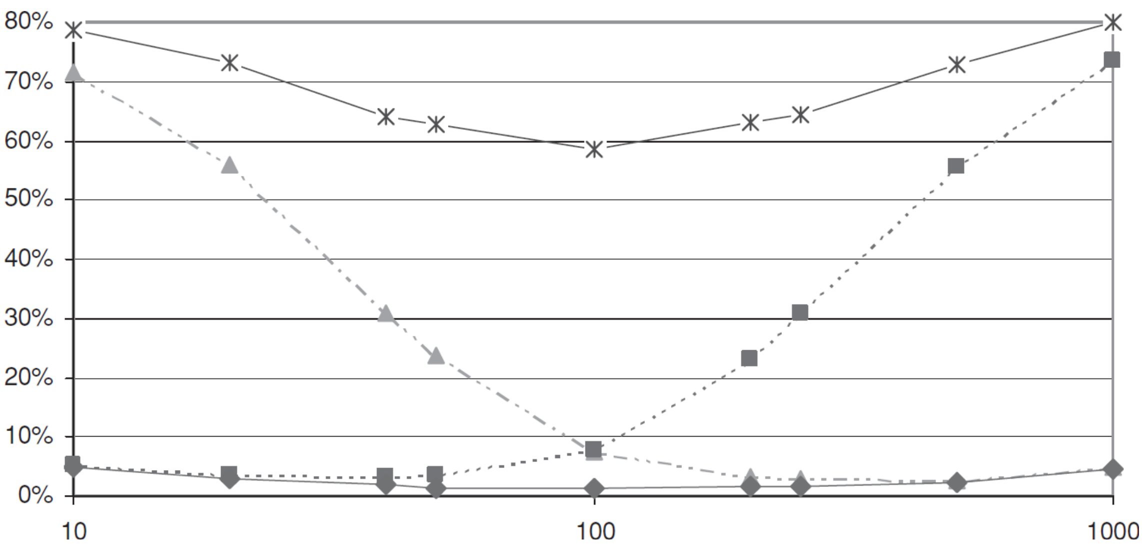
$$\begin{aligned}
\boldsymbol{V}_{\text{firm\&time}} &= \boldsymbol{V}_{\text{firm}} + \boldsymbol{V}_{\text{time}} - \boldsymbol{V}_{\text{White}} \\
\boldsymbol{V}_{\text{firm}} &:= \text{capture the correlation caused by firm effect} \\
\boldsymbol{V}_{\text{time}} &:= \text{capture the correlation caused by time effect} \\
\boldsymbol{V}_{\text{White}} &:= \text{avoid double-counting of diagonal term}
\end{aligned}$$

	firm 1			firm 2			firm 3		
firm 1	$\varepsilon_{11}{}^2$	$\varepsilon_{11}\varepsilon_{12}$	$\varepsilon_{11}\varepsilon_{13}$	$\varepsilon_{11}\varepsilon_{21}$	0	0	$\varepsilon_{11}\varepsilon_{31}$	0	0
	$\varepsilon_{12}\varepsilon_{11}$	$\varepsilon_{12}{}^2$	$\varepsilon_{12}\varepsilon_{13}$	0	$\varepsilon_{12}\varepsilon_{22}$	0	0	$\varepsilon_{12}\varepsilon_{32}$	0
	$\varepsilon_{11}\varepsilon_{13}$	$\varepsilon_{13}\varepsilon_{12}$	$\varepsilon_{13}{}^2$	0	0	$\varepsilon_{13}\varepsilon_{23}$	0	0	$\varepsilon_{12}\varepsilon_{32}$
firm 2	$\varepsilon_{21}\varepsilon_{11}$	0	0	$\varepsilon_{21}{}^2$	$\varepsilon_{21}\varepsilon_{22}$	$\varepsilon_{21}\varepsilon_{23}$	$\varepsilon_{21}\varepsilon_{31}$	0	0
	0	$\varepsilon_{22}\varepsilon_{12}$	0	$\varepsilon_{22}\varepsilon_{21}$	$\varepsilon_{22}{}^2$	$\varepsilon_{22}\varepsilon_{23}$	0	$\varepsilon_{22}\varepsilon_{32}$	0
	0	0	$\varepsilon_{22}\varepsilon_{12}$	$\varepsilon_{21}\varepsilon_{23}$	$\varepsilon_{23}\varepsilon_{22}$	$\varepsilon_{23}{}^2$	0	0	$\varepsilon_{23}\varepsilon_{33}$
firm 3	$\varepsilon_{31}\varepsilon_{11}$	0	0	$\varepsilon_{31}\varepsilon_{21}$	0	0	$\varepsilon_{31}{}^2$	$\varepsilon_{31}\varepsilon_{32}$	$\varepsilon_{31}\varepsilon_{33}$
	0	$\varepsilon_{32}\varepsilon_{12}$	0	0	$\varepsilon_{32}\varepsilon_{22}$	0	$\varepsilon_{32}\varepsilon_{31}$	$\varepsilon_{32}{}^2$	$\varepsilon_{32}\varepsilon_{33}$
	0	0	$\varepsilon_{33}\varepsilon_{13}$	0	0	$\varepsilon_{33}\varepsilon_{23}$	$\varepsilon_{31}\varepsilon_{33}$	$\varepsilon_{33}\varepsilon_{32}$	$\varepsilon_{33}{}^2$

