パネル分析におけるFama-MacBethとCluster-robustの手法の理論と応用

太　田　浩　司

要約

本稿では，パネル推定で，誤差項が標準的仮定を満たしていない場合の対処法として会計・ファイナンスの実証研究で頻繁に用いられているFama-MacBeth（FM）およびCluster-robust（CR）の手法について，理論的分析を行うと共に，その使用に関する実務的指針を述べている。本稿の理論的分析からは，FMとCRの手法ではその発想が全く異なっているが，説明変数の変動が年度間でそれ程大きくない場合には，FMの手法は，年度に関するOne-way CRの手法と同様に，誤差項の未知のクロスセクショナルな相関に対処することができるという結論を得ている。さらに，企業数に比して年度数の少ないショート・パネルを用いた場合には，FMの手法，Two-way CRの手法，年度に関するOne-way CRの手法の何れの手法を用いても，得られる標準誤差には大差がないという考察を導いており，実証分析においてもそれを支持する結果を得ている。

本稿の結果は，その広範な普及にも関わらずこれまで理論的な分析が殆ど行われてこなかったFMの手法を理論的に解明し，CRの手法との関連性を導いている点で，今後の両手法の使用に関して一定の指針を与えるものと考えられる。

キーワード：Fama-MacBeth (1973)手法，Cluster-robust手法，頑健性のある標準誤差

１．はじめに

会計・ファイナンス領域における実証研究では，企業と年度の２つのディメンションから成るパネル・データを用いた分析が頻繁に行われるが，企業と年度のパネル分析では，同一企業の異なる年度や同一年度の異なる企業の誤差項間に相関が生じることが多く，誤差項の標準的仮定（均一分散で互いに無相関）が満たされないケースが多い。そのような場合には，推定モ

デルの係数推定値の分散が正しく求められず（有効推定量でなくなってしまう），統計的検定が適切に行われているという保証が得られない。従って，本来有意ではない係数を有意であると判断してしまって，結果として誤った結論を導いてしまう可能性がある。

このような，企業と年度のパネル推定において誤差項が標準的仮定を満たしていないと思われる場合の対処法として，会計・ファイナンスの実証研究で最も頻繁に用いられているのが，Fama-MacBeth（以下FM）の手法ならびにCluster-robust（以下CR）の手法である。

最初に，FMの手法は，Fama and MacBeth (1973)によって提案されたもので，会計・ファイナンスの実証研究において古くから頻繁に用いられている。しかしながら，FMの手法は，その広範な普及にもかかわらず，理論的背景についてはCochrane (2005)で若干取り上げられている程度で，十分な研究が行われていない。また，FMの手法には幾つかのバリエーションが存在しているが，何れもアドホックなものであり，その理論的考察は殆ど行われていない。

次に，CRの手法には，One-way CRとその応用版であるTwo-way CRの２つの手法が存在する。One-way CRは，Liang and Zeger (1986)やArellano (1987)によって提案されたが，会計・ファイナンスの実証分野ではあまり注目を集めなかった。ところが，Cameron *et al.* (2011)やThompson (2011)によりその応用版であるTwo-way CRの手法が提示され，また，Petersen (2009)によってその統計ソフトプログラムが公開されたことにより，近年の実証研究において，その使用が急速に拡大している。しかしながら，CRの手法については最近の研究が多いこともあってか，その理論的背景や実際に使用する際の問題点についてはあまり知られておらず，その適切な使用方法は未だ十分に認知されていない。

そこで本稿では，企業と年度のパネル推定において誤差項が標準的仮定を満たしていない場合の対処法として，実証研究で最も頻繁に用いられているFMの手法ならびにCRの手法について，第２章で，FMの手法に関する理論的分析を行い，第３章で，CRの手法についてその理論的概要および使用に関する実務的指針について叙述し，第４章でFMの手法とCRの手法の特徴を比較している。また，第５章で，実際のデータを用いて，これらの手法から得られる標準誤差を用いることによって，係数の有意性検定の結果がどのような影響を受けるのかを具体例を用いて調査し，最後に，第６章で，本稿の総括を行っている。

２．Fama-MacBethの手法

FMの手法は，会計・ファイナンスの実証研究で広範に使用されているにもかかわらず，その理論的背景については，Cochrane (2005, pp. 245-251)で取り上げられている程度で，十分な研究が行われているとはいえない。そこで，本章の第１節では，FMの手法の概要について述べ，第２節以降では，FMの標準誤差が厳密に正しく推定されるための条件，適度に正しく推定されるための条件，そして，FMのバリエーションである調整済みFMの手法について順

番に説明する。

（1）Fama-MacBethの手法の概要

最初に，企業と年度のパネル・データを用いる実証研究では，通常，以下のような回帰モデルが使用される。

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \varepsilon_{it}\ ,$$

ただし，i (1, 2, ⋯ , N)は個別企業を表し，t (1, 2, ⋯ , T)は年度を表している。

このパネル・データをプールした時の回帰式を，

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

と表わすとする。このとき，Pooled OLSの係数推定量とその分散推定量は，

$$\hat{\boldsymbol{\beta}}_{Pooled} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y},$$

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{Pooled}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1},\ \text{ただし，}\ E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \boldsymbol{\Omega},$$

と表わされる。

しかしながら，このままでは，誤差項の共分散行列$\boldsymbol{\Omega}$が未知であるので，係数の有意性検定を適切に行うことができない。そこで，FMの手法では，サンプルを，年度毎に分割し，

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{bmatrix},\quad \mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_T \end{bmatrix},\quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_T \end{bmatrix},$$

これを各年クロスセクションで回帰するのである。

$$\begin{aligned} t = 1:&\ \ \mathbf{y}_1 = \mathbf{x}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1, \\ t = 2:&\ \ \mathbf{y}_2 = \mathbf{x}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2, \\ \vdots\ \ &\ \ \ \ \ \ \ \ \vdots \\ t = T:&\ \ \mathbf{y}_T = \mathbf{x}_T\boldsymbol{\beta}_T + \boldsymbol{\varepsilon}_T. \end{aligned}$$

そして，得られた，T個の$\hat{\boldsymbol{\beta}}_t$ $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \ldots, \hat{\boldsymbol{\beta}}_T)$を用いて，

$$\hat{\boldsymbol{\beta}}_{FM} = \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{\beta}}_t\ , \tag{1a}$$

$$\mathrm{Est.Var}[\hat{\boldsymbol{\beta}}_{FM}] = \frac{1}{T(T-1)}\sum_{t=1}^{T}(\hat{\boldsymbol{\beta}}_t - \hat{\boldsymbol{\beta}}_{FM})(\hat{\boldsymbol{\beta}}_t - \hat{\boldsymbol{\beta}}_{FM})'\ \text{（標本平均の分散）}, \tag{1b}$$

を求めて，$\hat{\boldsymbol{\beta}}_{FM}$ の有意性検定を行うのである。

（2）Fama-MacBethの標準誤差が厳密に正しく推定されるための条件

FMの標準誤差は(1b)から求められるが，このFMの標準誤差が厳密に正しく推定されるた

めには，T個の $\hat{\boldsymbol{\beta}}_t$ $(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \dots, \hat{\boldsymbol{\beta}}_T)$ に関して，以下の条件が成り立つ必要がある。

A1: $E[\hat{\boldsymbol{\beta}}_t] = \boldsymbol{\beta}$ 　for all t $(t = 1, 2, \dots, T)$,

A2: $\mathrm{Var}[\hat{\boldsymbol{\beta}}_t]$が一定である　for all t $(t = 1, 2, \dots, T)$,

A3: $\mathrm{Cov}[\hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s] = \mathbf{0}$ 　for all $t \neq s$ $(t, s = 1, 2, \dots, T)$.