- correlation structure

$$\begin{aligned} X_{it} &= \gamma_i+\delta_t+\eta_{it} \qquad Var(\gamma)=Var(\delta)=Var(\eta) \\ \varepsilon_{it} &= \mu_i+\zeta_t+v_{it} \qquad Var(\mu)=Var(\zeta)=Var(v) \end{aligned}$$

  - 1-D clustered $SE \Rightarrow {}^{\exists}\text{bias}^{\downarrow}$; $T^{\uparrow} \Rightarrow \text{bias}^{\uparrow} \rightarrow$ 5% for T=10, 73% for T=100



  - 2-D clustered $SE \Rightarrow {}^{\nexists}\text{bias}$ (enough T and N); but $T^{\downarrow}$ or $N^{\downarrow} \Rightarrow {}^{\exists}\text{bias}^{\downarrow}$ in SE $\rightarrow$ Cameron, Gelbach, and Miller (2006), Thompson (2006)
  - $T^{\downarrow}$ or $N^{\downarrow} \Rightarrow$ 2-D clustered $SE \approx$ 1-D clustered SE with clustering $\max(T,N)$

4. ${}^{\exists}$firm effect but not fixed (non-permanent, i.e. decay over time)

- ${}^{\exists}$firm effect & serial correlation

$$\begin{aligned} \eta_{it} &= \zeta_{it} && \text{if } t=1 \\ &= \phi\eta_{it-1} + \sqrt{1-\phi^2}\,\zeta_{it} && \text{if } t>1 \end{aligned}$$

$$\begin{aligned} \Rightarrow Corr(\varepsilon_{it},\varepsilon_{it-k}) &= \frac{Cov(\gamma_i+\eta_{it},\gamma_i+\eta_{it-k})}{\sqrt{Var(\gamma_i+\eta_{it})}\sqrt{Var(\gamma_i+\eta_{it-k})}} \\ &= \frac{\sigma_\gamma^2+\phi^k\sigma_\eta^2}{\sigma_\gamma^2+\sigma_\eta^2} \\ &= \rho_\varepsilon+(1-\rho_\varepsilon)\phi^k \end{aligned}$$

Panel A: OLS and clustered standard errors

Avg(β_{OLS}) Std(β_{OLS}) Avg(SE_{OLS}) % Sig(T_{OLS}) Avg(SE_C) % Sig(T_C)	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
OLS	0.9994	1.0001	1.0009	0.9991
	0.0513	0.0659	0.0566	0.0677
	0.0283	0.0283	0.0283	0.0253
	[0.1578]	[0.2746]	[0.1996]	[0.3302]
	0.0508	0.0668	0.0569	0.0670
	[0.0114]	[0.0086]	[0.0104]	[0.0098]
OLS with firm dummies	1.0007	1.0003	1.0013	1.0046
	0.0299	0.0517	0.0442	0.1881
	0.0298	0.0298	0.0298	0.1101
	[0.0096]	[0.1382]	[0.0802]	[0.1288]
	0.0298	0.0516	0.0441	0.1886
	[0.0100]	[0.0098]	[0.0092]	[0.0108]