最初に，A1は，

$$E[\hat{\boldsymbol{\beta}}_t] = \boldsymbol{\beta}_t + (\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}'E[\boldsymbol{\varepsilon}_t] = \boldsymbol{\beta}_t = \boldsymbol{\beta},$$

が全ての年度において成り立つという条件である。これは，どの年度においても係数パラメータ$\boldsymbol{\beta}$は不変である，

$$\mathbf{y}_t = \mathbf{x}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t \quad \text{for all } t \ (t{=}1, 2, \dots\dots, T)$$

という関係が成り立っているということを意味しており，最初のモデル設定から自明である。

次に，A2は，

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_t] = (\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']\mathbf{x}_t(\mathbf{x}_t'\mathbf{x}_t)^{-1},$$

が全ての年度において一定であるということなので，

$$\mathbf{x}_t = \mathbf{x},\ E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \boldsymbol{\Sigma} \quad \text{for all } t \ (t = 1, 2, \dots, T),$$

という仮定が必要である。つまり，A2は，どの年度においても説明変数と誤差項の共分散行列は不変であるということを意味している。

最後に，A3は，

$$\mathrm{Cov}[\hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s] = (\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s']\mathbf{x}_s(\mathbf{x}_s'\mathbf{x}_s)^{-1} = \mathbf{0} \quad \text{for all } t \neq s \ (t, s = 1, 2, \dots, T),$$

ということなので，異なる年度間の誤差項には相関がない，

$$E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s'] = \mathbf{0} \quad \text{for all } t \neq s \ (t, s = 1, 2, \dots, T),$$

ということを意味している。

以上をまとめると，FMの標準誤差が厳密に正しく推定されるためには，回帰モデルのデータ構造が，以下のように表される必要があるといえる。

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

ただし，

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \\ \vdots \\ \mathbf{x} \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_T \end{bmatrix}, \quad E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \boldsymbol{\Sigma}, \quad E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Sigma} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma} \end{pmatrix}.$$

このとき，Pooled OLSの係数推定量とその分散推定量は，

$$\hat{\boldsymbol{\beta}}_{Pooled} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\frac{1}{T}\sum_{t=1}^{T}\mathbf{y}_t, \tag{2a}$$

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{Pooled}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{T}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}, \tag{2b}$$

と表わされる。

一方，FMの係数推定量とその漸近分散推定量は，次のように表される。最初に，係数推定量は，

$$\hat{\boldsymbol{\beta}}_{FM} = \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{\beta}}_t = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'\mathbf{y}_t = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\frac{1}{T}\sum_{t=1}^{T}\mathbf{y}_t, \tag{3a}$$

となる。次に，漸近分散推定量は，大数の弱法則から，$\mathrm{plim}_{T\to\infty}\hat{\boldsymbol{\beta}}_{FM} = \boldsymbol{\beta}$が成り立つので，

$$\begin{aligned}\mathrm{Asy.Var}[\hat{\boldsymbol{\beta}}_{FM}] &= \frac{1}{T(T-1)}\sum_{t=1}^{T}(\hat{\boldsymbol{\beta}}_t - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_t - \boldsymbol{\beta})' \\ &= \frac{1}{T^2}\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']\mathbf{x}_t(\mathbf{x}_t'\mathbf{x}_t)^{-1} \\ &= \frac{1}{T}(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1},\end{aligned} \tag{3b}$$

となる。

(2a) (2b) と (3a) (3b) が同一であることから，A1-3の条件の下では，FMの標準誤差は正しい推定量であるといえる。

このように，FMの標準誤差が厳密に正しく推定されるためには，説明変数と誤差項の共分散行列が全ての年度で同じであり，かつ，異なる年度間では誤差項に相関がないという強い仮定が必要となってくる。その一方で，FMの手法では，$\boldsymbol{\Sigma}$を特定化する必要がなく，誤差項がクロスセクショナルにどのような構造であっても，係数推定量の分散を正しく推定することができるというメリットも存在している。

（3）Fama-MacBethの標準誤差が適度に正しく推定されるための条件

前節のFMの標準誤差が厳密に正しく推定されるための条件では，A2で，どの年度においても説明変数と誤差項の共分散行列は不変であるという非常に厳しい仮定がおかれており，これは現実のデータにはそぐわないものである。そこで，A2の仮定を緩和して，$\hat{\boldsymbol{\beta}}_t$の分散が有限でさえあれば良いというものに変えた場合の影響について考える。

B1:　$E[\hat{\boldsymbol{\beta}}_t] = \boldsymbol{\beta}$　for all t $(t = 1, 2, \ldots, T)$,

B2:　$\mathrm{Var}[\hat{\boldsymbol{\beta}}_t]$が有限である　for all t $(t = 1, 2, \ldots, T)$,

B3:　$\mathrm{Cov}[\hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s] = \mathbf{0}$　for all $t \neq s$ $(t, s = 1, 2, \ldots, T)$.

これらの仮定を満たす回帰モデルのデータ構造は，以下のように表される。

$$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon},$$

ただし，

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_T \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_T \end{bmatrix}, \quad E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] = \boldsymbol{\Sigma}_t, \quad E[\boldsymbol{\varepsilon\varepsilon}'] = \boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_T \end{pmatrix}.$$

このデータ構造からもわかるように，B1は自明であり，B2も通常成り立つ条件であるので，実際に制約となっているのは，B3の異なる年度間の誤差項には相関がないという条件だけである。

このとき，Pooled OLSの係数推定量とその分散推定量は，

$$\hat{\boldsymbol{\beta}}_{Pooled} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_t'\mathbf{x}_t\right)^{-1}\mathbf{x}_t'\mathbf{y}_t, \tag{4a}$$

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{Pooled}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{T^2}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_t'\mathbf{x}_t\right)^{-1}\mathbf{x}_t'\boldsymbol{\Sigma}_t\mathbf{x}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_t'\mathbf{x}_t\right)^{-1}, \tag{4b}$$

と表わされる。

一方，FMの係数推定量とその漸近分散推定量は，次のように表される。最初に，係数推定量は，

$$\hat{\boldsymbol{\beta}}_{FM} = \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{\beta}}_t = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'\mathbf{y}_t, \tag{5a}$$

と表わされる。次に，漸近分散推定量は，B2より$\mathrm{Var}[\hat{\boldsymbol{\beta}}_t]$が有限であるので，

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_t] = \sigma_{\beta t}^2 \le \sigma_\beta^2 \quad \text{for all } t\ (t = 1, 2, \dots, T),$$

と表わされ，さらに，B3より$\mathrm{Cov}[\hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s] = \mathbf{0}$であるので，

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{FM}] = \frac{\sigma_{\beta 1}^2 + \sigma_{\beta 2}^2 + \cdots + \sigma_{\beta T}^2}{T^2} \le \frac{\sigma_\beta^2 + \sigma_\beta^2 + \cdots + \sigma_\beta^2}{T^2} = \frac{\sigma_\beta^2}{T}.$$

従って，チェビシェフの不等式より，任意の正の実数$k > 0$について，

$$\mathrm{Prob}(|\hat{\boldsymbol{\beta}}_{FM} - E[\hat{\boldsymbol{\beta}}_{FM}]| \ge k) \le \frac{\mathrm{Var}[\hat{\boldsymbol{\beta}}_{FM}]}{k^2},$$

$$\mathrm{Prob}(|\hat{\boldsymbol{\beta}}_{FM} - \boldsymbol{\beta}| \ge k) \le \frac{\sigma_\beta^2}{Tk^2},$$

が成り立つ。ここで，$T \to \infty$の極限をとると，

$$\lim_{T\to\infty}\mathrm{Prob}(|\hat{\boldsymbol{\beta}}_{FM} - \boldsymbol{\beta}| \ge k) = 0.$$

つまり，$\hat{\boldsymbol{\beta}}_t$の分散が異なっていても，$\mathrm{plim}_{T\to\infty}\hat{\boldsymbol{\beta}}_{FM} = \boldsymbol{\beta}$ が成り立つのである。従って，FMの漸近分散推定量は，

$$\begin{aligned}\text{Asy.Var}[\hat{\boldsymbol{\beta}}_{FM}] &= \frac{1}{T(T-1)}\sum_{t=1}^{T}(\hat{\boldsymbol{\beta}}_t-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_t-\boldsymbol{\beta})' \\ &= \frac{1}{T^2}\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']\mathbf{x}_t(\mathbf{x}_t'\mathbf{x}_t)^{-1} \\ &= \frac{1}{T^2}\sum_{t=1}^{T}(\mathbf{x}_t'\mathbf{x}_t)^{-1}\mathbf{x}_t'\boldsymbol{\Sigma}_t\mathbf{x}_t(\mathbf{x}_t'\mathbf{x}_t)^{-1},\end{aligned} \tag{5b}$$

と表わされる。

ここで，Pooled OLSの係数推定量(4a)および分散推定量(4b)と，FMの係数推定量(5a)および漸近分散推定量(5b)を比較すると，両者の推定量が大変よく似ており，その唯一の違いは，FMが$\mathbf{x}_t'\mathbf{x}_t$を用いているところを，Pooled OLSではその平均値である，$\frac{1}{T}\sum_{t=1}^{T}\mathbf{x}_t'\mathbf{x}_t$ を用いている点だけであるということがわかる。つまり，係数推定量では $\mathbf{x}_t'\mathbf{y}_t$，分散推定量では $\mathbf{x}_t'\boldsymbol{\Sigma}_t\mathbf{x}_t$ に対する重み付けだけが，Pooled OLSとFMでは異なっているのである。

このことは，実用面からいうと，年度に亘って説明変数にそれほど大きな変化がない場合には，FMの手法を用いることによって，誤差項の未知のクロスセクショナルな相関$\boldsymbol{\Sigma}_t$に対処できるということを意味している。

(4) 調整済みFama-MacBethの手法

前節では，FMの標準誤差が適度に正しく推定されるための条件として，B1-3の３つを挙げているが，実際に制約となっているのは，B3の異なる年度間の誤差項には相関がないという条件だけである。調整済みFMの標準誤差（Adjusted Fama-MacBeth Standard Errors）は，このB3の条件が満たされない場合の対処方法として提案されているものである。