$^{+}$firm dummy
$\Rightarrow$ $^{\nexists}$bias in OLS SE

Panel B: GLS estimates with and without clustered standard errors

Avg(β_{GLS}) Std(β_{GLS}) Avg(SE_{GLS}) % Sig(T_{GLS}) Avg(SE_{GLS-C}) % Sig(T_{GLS-C})	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
GLS	1.0005	1.0003	1.0012	1.0006
	0.0284	0.0475	0.0408	0.0731
	0.0283	0.0283	0.0283	0.0580
	[0.0090]	[0.1240]	[0.0730]	[0.0388]
	0.0282	0.0474	0.0408	0.0721
	[0.0090]	[0.0100]	[0.0090]	[0.0112]

correlated residual $\rightarrow$ efficiency; GLS$\gtrsim$OLS
decay serial correlation $\rightarrow$ β=efficient but SE=biased

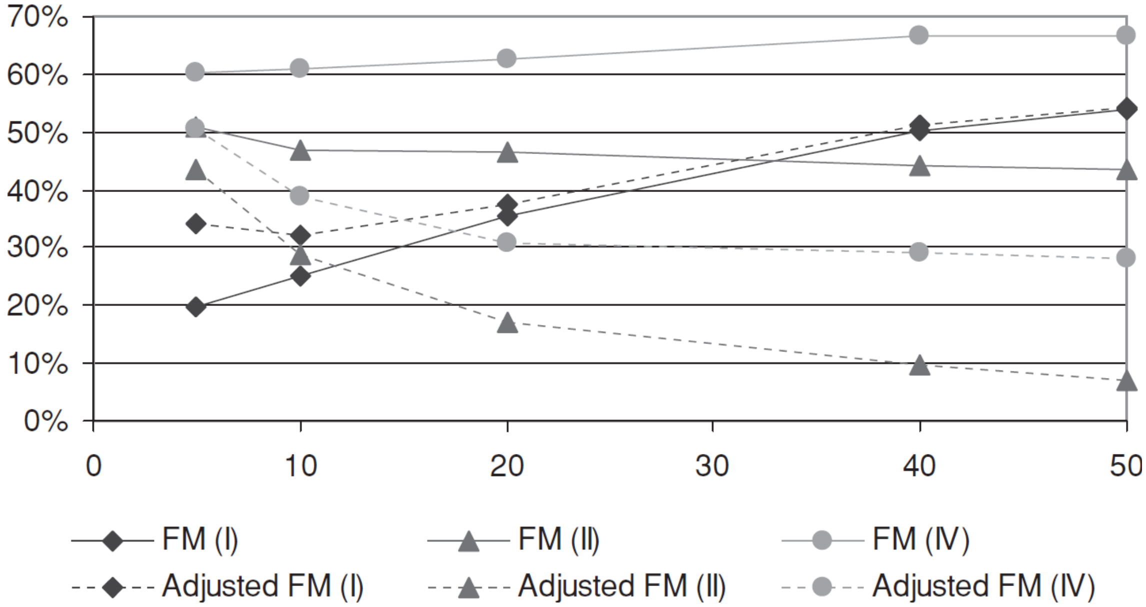
- fixed effect (hereafter FE) $\rightarrow$ significant minority
- efficiency; in many cases, fixed effect is better than OLS $\rightarrow$ $Var(\beta_{FIX}) \leqq Var(\beta_{OLS})$
- $^{\exists}$fixed firm effect $\Rightarrow$ SE of $\rightarrow$ OLS=bias, OLS+FE & clustered & clustered+FE=unbias
  - Kezdi (2004,)
- $^{\exists}$decaying firm effect $\Rightarrow$ OLS+FE=bias, clustered & clustered+FE=unbias
- $^{\exists}$decaying (temporary) firm effect $\Rightarrow$ $^{\exists}$bias$^{\downarrow}$ in SE of OLS+FE (depends on ϕ)
  - Baker, Stein, and Wurgler (2003, QJE), Wooldridge (2003, AER)
- GLS; improve efficiency of estimator; $Var(\beta_{GLS}) \leqq Var(\beta_{OLS})$ $\rightarrow$ Wooldridge (2007)
  - rarely(<3%) done in finance paper $\rightarrow$ Maksimovic and Phillips (2002, JF), Gentry, Kemsley, and Mayer (2003, JF), Almazan et al. (2004, JFE)