調整方法は幾つかあるが，何れも，各年の係数推定値 $\hat{\beta}_t$ の自己相関係数を用いて，通常のFMの標準誤差を調整するというものである。なお，以下で述べる調整方法は，各説明変数の係数推定量に関してひとつずつ個別に行われるということには注意されたい。

最初に，Chakravarty *et al.* (2004)では，T個の$\hat{\beta}_t$の1階の自己相関係数,

$$\text{Corr}[\hat{\beta}_t,\hat{\beta}_{t-1}]=\rho,$$

を求めて，通常のFMの分散推定量に$(1+\rho)/(1-\rho)$を乗じるという方法が提案されている,

$$\text{adj.Var}[\hat{\beta}_{FM}]=\frac{1+\rho}{1-\rho}\frac{1}{T(T-1)}\sum_{t=1}^{T}(\hat{\beta}_t-\hat{\beta}_{FM})^2.$$

次に，Fama and French (2002)では,

$$\text{Var}[\hat{\beta}_t]=\sigma^2,\quad \text{Corr}[\hat{\beta}_t,\hat{\beta}_{t-s}]=\rho^s,$$

と仮定すると，T が大きい場合には，T個の $\hat{\beta}_t$ の合計の分散が,

$$\text{adj.Var}[\sum_{t=1}^{T}\hat{\beta}_t]\simeq T\sigma^2+2\sigma^2T\frac{\rho}{1-\rho},$$

で近似されることを示している。そして，ρ ＝ 0.75と仮定すると，$\text{adj.Var}[\sum_{t=1}^{T}\hat{\beta}_t]=7T\sigma^2$であるので，$\rho$ ＝ 0と仮定する通常のFMの分散推定量を7倍するという方法を提案している[1)]。

最後に，FMの分散推定量に，Newey and West（1987）が提案した誤差項の系列相関に対して頑健性のある標準誤差を求める手法と同様の調整を行うという方法も提案されている。初めに，被説明変数をT個の$\hat{\beta}_t$，説明変数を定数項ベクトルとする以下の回帰式を設定する，

$$\mathbf{y}=\mathbf{x}\beta_{FM}+\boldsymbol{\varepsilon},\text{ただし，}\mathbf{y}=(\hat{\beta}_1\quad\hat{\beta}_2\quad\cdots\quad\hat{\beta}_T)',\quad\mathbf{x}=(1\quad 1\quad\cdots\quad 1)'. \tag{6}$$

そして，(6)を推定すると，

$$\hat{\beta}_{FM}=(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y}=\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}_t,$$

$$\text{Var}[\hat{\beta}_{FM}]=(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x}(\mathbf{x}'\mathbf{x})^{-1}=\frac{1}{T^2}\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x},\text{ただし，}E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']=\boldsymbol{\Sigma},$$

が得られる。

このとき，(6)の誤差項が標準的仮定（均一分散で互いに無相関：$E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']=\sigma^2\mathbf{I}$）を満たすとすると，

$$\text{Est.Var}[\hat{\beta}_{FM}]=\frac{1}{T}\hat{\sigma}^2=\frac{1}{T}\frac{1}{(T-1)}\sum_{t=1}^{T}\hat{\varepsilon}_t^2=\frac{1}{T}\frac{1}{(T-1)}\sum_{t=1}^{T}(\hat{\beta}_t-\hat{\beta}_{FM})^2,$$

となり，(1b)と一致して，通常のFMの分散推定量が得られる。

一方，Newey-West流の方法では，(6)の誤差項に系列相関があると仮定し，適当な最大ラグLを設定して重みを付けると，

$$\text{adj.Var}[\hat{\beta}_{FM}]=\frac{1}{T^2}\mathbf{x}'\hat{\boldsymbol{\Sigma}}\mathbf{x},$$

ただし，

$$\hat{\boldsymbol{\Sigma}}=\begin{pmatrix}(\hat{\beta}_1-\hat{\beta}_{FM})^2 & \cdots & \left(1-\frac{L}{L+1}\right)(\hat{\beta}_1-\hat{\beta}_{FM})(\hat{\beta}_{L+1}-\hat{\beta}_{FM}) & \mathbf{0}\\ \vdots & & & \left(1-\frac{L}{L+1}\right)(\hat{\beta}_{T-L}-\hat{\beta}_{FM})(\hat{\beta}_T-\hat{\beta}_{FM})\\ \left(1-\frac{L}{L+1}\right)(\hat{\beta}_{L+1}-\hat{\beta}_{FM})(\hat{\beta}_1-\hat{\beta}_{FM}) & & \ddots & \vdots\\ \mathbf{0} & \left(1-\frac{L}{L+1}\right)(\hat{\beta}_T-\hat{\beta}_{FM})(\hat{\beta}_{T-L}-\hat{\beta}_{FM}) & \cdots & (\hat{\beta}_T-\hat{\beta}_{FM})^2\end{pmatrix},$$

が得られる。これが，Newey-West調整されたFMの分散推定量である。

例えば，L ＝ 1の場合は，

$$\text{adj.Var}[\hat{\beta}_{FM}]=\frac{1}{T^2}\sum_{t=1}^{T}(\hat{\beta}_t-\hat{\beta}_{FM})^2+\frac{1}{T^2}\sum_{t=2}^{T}(\hat{\beta}_t-\hat{\beta}_{FM})(\hat{\beta}_{t-1}-\hat{\beta}_{FM}),$$

と表わされる。右辺の第1項は，(1b)の通常のFMの分散推定量と同じであり，第2項は，$\hat{\beta}_t$の1階の自己共分散$\text{Cov}[\hat{\beta}_t,\hat{\beta}_{t-1}]$に$1/T$を乗じたものである。

1）実際には，通常のFM の標準誤差を2.5倍している。

このことからもわかるように，T個の$\hat{\beta}_t$を定数項ベクトルに回帰して得られる残差は，$\hat{\beta}_t - \hat{\beta}_{FM}$であるので，(6)の残差の系列相関を考慮するということは，$\hat{\beta}_t$の自己相関に対処することと密接な関係があるのである。

以上，３種類の調整済みFMの標準誤差について述べたが，これらは何れも，アドホックな対処方法であり，その性質に関する理論的な研究は未だ行われていない（Gow *et al.* 2010）。そもそも，係数推定値の時系列相関に対処することが，原データにおける，誤差項の異なる年度間の相関に対処することとどのような関係があるのかが明瞭ではない。また，これらの調整方法の有用性をシュミレーションを用いて検証した，Petersen (2009)およびGow *et al.* (2010)からは，これらの調整方法のパフォーマンスが悪いという結果が得られている。

このように，調整済みFMの標準誤差は，理論的根拠がなく，またシミュレーション結果も芳しくないことを考慮すると，安易に用いるべきではないであろう。

３．Cluster-robustの手法

企業と年度のパネル・データを用いる分析では，誤差項が，企業や年度でクラスタリングされるので，各クラスター内で誤差項に相関が生じているのではないかと考えるのは自然なことである。そのような誤差項のクラスタリングに対して頑健性のある標準誤差を求めるのが，CRの手法である。また，CRの手法には，企業と年度のどちらか片方だけのディメンションのクラスタリングに対処するOne-way CRの手法と，企業と年度の両方のディメンションのクラスタリングに対処するTwo-way CRの手法の２つがある[2)]。

そこで，本章では，第１節および第２節で，それぞれ，One-way CRの手法とTwo-way CRの手法の理論的概要を述べる。さらに，第３節で，CRの手法を実際に用いる際に問題となる点について叙述する。

(1) One-way Cluster-robustの手法

企業と年度のパネル・データを用いる分析で使用されるOne-way CRの手法には，企業に関するクラスタリングに対処する手法と，年度に関するクラスタリングに対処する手法の２種類があるが，本節では，企業に関するOne-way CRの手法について説明する。なお，年度に関するOne-way CRの手法も，企業についての場合と同様である。

最初に，パネル・データを用いた場合の回帰モデルは，以下のように表される。

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it},$$

ただし，i (1, 2, … , N)は個別企業を表し，t (1, 2, … , T)は年度を表しており，観測値数はバ

2) CRの手法には，さらに多元的にクラスタリングを行うMultiway CR というものも存在するが，基本的にOne- and Two-way CR を発展させたもので，その発想は同じである（Cameron *et al.* 2011）。

ランスト・パネルでNTであるとする。

次に，このパネル・データをプールした時の回帰式を，

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon},$$

で表わすとする。

このとき，誤差項のクラスタリングが企業に関してあるとすると，誤差項の共分散行列は，以下のように表される，

$$E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']=\begin{pmatrix}\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1' & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\varepsilon}_2\boldsymbol{\varepsilon}_2' & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\varepsilon}_N\boldsymbol{\varepsilon}_N'\end{pmatrix}=\boldsymbol{\Sigma},\quad \text{ただし，}\quad \boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'=\begin{pmatrix}\varepsilon_{i1}^2 & \varepsilon_{i1}\varepsilon_{i2} & \cdots & \varepsilon_{i1}\varepsilon_{iT}\\ \varepsilon_{i2}\varepsilon_{i1} & \varepsilon_{i2}^2 & \cdots & \varepsilon_{i2}\varepsilon_{iT}\\ \vdots & \vdots & \ddots & \vdots\\ \varepsilon_{iT}\varepsilon_{i1} & \varepsilon_{iT}\varepsilon_{i2} & \cdots & \varepsilon_{iT}^2\end{pmatrix}.$$

また，回帰モデルの係数推定量$\hat{\boldsymbol{\beta}}$の分散推定量は，

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}]=\frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X}}{NT}\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}, \tag{7}$$

と表わされる。

One-way CRの手法による，誤差項のクラスタリングに対して頑健性のある漸近分散推定量を求める方法は，その発想は，White（1980）やNewey and West（1987）が提案した，誤差項の不均一分散や系列相関に対して頑健な漸近分散推定量を求める方法と同じである。

今，$\hat{\boldsymbol{\beta}}$が$\boldsymbol{\beta}$の一致推定量であれば，Slutskyの定理を用いて，

$$\begin{aligned}\operatorname{plim}\hat{\varepsilon}_{it}&=\operatorname{plim}(y_{it}-\mathbf{x}_{it}'\hat{\boldsymbol{\beta}})\\ &=y_{it}-\mathbf{x}_{it}'(\operatorname{plim}\hat{\boldsymbol{\beta}})\\ &=y_{it}-\mathbf{x}_{it}'\boldsymbol{\beta}\\ &=\varepsilon_{it},\end{aligned}$$

となるので，残差$\hat{\varepsilon}_{it}$は，誤差項ε_{it}の一致推定量であるといえる。

そこで，(7)の$\boldsymbol{\Sigma}$の一致推定量として，誤差項ε_{it}の一致推定量である残差$\hat{\varepsilon}_{it}$を用いる，

$$\hat{\boldsymbol{\Sigma}}=\begin{pmatrix}\hat{\boldsymbol{\varepsilon}}_1\hat{\boldsymbol{\varepsilon}}_1' & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \hat{\boldsymbol{\varepsilon}}_2\hat{\boldsymbol{\varepsilon}}_2' & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \hat{\boldsymbol{\varepsilon}}_N\hat{\boldsymbol{\varepsilon}}_N'\end{pmatrix},\text{ただし，}\ \hat{\boldsymbol{\varepsilon}}_i\hat{\boldsymbol{\varepsilon}}_i'=\begin{pmatrix}\hat{\varepsilon}_{i1}^2 & \hat{\varepsilon}_{i1}\hat{\varepsilon}_{i2} & \cdots & \hat{\varepsilon}_{i1}\hat{\varepsilon}_{iT}\\ \hat{\varepsilon}_{i2}\hat{\varepsilon}_{i1} & \hat{\varepsilon}_{i2}^2 & \cdots & \hat{\varepsilon}_{i2}\hat{\varepsilon}_{iT}\\ \vdots & \vdots & \ddots & \vdots\\ \hat{\varepsilon}_{iT}\hat{\varepsilon}_{i1} & \hat{\varepsilon}_{iT}\hat{\varepsilon}_{i2} & \cdots & \hat{\varepsilon}_{iT}^2\end{pmatrix}.$$

そして，上記の$\hat{\boldsymbol{\Sigma}}$を用いることによって，$\hat{\boldsymbol{\beta}}$の漸近分散推定量である，

$$\begin{aligned}\mathrm{Asy.Var}[\hat{\boldsymbol{\beta}}]&=\frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{\mathbf{X}'\hat{\boldsymbol{\Sigma}}\mathbf{X}}{NT}\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\\ &=\frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\hat{\boldsymbol{\varepsilon}}_i\hat{\boldsymbol{\varepsilon}}_i'\mathbf{X}_i\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1},\end{aligned} \tag{8}$$

ただし，$\mathbf{X}_i' = (\mathbf{x}_{i1} \quad \mathbf{x}_{i2} \quad \cdots \quad \mathbf{x}_{iT})$，$\hat{\boldsymbol{\varepsilon}}_i' = (\hat{\varepsilon}_{i1} \quad \hat{\varepsilon}_{i2} \quad \cdots \quad \hat{\varepsilon}_{iT})$，
が求められる。

この漸近分散推定量に基づいて，誤差項の企業に関するクラスタリングに対して頑健な，One-way CRの標準誤差が得られるのである。

最後に，One-way CRの手法の前提となっている $\hat{\boldsymbol{\beta}}$ の一致性について述べる。$\hat{\boldsymbol{\beta}}$ の一致性については，(7)の右辺の積のオーダーについて考える。$(1/NT)$ のオーダーは$O(N^{-1})$であり，$(\mathbf{X}'\mathbf{X}/NT)^{-1}$のオーダーは通常$O(1)$であると想定されているので，問題となるのは$(\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X}/NT)$のオーダーだけである。このとき，$\boldsymbol{\Sigma}$の全要素の個数は$(NT)^2$であるので，$T$を固定して$N$が非常に大きくなると考えると，$\boldsymbol{\Sigma}$のゼロでない要素の個数が全要素の個数に占める割合は小さくなり，その極限は，

$$\lim_{\substack{T\ fixed \\ N\to\infty}} \frac{NT^2}{(NT)^2} = \lim_{\substack{T\ fixed \\ N\to\infty}} \frac{1}{N} = 0,$$

となる。従って，Tが一定でNが十分に大きい場合には，$(\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X}/NT)$のオーダーは，分子が$O(N)$で分母が$O(N^{-1})$なので，$O(1)$で有限となる。よって，$\hat{\boldsymbol{\beta}}$ は $\boldsymbol{\beta}$ の一致推定量であるといえる。

(2) Two-way Cluster-robustの手法

Two-way CRの手法は，前節のOne-way CRの手法の応用版で，誤差項が企業と年度の両方のディメンションに関してクラスタリングしている場合に用いられる手法である（Cameron and Miller 2010; Cameron *et al.* 2011）。

図１は，企業３社（firm = *a, b, c*），年度３年（t = 1, 2, 3）のパネル・データを用いた場合の誤差項のクラスタリングの構造を，共分散行列$\boldsymbol{\Sigma}$を用いて表したものである。なお，下添字の数字は年度を表している。例えば，a_1は企業aの年度１に関する誤差項，c_3は企業cの年度３に関する誤差項を意味している。相関があると仮定されているセルは灰色に塗られており，無相関と仮定されているセルは白色のままである。

図１(a)は，誤差項が企業に関してOne-wayクラスタリングしている場合の誤差項の共分散行列の構造$\boldsymbol{\Sigma}_{\text{firm}}$,(b)は，誤差項が年度に関してOne-wayクラスタリングしている場合の誤差項の共分散行列の構造 $\boldsymbol{\Sigma}_{\text{year}}$ を表している。$\boldsymbol{\Sigma}_{\text{firm}}$,では同一企業の誤差項間に，$\boldsymbol{\Sigma}_{\text{year}}$ では同一年度の誤差項間にクラスタリングが生じていると仮定されていることが見て取れる。

一方，図１(c)は，誤差項が企業と年度の両方に関してTwo-wayクラスタリングしている場合の共分散構造$\boldsymbol{\Sigma}_{\text{Two-way}}$ を示している。$\boldsymbol{\Sigma}_{\text{Two-way}}$ は，基本的に,$\boldsymbol{\Sigma}_{\text{firm}}$ と $\boldsymbol{\Sigma}_{\text{year}}$ を足し合わせることによって容易に構築されるのだが，その場合に，共分散行列の対角要素が重なり合ってしまう((c)の濃い灰色の部分)。そこで，この対角要素の重複部分を１回差し引いてやるのだが，この重複部分は，誤差項に不均一分散があるとする場合の共分散構造 $\boldsymbol{\Sigma}_{\text{hetro}}$ と同じである。

図1　Cluster-robustの手法における誤差項の共分散構造

(a)　企業に関するOne-way クラスタリング $\mathbf{\Sigma}_{\mathrm{firm}}$

		firm *a*			firm *b*			firm *c*		
		t=1	t=2	t=3	t=1	t=2	t=3	t=1	t=2	t=3
	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3	a_1c_1	a_1c_2	a_1c_3
firm *a*	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3	a_2c_1	a_2c_2	a_2c_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3	a_3c_1	a_3c_2	a_3c_3
	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3	b_1c_1	b_1c_2	b_1c_3
firm *b*	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3	b_2c_1	b_2c_2	b_2c_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2	b_3c_1	b_3c_2	b_3c_3
	t=1	c_1a_1	c_1a_2	c_1a_3	c_1b_1	c_1b_2	c_1b_3	c_1^2	c_1c_2	c_1c_3
firm *c*	t=2	c_2a_1	c_2a_2	c_2a_3	c_2b_1	c_2b_2	c_2b_3	c_2c_1	c_2^2	c_2c_3
	t=3	c_3a_1	c_3a_2	c_3a_3	c_3b_1	c_3b_2	c_3b_3	c_3c_1	c_3c_2	c_3^2