Panel C: Fama-MacBeth standard errors

Avg(β_{FM}) Std(β_{FM}) Avg(SE_{FM}) % Sig(T_{FM}) Avg(SE_{FM-AR1}) % Sig(T_{FM-AR1})	I	II	III	IV
ρ_X/ρ_ε	0.50/0.50	0.00/0.00	0.25/0.25	0.60/0.35
ϕ_X/ϕ_ε	0.00/0.00	0.90/0.90	0.75/0.75	0.99/0.81
Fama-MacBeth	0.9995	1.0001	1.0008	0.9991
	0.0514	0.0660	0.0567	0.0667
	0.0239	0.0187	0.0221	0.0138
	[0.2510]	[0.4696]	[0.3350]	[0.6094]
	0.0224	0.0389	0.0376	0.0289
	[0.3222]	[0.2876]	[0.2098]	[0.3900]
Avg(first-order autocorrelation)	−0.1157	0.4395	0.3250	0.4389

very imprecise SE because of negative correlation in cross-sectional β for each year

- adjusted Fama-MacBeth SE; $SE_{aFM}=SE_{FM}\times(1+\vartheta)/(1-\vartheta)$, $\vartheta=Corr(\beta_t,\beta_{t-1})$
  - Christopherson, Ferson, and Glassman (1998, RFS)
  - Graham, Lemmon, and Schallheim (1998, JF)
  - Chen, Hong, and Stein (2001, JFE)
  - Cochrane (2001)
  - Lakonishok and Lee (2001, RFS)
  - Fama and French (2002, RFS)
  - Kemsley and Nissim (2002, JF)
  - Bakshi, Kapadia, and Madan (2003, RFS)
  - Pastor and Veronesi (2003, JF)
  - Chakravarty, Gulen, and Mayhew (2004, JF)
  - Nagel (2005, JFE)
  - Schultz and Loughran (2005, JFE)
- 90% confidence interval for $\vartheta=(-0.60,0.41)$, $E(\vartheta)=-0.12$ $\rightarrow$ $\mathbb{P}(|t|>2.58)=25\sim32\%$
- negative $\vartheta\Rightarrow$ magnitude of bias$^{\uparrow}$; reason=estimated autocorr.$\neq$population autocorr.
  - positive correlation in β_t and β_{t-1} $\Rightarrow$ Fama-MacBeth SE bias$^{\downarrow}$
  - however in sample $E(\beta)$ is also influenced by firm effect $\rightarrow$ asymptotic covariance between them will be zero $\rightarrow$ adjustment based on this measure cannot be justified
  - $\phi>0$ $\Rightarrow$ performance$^{\uparrow}$ of SE of aFM



- case (1, ◆) → FM≳aFM & $T^{\uparrow}$=useless
- case (2, ▲) → aFM≳FM & $T^{\uparrow}$⇒magnitude of bias$^{\downarrow}$
- case (3, ●) → similar

● $^{\exists}$fixed & temporary firm effect ⇒ SE_{aFM}=biased but less than SE_{FM}

5. Empirical Applications

● 2 examples of panel analysis
  - asset pricing; usually time-effect matters
  - corporate finance; usually firm-effect matters

● in real world → data structure=unknown

● in this chapter
  - how the different methods for estimating *SE* compare
  - confirms that the methods used by some papers could have produced *SE*
  - shows what can be learned from the different *SE* estimates

● *SE* estimated by; OLS, White HCC, 1-D clustered by t, i & 2-D clustered
  - comparison between White HCC and clustered *SE*; no heteroskedasticity problem
  - if clustered *SE*≫White HCC *SE*⇒$^{\exists}$significant firm effect

I. OLS β + time dummy + White *SE*
II. OLS β + time dummy + firm clustered *SE*
III. OLS β + time dummy + time clustered *SE*
IV. OLS β + time dummy + 2-D clustered *SE*
V. Fama-MacBeth β + Fama-MacBeth *SE*

● Daniel and Titman (2006, JF) → $\text{return}_{it}=f(\text{B/M}_{it-1})$?
  - return=monthly, B/M=yearly ⇒ highly persistent
  - White *SE*≈clustered *SE* by firm ∵ autocorr. in residual=effectively 0
  - clustered *SE* by time≫other *SE* ∵ heavy time effect in the data