(b)　年度に関するOne-way クラスタリング $\mathbf{\Sigma}_{\mathrm{year}}$

		firm *a*			firm *b*			firm *c*		
		t=1	t=2	t=3	t=1	t=2	t=3	t=1	t=2	t=3
	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3	a_1c_1	a_1c_2	a_1c_3
firm *a*	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3	a_2c_1	a_2c_2	a_2c_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3	a_3c_1	a_3c_2	a_3c_3
	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3	b_1c_1	b_1c_2	b_1c_3
firm *b*	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3	b_2c_1	b_2c_2	b_2c_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2	b_3c_1	b_3c_2	b_3c_3
	t=1	c_1a_1	c_1a_2	c_1a_3	c_1b_1	c_1b_2	c_1b_3	c_1^2	c_1c_2	c_1c_3
firm *c*	t=2	c_2a_1	c_2a_2	c_2a_3	c_2b_1	c_2b_2	c_2b_3	c_2c_1	c_2^2	c_2c_3
	t=3	c_3a_1	c_3a_2	c_3a_3	c_3b_1	c_3b_2	c_3b_3	c_3c_1	c_3c_2	c_3^2

(c)　企業と年度に関するTwo-way クラスタリング $\mathbf{\Sigma}_{\mathrm{Two\text{-}way}}$

		firm *a*			firm *b*			firm *c*		
		t=1	t=2	t=3	t=1	t=2	t=3	t=1	t=2	t=3
	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3	a_1c_1	a_1c_2	a_1c_3
firm *a*	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3	a_2c_1	a_2c_2	a_2c_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3	a_3c_1	a_3c_2	a_3c_3
	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3	b_1c_1	b_1c_2	b_1c_3
firm *b*	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3	b_2c_1	b_2c_2	b_2c_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2	b_3c_1	b_3c_2	b_3c_3
	t=1	c_1a_1	c_1a_2	c_1a_3	c_1b_1	c_1b_2	c_1b_3	c_1^2	c_1c_2	c_1c_3
firm *c*	t=2	c_2a_1	c_2a_2	c_2a_3	c_2b_1	c_2b_2	c_2b_3	c_2c_1	c_2^2	c_2c_3
	t=3	c_3a_1	c_3a_2	c_3a_3	c_3b_1	c_3b_2	c_3b_3	c_3c_1	c_3c_2	c_3^2

（注）灰色のエリアだけ相関があると仮定しており，白色のエリアは無相関と仮定している。なお，(c)の対角要素の濃い灰色のエリアは，(a)と(b)を足し合わせた場合の重なる領域であり，誤差項に不均一分散がある場合の共分散行列$\mathbf{\Sigma}_{\mathrm{hetro}}$と同じである。従って，$\mathbf{\Sigma}_{\mathrm{Two\text{-}way}} = \mathbf{\Sigma}_{\mathrm{firm}} + \mathbf{\Sigma}_{\mathrm{year}} - \mathbf{\Sigma}_{\mathrm{hetro}}$という関係が存在している。

つまり，誤差項が企業と年度の両方に関してクラスタリングがある場合の，誤差項の共分散行列は，

$$\boldsymbol{\Sigma}_{\text{Two-way}} = \boldsymbol{\Sigma}_{\text{firm}} + \boldsymbol{\Sigma}_{\text{year}} - \boldsymbol{\Sigma}_{\text{hetro}}$$

で表現されるのである。

従って，Two-way CRの$\boldsymbol{\Sigma}$の一致推定量は，One-wayの場合を応用して，誤差項ε_{it}の一致推定量である残差$\hat{\varepsilon}_{it}$を用いて，以下のように表せる。

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\varepsilon}}_1\hat{\boldsymbol{\varepsilon}}_1' & \hat{\boldsymbol{\varepsilon}}_1\hat{\boldsymbol{\varepsilon}}_2' & \dots & \hat{\boldsymbol{\varepsilon}}_1\hat{\boldsymbol{\varepsilon}}_N' \\ \hat{\boldsymbol{\varepsilon}}_2\hat{\boldsymbol{\varepsilon}}_1' & \hat{\boldsymbol{\varepsilon}}_2\hat{\boldsymbol{\varepsilon}}_2' & \cdots & \hat{\boldsymbol{\varepsilon}}_2\hat{\boldsymbol{\varepsilon}}_N' \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\boldsymbol{\varepsilon}}_N\hat{\boldsymbol{\varepsilon}}_1' & \hat{\boldsymbol{\varepsilon}}_N\hat{\boldsymbol{\varepsilon}}_2' & \cdots & \hat{\boldsymbol{\varepsilon}}_N\hat{\boldsymbol{\varepsilon}}_N' \end{pmatrix}$$

ただし，$\hat{\boldsymbol{\varepsilon}}_i\hat{\boldsymbol{\varepsilon}}_i' = \begin{pmatrix} \hat{\varepsilon}_{i1}^2 & \hat{\varepsilon}_{i1}\hat{\varepsilon}_{i2} & \dots & \hat{\varepsilon}_{i1}\hat{\varepsilon}_{iT} \\ \hat{\varepsilon}_{i2}\hat{\varepsilon}_{i1} & \hat{\varepsilon}_{i2}^2 & \cdots & \hat{\varepsilon}_{i2}\hat{\varepsilon}_{iT} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\varepsilon}_{iT}\hat{\varepsilon}_{i1} & \hat{\varepsilon}_{iT}\hat{\varepsilon}_{i2} & \cdots & \hat{\varepsilon}_{iT}^2 \end{pmatrix}$ and $\hat{\boldsymbol{\varepsilon}}_i\hat{\boldsymbol{\varepsilon}}_j' = \begin{pmatrix} \hat{\varepsilon}_{i1}\hat{\varepsilon}_{j1} & 0 & \dots & 0 \\ 0 & \hat{\varepsilon}_{i2}\hat{\varepsilon}_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\varepsilon}_{iT}\hat{\varepsilon}_{jT} \end{pmatrix}$

そして，上記の$\hat{\boldsymbol{\Sigma}}$を用いることによって，$\hat{\boldsymbol{\beta}}$の漸近分散推定量である，

$$\begin{aligned} \text{Asy.Var}[\hat{\boldsymbol{\beta}}] &= \frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{\mathbf{X}'\hat{\boldsymbol{\Sigma}}\mathbf{X}}{NT}\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1} \\ &= \frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\hat{\boldsymbol{\varepsilon}}_i\hat{\boldsymbol{\varepsilon}}_i'\mathbf{X}_i\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1} \\ &\quad + \frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{X}_t'\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'\mathbf{X}_t\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1} \\ &\quad - \frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2\mathbf{x}_{it}\mathbf{x}_{it}'\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}, \end{aligned} \tag{9}$$

ただし，

$\mathbf{X}_i' = (\mathbf{x}_{i1}\ \mathbf{x}_{i2} \cdots \mathbf{x}_{iT})$, $\mathbf{X}_t' = (\mathbf{x}_{1t}\ \mathbf{x}_{2t} \cdots \mathbf{x}_{Nt})$, $\hat{\boldsymbol{\varepsilon}}_i' = (\hat{\varepsilon}_{i1}\ \hat{\varepsilon}_{i2} \cdots \hat{\varepsilon}_{iT})$, $\hat{\boldsymbol{\varepsilon}}_t' = (\hat{\varepsilon}_{1t}\ \hat{\varepsilon}_{2t} \cdots \hat{\varepsilon}_{Nt})$,

が求められる。

この漸近分散推定量に基づいて，誤差項の企業と年度の両方のクラスタリングに対して頑健な，Two-way CRの標準誤差が得られるのである。

最後に，Two-way CRの手法における$\hat{\boldsymbol{\beta}}$の一致性について述べる。One-way CRの場合と同様に，問題となるのは$(\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X}/NT)$のオーダーである。$\boldsymbol{\Sigma}$の全要素の個数は$(NT)^2$であるので，$N$と$T$の両方が非常に大きくなると考えると，$\boldsymbol{\Sigma}$のゼロでない要素の個数が全要素の個数に占める割合は小さくなり，その極限は，

$$\lim_{\substack{N\to\infty \\ T\to\infty}} \frac{NT^2 + N^2T - NT}{(NT)^2} = \lim_{\substack{N\to\infty \\ T\to\infty}}\left(\frac{1}{N} + \frac{1}{T} - \frac{1}{NT}\right) = 0,$$

となる。従って，N と T の両方が十分に大きい場合には，$(\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X} / NT)$のオーダーは$O(1)$で有限となるので，$\hat{\boldsymbol{\beta}}$ は $\boldsymbol{\beta}$ の一致推定量であるといえる。

(3) Cluster-robustの手法を用いる際の問題点

(a) クラスター数が少ない場合の調整方法

最初に，CRの手法を用いる際に問題となるのが，クラスター数が少ない場合のCRの標準誤差には，真の標準誤差を過小に推定してしまう過小推定バイアスが存在するということである（Cameron *et al.* 2008; Cameron and Miller 2010）。つまり，クラスター数が少ないケースでCRの手法を用いると，本来なら棄却されない係数推定値を，誤って棄却してしまう可能性があるのである。