Table 6
Asset pricing application: Equity returns and asset tangibility

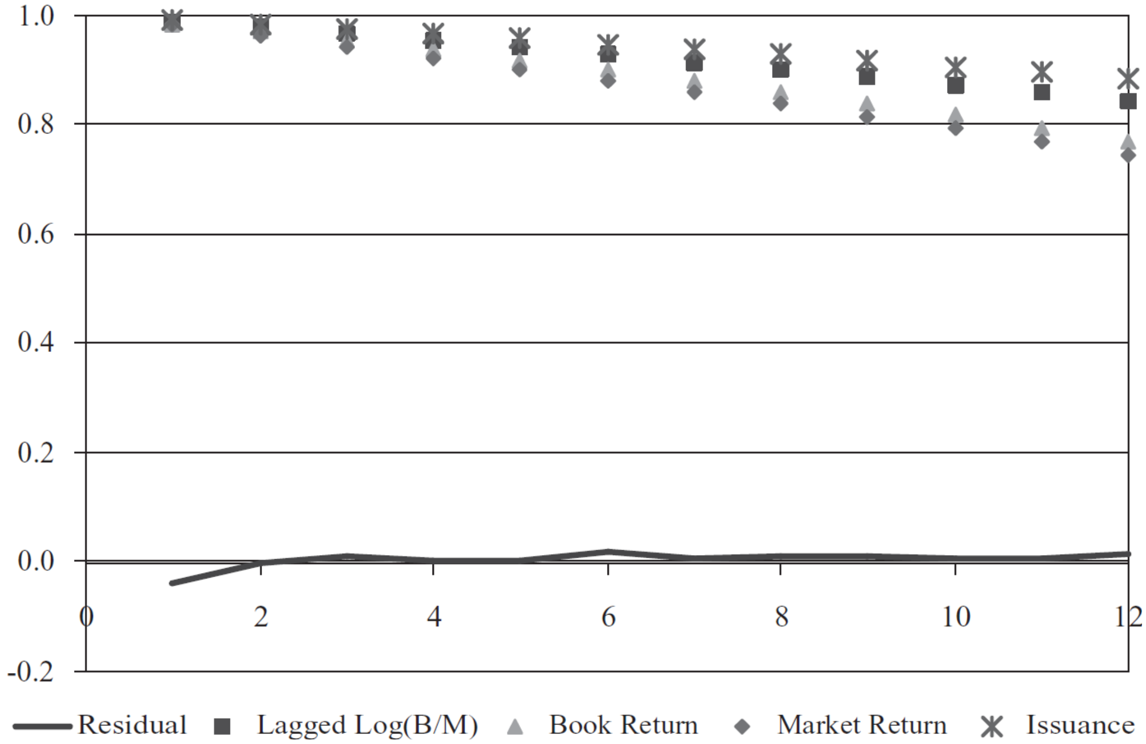
	I	II	III	IV	V
$\text{Log}(B/M)_{t-5}$	0.2456**	0.2456**	0.2456**	0.2456**	0.2064**
	(0.0238)	(0.0247)	(0.0924)	(0.0926)	(0.0795)
Log(Book Return) (last five years)	0.2482**	0.2482**	0.2482**	0.2482**	0.2145**
	(0.0385)	(0.0396)	(0.0859)	(0.0864)	(0.0788)
Market return (last five years)	−0.3445*	−0.3445*	−0.3445*	−0.3445*	−0.3310**
	(0.0257)	(0.0261)	(0.1000)	(0.1001)	(0.0893)
Share issuance	−0.5245**	−0.5245**	−0.5245**	−0.5245**	−0.5143**
	(0.0426)	(0.0427)	(0.1440)	(0.1441)	(0.1235)
R^2	0.0008	0.0008	0.0008	0.0008	0.0008
Coefficient estimates	OLS	OLS	OLS	OLS	FM
Standard errors	White	$CL-F$	$CL-T$	$CL-F\&T$	FM

- temporary firm effect; understanding is simple (time decaying)
- non-constant time effect; difficult (time dummy SE=clustered SE by time $\rightarrow$ $^{\exists}$constant effect, otherwise i.e. time dummy $SE \neq$clustered SE by time$\Rightarrow$ $^{\exists}$non-constant effect
- economic intuition can be applied into the analysis$\Rightarrow$matter of sorting order
- It can be found the structure of given data by comparing the magnitude of estimated SE of many models.
- III and V=similar $\because$ $^{\exists}$firm effect

Table 7
Corporate finance application: Capital structure regressions (1965–2003)