それでは，実際どのくらいのクラスター数があれば良いのかということであるが，この点に関しては，研究者の間で意見にバラつきがみられる。というのも，この適切なクラスター数は，全てモンテカルロ・シミュレーションの結果に基づいているので，シミュレーションモデルの仮定や，許容されるバイアスをどの程度までと定めるかによって結論が変わってくるからである。

例えば，Thompson（2011）は，Two-way CRのシミュレーションで，真の棄却率が５%のときに10%の棄却率までを許容範囲とすると，企業と年度の数が共に25個以上あれば許容範囲内に収まるという結果を報告している。一方，Petersen（2009）は，One-way CRのシミュレーションで，年度数が５個，10個，20個，40個の場合に，真の標準誤差を，それぞれ約27%，16%，８%，１%過小推定するという結果を示している。また，Two-way CRのシミュレーションでは，真の棄却率が１%のときに，企業と年度の数が共に100個の場合には１%の棄却率で正しく推定されるが，年度数が10個に減ると棄却率が５%にまで上昇するという結果を得ている。さらに，Gow *et al.*（2010）では，Two-way CRのシミュレーションで，真の棄却率が１%のときに，年度数が10個の場合には棄却率が3.3%に達するという結果を報告している。しかしながら，彼らは，それでも他の手法による棄却率よりは低いので，クラスター数が少ないからといってTwo-way CRを用いない方が良いということにはならないと述べている。

以上の先行研究の結果からは，クラスター数がおおよそ20～30個あれば，一般に問題なくCRの手法が利用可能であるといえる。しかしながら，20～30個以下である場合には，何らかの調整が必要であると思われる。

現在，クラスター数が少ない場合の調整方法として，幾つかの手法が提案されているが，その多くは，One-way CRに関するものである（Cameron *et al.* 2008; Angrist and Pischke 2009; Cameron and Miller 2010）。そこで以下では，実施が容易で，かつOne-wayとTwo-wayの両方のCR手法に利用可能な調整方法について説明する（Cameron *et al.* 2011）。

今仮に，企業数$(i = 1, 2, \cdots, N)$に比して年度数$(t = 1, 2, \cdots, T)$が少ないショート・パネル・データを使って，年度に関してOne-way CRの手法を用いたとする。なお，説明変数の数は k

個とする。このとき，$\hat{\boldsymbol{\beta}}$ の通常の漸近分散推定量は，(8)と同様に，以下のように表わされる。

$$\text{Asy.Var}[\hat{\boldsymbol{\beta}}]=\frac{1}{NT}\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1}\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{X}_t'\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'\mathbf{X}_t\right)\left(\frac{\mathbf{X}'\mathbf{X}}{NT}\right)^{-1},$$

ただし，$\mathbf{X}_t'=(\mathbf{x}_{1t}\quad \mathbf{x}_{2t}\quad \cdots\quad \mathbf{x}_{Nt})$，$\hat{\boldsymbol{\varepsilon}}_t'=(\hat{\varepsilon}_{1t}\quad \hat{\varepsilon}_{2t}\quad \cdots\quad \hat{\varepsilon}_{Nt})$.

このとき，Tが小さい場合には，残差ベクトル$\hat{\boldsymbol{\varepsilon}}_t$の代わりに，

$$\sqrt{c}\hat{\boldsymbol{\varepsilon}}_t,\quad \text{ただし,}\quad c=\frac{T}{T-1}\frac{NT-1}{NT-k},$$

を用いて，残差を大きく修正するのである。

つまり，Nが十分に大きいときには，修正された標準誤差は，修正前の標準誤差よりも$\sqrt{T/(T-1)}$ 倍大きくなっているのである。さらに，検定統計量の有意水準を求める際には，標準正規分布を用いるのではなく，自由度$T-1$の t 分布を用いるのである。

Two-way CRの手法を用いた場合の調整方法は，One-wayの場合の応用版であり，(9)の右辺第１項の残差ベクトル $\hat{\boldsymbol{\varepsilon}}_i$，第２項の残差ベクトル $\hat{\boldsymbol{\varepsilon}}_t$，第３項の残差 $\hat{\varepsilon}_{it}$の代わりに，それぞれ，

$$\sqrt{c_1}\hat{\boldsymbol{\varepsilon}}_i,\quad \sqrt{c_2}\hat{\boldsymbol{\varepsilon}}_t,\quad \sqrt{c_3}\hat{\varepsilon}_{it},$$

ただし，$c_1=\dfrac{N}{N-1}\dfrac{NT-1}{NT-k},\quad c_2=\dfrac{T}{T-1}\dfrac{NT-1}{NT-k},\quad c_3=\dfrac{NT}{NT-1}\dfrac{NT-1}{NT-k},$

を用いる。そして，検定統計量の有意水準を求める際には，自由度$\min(N,\ T)-1$ の t 分布を用いるのである。

(b) Two-way Cluster-robustの漸近分散推定量の正定値性

次に注意しなければならないのが，CRの手法から得られる漸近分散推定量の正定値性に関する問題である。One-way CRおよびTwo-way CRの手法による漸近分散推定量は，それぞれ，(8)と(9)で表されるが，これらが正定値行列でない場合には，個々の係数の分散が負になったり，複数の係数の線形制約に関する検定統計量が負になってしまう可能性がある。そして，(8)および(9)が正定値行列となるかどうかは，誤差項の共分散行列の一致推定量である$\hat{\boldsymbol{\Sigma}}$が正定値行列であるかどうかによって決定される。

そして，One-way CRの$\hat{\boldsymbol{\Sigma}}$は正定値行列となるので，One-way CRの手法から得られる漸近分散推定量の正定値性は保証されている[3)]。一方，Two-way CRの手法から得られる漸近分散

3）One-way CR の$\hat{\boldsymbol{\Sigma}}$ が正定値行列となることの証明は次の通りである。(8)の$\hat{\boldsymbol{\Sigma}}$ は，

$$\hat{\boldsymbol{\Sigma}}=\mathbf{a}_1\mathbf{a}_1'+\mathbf{a}_2\mathbf{a}_2'+\cdots+\mathbf{a}_N\mathbf{a}_N',$$

ただし，$\mathbf{a}_1'=(\hat{\boldsymbol{\varepsilon}}_1'\quad \mathbf{0}'\quad \cdots\quad \mathbf{0}')$，$\mathbf{a}_2'=(\mathbf{0}'\quad \hat{\boldsymbol{\varepsilon}}_2'\quad \mathbf{0}'\quad \cdots\quad \mathbf{0}')$，$\cdots$ ，$\mathbf{a}_N'=(\mathbf{0}'\quad \cdots\quad \mathbf{0}'\quad \hat{\boldsymbol{\varepsilon}}_N')$,と表せる。

$\mathbf{a}\mathbf{a}'$ は，任意のベクトル $\mathbf{x}$ に対して，$\mathbf{x}'(\mathbf{a}\mathbf{a}')\mathbf{x}=(\mathbf{x}'\mathbf{a})(\mathbf{x}'\mathbf{a})'=\mathbf{y}\mathbf{y}'>0$となるので正定値行列である。従って，$\hat{\boldsymbol{\Sigma}}$ は正定値行列であるといえる。

推定量については，One-wayの場合とは異なって，正定値性が保証されていない。そこで，その対処法として，Two-way CRの漸近分散推定量をスペクトラル分解して，

$$\mathrm{Asy.Var}[\hat{\boldsymbol{\beta}}] = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}',\quad \text{ただし，}\mathbf{C} = (\mathbf{c}_1 \quad \mathbf{c}_2 \quad \cdots \quad \mathbf{c}_K),\quad \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_K \end{pmatrix},$$

を得る。そして，

$$\boldsymbol{\Lambda}^+ = \mathrm{Diag}\left[\lambda_1^+ \quad \lambda_2^+ \quad \cdots \quad \lambda_K^+\right],\ \text{ただし，}\lambda_k^+ = \max(0, \lambda_k),$$

というような$\boldsymbol{\Lambda}^+$を作成して，

$$\mathrm{Asy.Var}[\hat{\boldsymbol{\beta}}]^+ = \mathbf{C}\boldsymbol{\Lambda}^+\mathbf{C}',$$

を代わりに用いるという方法が提案されている（Cameron *et al.* 2011）。つまり負の固有値はゼロとする新たな対角行列$\boldsymbol{\Lambda}^+$を作成して，それを用いてTwo-way CRの漸近分散推定量とするのである。

ただし，この負の固有値をゼロと置き換えることによって正定値性を担保する方法は，非常にアドホックであり，根本的な解決策とはなっていないということには注意が必要である。

(c) ショート・パネル・データにCluster-robustの手法を用いた場合の影響

最後に，企業数（$i = 1, 2, \cdots, N$）に比して年度数（$t = 1, 2, \cdots, T$）が非常に少ないショート・パネル・データを用いた場合の，One-way CRとTwo-way CRの分散推定量の一般的な関係について述べる。