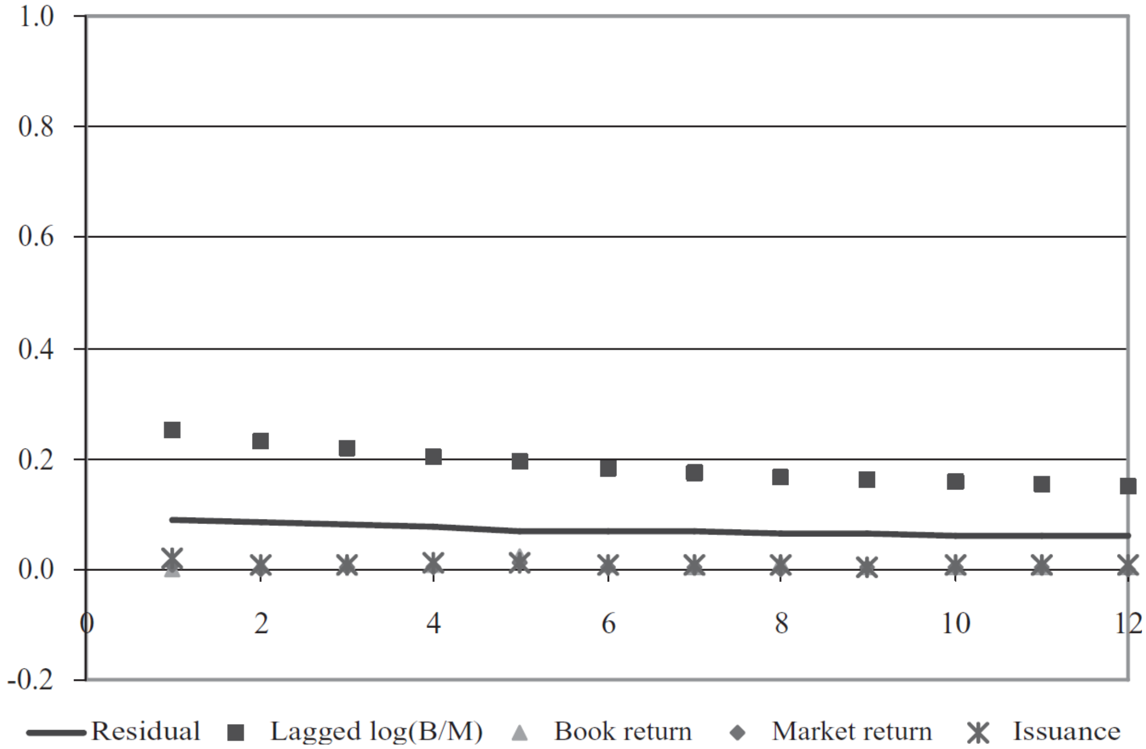
	I	II	III	IV	V
Ln(MV assets)	0.0460**	0.0460*	0.0460**	0.0460*	0.0394**
	(0.0055)	(0.0184)	(0.0074)	(0.0191)	(0.0076)
Ln(1 + Firm age)	−0.0432**	−0.0432	−0.0432**	−0.0432	−0.0479**
	(0.0084)	(0.0297)	(0.0067)	(0.0293)	(0.0077)
Profits/sales	−0.0330**	−0.0330	−0.0330**	−0.0330	−0.0299**
	(0.0107)	(0.0359)	(0.0098)	(0.0357)	(0.0097)
Tangible assets	0.1043**	0.1043**	0.1043**	0.1043**	0.1158**
	(0.0057)	(0.0197)	(0.0083)	(0.0206)	(0.0096)
Market-to-book (assets)	−0.0251**	−0.0251**	−0.0251**	−0.0251**	−0.0272**
	(0.0006)	(0.0020)	(0.0013)	(0.0023)	(0.0016)
Advertising/sales	−0.3245**	−0.3245	−0.3245**	−0.3245	−0.3965*
	(0.0841)	(0.2617)	(0.0814)	(0.2609)	(0.1712)
R&D/sales	−0.3513**	−0.3513*	−0.3513**	−0.3513*	−0.3359**
	(0.0469)	(0.1544)	(0.0504)	(0.1555)	(0.0501)
R&D > 0 (= 1 if yes)	0.0177**	0.0177*	0.0177**	0.0177*	0.0126**
	(0.0024)	(0.0076)	(0.0025)	(0.0077)	(0.0034)
R-squared	0.1360	0.1360	0.1360	0.1360	0.1300
Coefficient estimates	OLS	OLS	OLS	OLS	FM
Standard errors	White	$CL-F$	$CL-T$	$CL-F\&T$	FM

Panel A: Within Firm



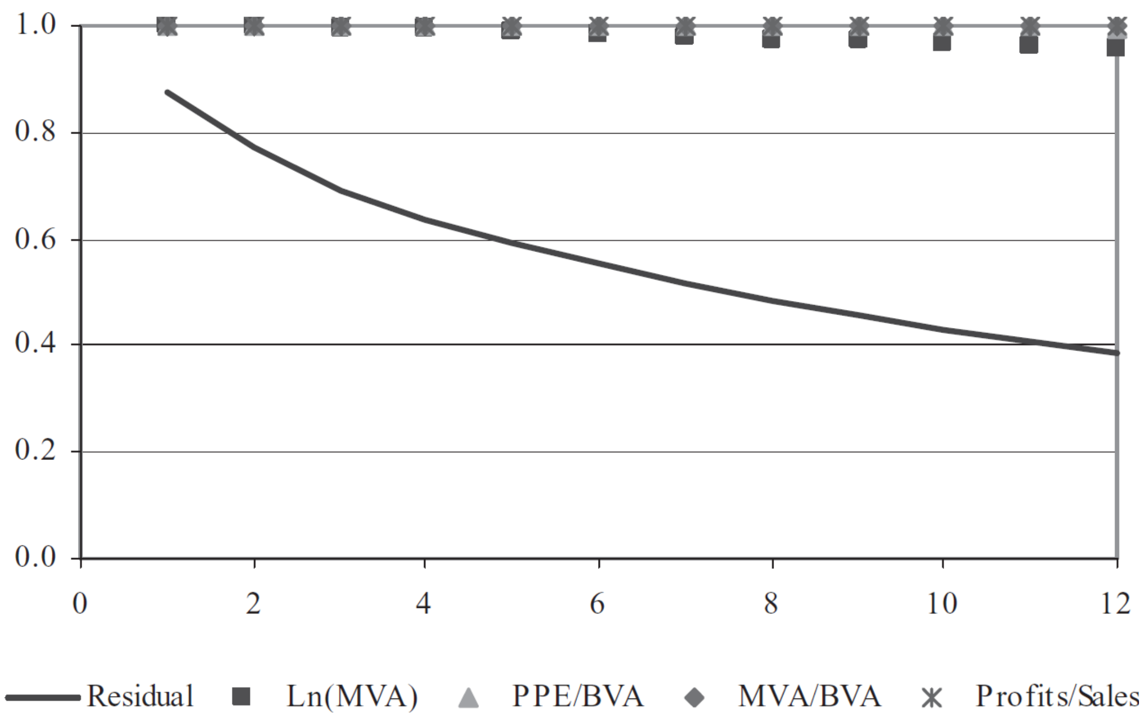
- ∄ residual autocorr. ∃ high autocorr. in B/M

Panel B: Within Month



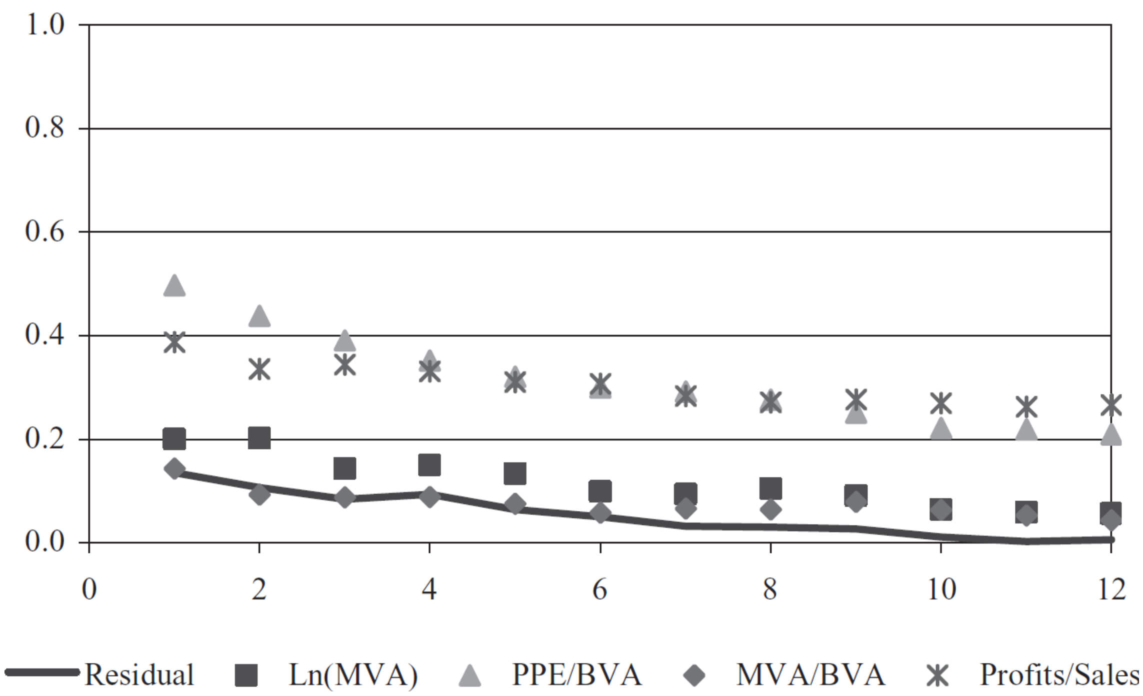
- ∃ correlation between industry at the same time $\Rightarrow$ time effect is more important