企業に関するOne-way CR，年度に関するOne-way CR，Two-way CRの手法から得られる分散推定量およびWhite（1980）の不均一分散に対して頑健性のある分散推定量を，それぞれ，$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (firm)}}]$，$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}]$，$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}]$および$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{White}}]$と表すとすると，Two-way CRの分散推定量は，(9)にあるように，

$$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}] = \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (firm)}}] + \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}] - \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{White}}], \tag{10}$$

と表現することができる。このとき，Tを固定して，Nを大きくすると，

$$\lim_{\substack{T\ fixed \\ N\to\infty}} \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}] = \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}],$$

となる。つまり，企業数が大きく年度数が小さいショート・パネルを用いた場合のTwo-way CRの標準誤差は，年度に関するOne-way CRの標準誤差と殆ど変わらなくなるのである（Thompson 2011）。

また別の表現をすれば，(10)の右辺の3つの項は，全て一致推定量である$\hat{\boldsymbol{\beta}}$の分散推定量であるので，第1項の$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (firm)}}]$は企業数が大きくなるとゼロに近付き，第2項の$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}]$は年度数が大きくなるゼロに近付き，第3項の$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{White}}]$は全体の観測値

数が大きくなると（すなわち企業数と年度数のどちらが大きくなっても）ゼロに近付く。従って，企業数が大きく年度数が小さいショート・パネルでは，第１項の $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (firm)}}]$ と第３項の $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{White}}]$ が非常に小さくなるので，結果として，$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}]$ は第２項の $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}]$ とほぼ一致するのである。

このことは，２つのディメンションから成るパネル・データでは，クラスター数の少ない方でCRの手法を用いることが重要であるということを意味しており，会計・ファイナンスの実証研究でしばしば遭遇する，企業数が大きく年度数が小さいショート・パネル・データを用いた分析では，年度に関してCRの手法を用いることが重要であるということを示唆している。また，ショート・パネルでは，$\mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}] \simeq \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}]$となるので，前項で述べた，Two-way CRの分散推定値が負になってしまうような場合には，Two-wayではなく年度に関するOne-way CRを用いることが，ひとつの解決策であるといえる。

４．Fama-MacBethの手法とCluster-robustの手法の特徴の比較

CRの手法に基づく分散推定量は，基本的に，White（1980）やNewey and West（1987）が提示したHC推定量（heteroscedasticity consistent estimator）やHAC推定量（heteroscedasticity and autocorrelation consistent estimator）と同じ発想で導かれている。すなわち，$\hat{\boldsymbol{\beta}}$が$\boldsymbol{\beta}$の一致推定量であれば残差は誤差項の一致推定量となるということを利用して，誤差項の共分散行列の一致推定量を求めているのである。

一方，FMの手法は，上記の推定量とはその発想が全く異なっており，誤差項の共分散行列を推定することなしに分散推定量を求めようとするものであり，CRの手法に対する一種の簡便法であるといえる。

両手法の使われ方としては，パネル分析における誤差項のクロスセクショナルな相関に対処するために，通常のFMの手法と年度に関するOne-way CRの手法が用いられ，誤差項のクロスセクションおよび時系列の相関に対処するために，調整済みFMの手法とTwo-way CRの手法が用いられる。ただし，第２章で論じたように，FMの手法は様々な問題を抱えており，通常のFMの手法は年度に亘って説明変数にそれほど大きな変化がない場合にのみ有効であり，調整済みFMの手法は明確な理論的根拠を欠いているという点には注意が必要である。

また，会計・ファイナンスの実証研究では，企業数が大きく年度数が小さいショート・パネル・データが用いられるケースが非常に多く，そのような場合には，前章で示したように，

$$\lim_{\substack{T\ fixed \\ N\to\infty}} \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}] = \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}], \tag{11}$$

となる。さらに，ショート・パネルで説明変数の変動が年度間でそれ程大きくない場合には，

$$\lim_{\substack{T\ fixed \\ N\to\infty}} \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{Two-way CR}}] = \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{One-way CR (year)}}] \approx \mathrm{Var}[\hat{\boldsymbol{\beta}}_{\text{FM}}], \tag{12}$$

表１　Fama-MacBethの手法とCluster-robustの手法の特徴の比較

	FMの手法	CRの手法
手法の特徴	誤差項の共分散行列を推定することなしに分散推定量を求めており，CRの手法に対する一種の簡便法である。	HC推定量やHAC推定量と同じ発想で，残差が誤差項の一致推定量となるということを利用して，誤差項の共分散行列の推定量を求めている。
誤差項のクロスセクショナルな相関に対応	通常のFMの手法（年度に亘って説明変数に大きな変化がない場合に有効）	年度に関するOne-way CRの手法
誤差項の時系列の相関に対応	N/A	企業に関するOne-way CRの手法
誤差項のクロスセクションおよび時系列の相関に対応	調整済みFMの手法（明確な理論的根拠がない）	Two-way CRの手法
企業数に比して年度数が小さいショート・パネル・データを使用	通常のFMの手法は，Two-way CRや年度に関するOne-way CRの手法と類似したものとなる。	Two-way CRと年度に関するOne-way CRの手法は一致する。

となるので，FMの手法，Two-way CRの手法，年度に関するOne-way CRの手法の何れの手法を用いても，同じような結果が得られるといえる。

最後に，本章で述べたFMの手法とCRの手法の特徴の比較を，表１でまとめている。

５．実証研究への応用

前章までの分析では，ショート・パネル・データを用いた場合には(11)が成立し，さらに，説明変数の変動が年度間でそれ程大きくない場合には(12)が成り立つという理論的帰結を得ている。そこで，本章では，年度間の変動が相対的に小さいと思われる例としては会計利益，大きいと思われる例としては株式リターンをそれぞれ説明変数とする回帰モデルを，ショート・パネル・データを使って実際に推定し，(11)および(12)が成立しているかを検証している。

具体的には，第１節では会計利益を説明変数とする包括利益情報の有用性に関する研究，第２節では株式リターンを説明変数とする保守主義に関する一連の研究で使用されているモデルを推定している。そして，One-way CR (firm)，One-way CR (year)，Two-way CRおよびFMの手法から得られる各種の標準誤差を用いることによって，係数推定値の統計的検定にどのような影響が生じるかを調査している。

(1) 会計利益を説明変数とする研究

わが国における包括利益情報の有用性を調査する研究としては，久保田・須田・竹原(2006)，井出(2006)，若林(2009；2010)がある。何れの論文も，企業と年度のショート・パネル・データを用いた分析を行っており，被説明変数としては年次リターン，説明変数としては，純利益，包括利益，その他の包括利益の構成要素等を用いている。そこで，本稿では，Chambers *et al.* (2007)が最初に提案し，若林(2009；2010)で用いられている以下のモデルを

用いて，各種の手法が係数の統計的検定に与える影響を調査している。

$$R_{it} = \alpha_0 + \alpha_1 NI_{it} + \alpha_2 DNEG*NI_{it} + \alpha_3 DYEAR_{it} + \varepsilon_{it}, \tag{13}$$

$$R_{it} = \alpha_0 + \alpha_1 CI_{it} + \alpha_2 DNEG*CI_{it} + \alpha_3 DYEAR_{it} + \varepsilon_{it}, \tag{14}$$

$$R_{it} = \alpha_0 + \alpha_1 NI_{it} + \alpha_2 DNEG*NI_{it} + \alpha_3 OCI_{it} + \alpha_4 DYEAR_{it} + \varepsilon_{it}, \tag{15}$$

ただし，

R_{it}: 期首３ヵ月後から期末３ヵ月後までの年次リターン，

NI_{it}: 当期純利益（少数株主損益控除前），

CI_{it}: 包括利益（少数株主損益控除前），

OCI_{it}: その他の包括利益，

$DNEG$: 利益が負の場合は１，それ以外はゼロのダミー変数，

$DYEAR$: 各年の年度ダミー変数.

なお，下添字 i と t は，それぞれ企業と年度を表しており，NI，CI，OCIは，期首の時価総額でデフレートされている。

サンプルの選択は，以下の基準で行っている。

(i) NEEDS FinancialQUESTで，2002-2011年の期間において，日本基準に準拠した連結財務諸表および株価に関するデータが入手可能である，

(ii) ３月決算の一般事業会社（銀行，証券，保険，その他金融業を除く）で，会計期間が12ヵ月である，

(iii) その他有価証券評価差額金，為替換算調整勘定，土地再評価差額金，繰延ヘッジ損益，年金債務調整額の何れかが計上されている。

これらの基準によって，10年間で延べ14,440企業年の観測値を得ている。さらに，異常値の影響を考慮して，各変数の上下１％を除去している。これによって，最終サンプルの観測値数は，13,656個となっている。

最初に，説明変数の年度間の変動を厳密に測定することは不可能ではあるが，その一つの尺度として，説明変数の各年平均値10個の平均とその標準偏差を，表２ Panel Aで示している。説明変数は全て時価総額デフレートされているので，例えばNIは，各年平均値の平均が時価総額の3.54%で，各年平均値の標準偏差は時価総額の2.68%であるということを意味している。全体的に，NI，CI，OCIといった各種利益の各年平均値の標準偏差は，後掲の表３ Panel Aの株式リターンの標準偏差と比べるとかなり小さくなっており，会計利益の各年平均値のバラつきが，株式リターンよりも相対的に小さいことが伺える。

次に，表２ Panel Bは，(13)(14)(15)の推定結果である。検定統計量としては，通常のOLS，One-way CR (firm)，One-way CR (year)，Two-way CR，そしてFMの合計５種類の t 値を載せている。なおFMの係数は，2002-2011年の各年で推定した係数推定値の平均値である.