Panel A: Within Firm



- $\exists$ high autocorrelation in dep. and indep. variable $\Rightarrow$ firm effect matters!

Panel B: Within Month



- cross sectional correlation$^{\downarrow}$ $\Rightarrow$ time effect is small in this case!

6. conclusion; again, this paper is about *SE*!!! not about β!

- empirical finance $\Rightarrow$ panel analysis is important
- but! OLS or White HCC for regression *SE* are biased in usual situation
- but!! appropriate remedies have rarely been proposed
- objective of this paper=to provide good guide for *SE* problems in panel analysis

$^{\exists}$firm effect $\Leftrightarrow$ $^{\exists}$serial correlation in X and ε $\Leftrightarrow$ highly persistent; usually in corporate finance data

- best=clustered *SE* by firm (whether the effect is fixed or not)
- good=fixed effect OLS, random effect GLS (only when firm effect is fixed)
- not good=simple OLS, White HCC, Newey-West for panel, Fama-MacBeth *SE*, first-order modified Fama-MacBeth *SE* (all of them will underestimate *SE*!)

$^{\exists}$time effect $\Leftrightarrow$ $^{\exists}$cross correlation $\Leftrightarrow$ low persistent but systematically influenced by macro or etc.; usually in asset pricing data

- best=Fama-MacBeth *SE*
- good=clustered *SE* by time (only when $T^{\uparrow}$ enough, otherwise$\Rightarrow$biased)
- not good=others

$^{\exists}$time effect & $^{\exists}$firm effect

- best=2-D clustered SE (only when $T^{\uparrow}$ and $N^{\uparrow}$)
- good=parametric on one dimension+clustered *SE* for other dimension $\rightarrow$ ex. time dummy+clustered *SE* by firm

if $^{\exists}$sufficient N and T $\Rightarrow$ 2-D clustered *SE*=always best
if not $\Rightarrow$ dummy for small dimension+clustered *SE* by large dimension is quasi-best

"OLS *SE*=bias" means that "$^{\exists}$remained information in the data"

- In this situation, researcher can improve the efficiency of β by using fixed effects, GLS, GMM, etc. to test.

comparing various *SE* estimates $\Rightarrow$ quickly observe the presence of firm/time effect

- if White *SE*$\gg$clustered *SE* by firm $\Rightarrow$ $^{\exists}$heavy firm effect
- if White *SE*$\gg$clustered *SE* by time $\Rightarrow$ $^{\exists}$heavy time effect
- if 2-D clustered *SE*$\gg$one of 1-D clustered *SE* $\Rightarrow$ $^{\exists}$firm effect & $^{\exists}$time effect

Which dependencies are most important will vary across data and thus researchers must consult their data.

- this paper gives; appropriate selecting method for *SE*, intuition as to the the deficiency of their models, guidance for improving their models

Firm Effect Concept!

Firm / Time	A	B	C	$\cdots$	N
2001	a	b	c	$\cdots$	n
2002	a	b	c	$\cdots$	n
2003	a	b	c	$\cdots$	n
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
T	a	b	c	$\cdots$	n

Time Effect Concept!

Firm / Time	A	B	C	$\cdots$	N
2001	1	1	1	$\cdots$	1
2002	2	2	2	$\cdots$	2
2003	3	3	3	$\cdots$	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
T	t	t	t	$\cdots$	t

2-D Effect Concept!

Firm / Time	A	B	C	$\cdots$	N
2001	1	1	1	$\cdots$	1
2002	2	2	2	$\cdots$	2
2003	3	3	3	$\cdots$	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
T	t	t	t	$\cdots$	t