Panel Bからは，大きく以下の三つのことが観察される。第一に，One-way CR (year)と

表2　各種の手法を用いた場合の包括利益モデルの推定結果

Panel A: 説明変数の各年の変動

説明変数	*NI*	*CI*	*OCI*	*DNEG*NI*	*DNEG*CI*
各年平均値の平均	0.0354	0.0347	−0.0006	−0.0236	−0.0150
各年平均値の標準偏差	0.0268	0.0477	0.0288	0.0158	0.0147

Panel B: 推定結果

説明変数	*NI*	*DNEG*NI*	*CI*	*DNEG*CI*	*OCI*
(13) 係数	1.594	−1.650			
通常のOLSの t 値	(36.14)**	(−27.68)**			
One-way CR (firm)の t 値	(28.02)**	(−21.78)**			
One-way CR (year)の t 値	(7.05)**	(−6.06)**			
Two-way CRの t 値	(7.04)**	(−6.06)**			
(13) FMの係数	1.637	−1.674			
FMの t 値	(7.07)**	(−6.51)**			
(14) 係数			1.287	−1.286	
通常のOLSの t 値			(32.76)**	(−23.70)**	
One-way CR (firm)の t 値			(26.13)**	(−19.36)**	
One-way CR (year)の t 値			(7.49)**	(−5.76)**	
Two-way CRの t 値			(7.49)**	(−5.76)**	
(14) FMの係数			1.352	−1.346	
FMの t 値			(7.39)**	(−6.17)**	
(15) 係数	1.593	−1.647			0.165
通常のOLSの t 値	(36.12)**	(−27.64)**			(2.76)**
One-way CR (firm)の t 値	(28.01)**	(−21.75)**			(2.35)*
One-way CR (year)の t 値	(7.01)**	(−6.00)**			(1.41)
Two-way CRの t 値	(7.00)**	(−5.99)**			(1.38)
(15) FMの係数	1.616	−1.653			0.304
FMの t 値	(7.03)**	(−6.41)**			(1.47)

（注）推定モデルは以下のようである。

$$R_{it} = \alpha_0 + \alpha_1 NI_{it} + \alpha_2 DNEG{*}NI_{it} + \alpha_3 DYEAR_{it} + \varepsilon_{it}, \quad (13)$$

$$R_{it} = \alpha_0 + \alpha_1 CI_{it} + \alpha_2 DNEG{*}CI_{it} + \alpha_3 DYEAR_{it} + \varepsilon_{it}, \quad (14)$$

$$R_{it} = \alpha_0 + \alpha_1 NI_{it} + \alpha_2 DNEG{*}NI_{it} + \alpha_3 OCI_{it} + \alpha_4 DYEAR_{it} + \varepsilon_{it}, \quad (15)$$

ただし，R_{it}：期首3ヵ月後から期末3ヵ月後までの年次リターン，NI_{it}：当期純利益（少数株主損益控除前），CI_{it}：包括利益（少数株主損益控除前），OCI_{it}：その他の包括利益，$DNEG$：利益が負の場合は1，それ以外はゼロのダミー変数，$DYEAR$：各年の年度ダミー変数。なお，下添字 i と t は，それぞれ企業と年度を表しており，NI，CI，OCIは，期首の時価総額でデフレートされている。

*　5%水準で有意　**　1%水準で有意。

Two-way CRによる t 値がほぼ同一であるということである。例えば，(13)のNIに関する t 値は，One-way CR（year）で7.05，Two-way CRで7.04であり，$DNEG{*}NI$の t 値は，両者とも同じ−6.06である。これと同様の傾向は，(14)(15)においても観察される。この結果は，企業数に比して年度数の少ないショート・パネルを用いた場合には，Two-way CRとOne-way CR（year）による標準誤差は等しくなるという(11)の考察を支持するものである。

第二に，FMの t 値と，One-way CR（year）およびTwo-way CRによる t 値に大差がないということである。先と同様に，(13)のNIの t 値を比べると，One-way CR（year），Two-way CR，FMによる t 値は，それぞれ，7.05，7.04，7.07であり，$DNEG{*}NI$の t 値は，それぞれ，−6.06，−6.06，−6.51である。これと同様の傾向は，(14)(15)においても観察される。この結果は，

説明変数の年度間の変動が小さい場合には，FMの標準誤差と，Two-way CRおよびOne-way CR（year）による標準誤差との間には大差はないという(12)の考察を支持するものである。

第三に，One-way CR（firm）の t 値は，他の三つの手法と比較すると非常に大きいということである。例えば，(13)の*NI*と*DNEG*NI*の t 値は，One-way CR（firm）の場合には，それぞれ，28.02と–21.78であり，他の三つの手法の t 値と比べて格段に大きい値となっている。これは，(11)の考察と表裏一体であるが，企業数に比して年度数の少ないショート・パネルを用いた場合には，誤差項の時系列の相関が統計的推論に与える影響は軽微であるということを示唆している。

最後に，表２の推定結果は，誤差項のクロスセクショナルな相関を考慮するOne-way CR（year），Two-way CR，FMの手法を用いると，(15)の*OCI*の係数は統計的に有意とはならず，その他の包括利益情報の有用性に疑義を呈する結果となっている.

(2) 株式リターンを説明変数とする研究

わが国における保守主義に関する研究としては，Ball *et al.*（2000），田澤(2004)，高田(2006)，Shuto and Takada（2010)などがある。これらの論文における推定モデルの基礎となっているのが，Basu（1997)およびPope and Walker（1999)で提案された，利益を被説明変数，リターンを説明変数とする保守主義の定量化モデルである。そして何れの論文も，企業と年度のショート・パネル・データを用いて，保守主義定量化モデルやそのバリエーション・モデルを推定している。そこで，本稿でも，田澤(2004)および高田(2006)で用いられている以下の保守主義モデルを用いて，各種の手法が係数の統計的検定に与える影響を調査している。

$$NI_{it} = \gamma_0 + \gamma_1 DR_{it} + \gamma_2 R_{it} + \gamma_3 DR{*}R_{it} + \varepsilon_{it}, \tag{16}$$

$$NI_{it} = \gamma_0 + \gamma_1 DR_{it} + \gamma_2 R_{it} + \gamma_3 DR{*}R_{it} + \gamma_4 R_{it-1} + \varepsilon_{it}, \tag{17}$$

ただし，

NI_{it}: 当期純利益，

R_{it}: 期首３ヵ月後から期末３ヵ月後までの年次リターン，

DR_{it}: R_{it}が負の場合は１，それ以外はゼロのダミー変数.

なお，下添字 i と t は，それぞれ企業と年度を表しており，*NI*は，期首の時価総額でデフレートされている。

サンプルの選択は，以下の基準で行っている。

(i) NEEDS FinancialQUESTで，1990-2011年の期間において，当期純利益および株価に関するデータが入手可能である。なお，当期純利益に関しては，連結財務諸表の値を用いており，無ければ個別財務諸表で代替している，

(ii) ３月決算の一般事業会社（銀行，証券，保険，その他金融業を除く）で，会計期間が12ヵ月である。

これらの基準によって，22年間で延べ38,190企業年の観測値を得ている。さらに，異常値の影響を考慮して，各変数の上下１％を除去している。これによって，最終サンプルの観測値数は，36,748個となっている。

最初に，前節同様に，説明変数の年度間の変動を表す一つの尺度として，説明変数の各年平均値22個の平均とその標準偏差を，表３ Panel Aで示している。例えば，Rの各年平均値の平均は0.71％で，各年平均値の標準偏差は24.41％である。Rは株式リターンであるので，このことは，時価総額の各年平均変化額の平均が時価総額の0.71％で，各年平均変化額の標準偏差が時価総額の24.41％にも上るということを意味している。全体的に，RやR_{t-1}といった株式リターンの各年平均値の標準偏差は，前掲の表２ Panel Aの会計利益の標準偏差と比べるとかなり大きくなっており，株式リターンの各年平均値のバラつきが，会計利益よりも相対的に大きいことが伺える。

次に，表３ Panel Bは，(16)(17)の推定結果である。検定統計量としては，表２と同様に，通常のOLS，One-way CR（firm），One-way CR（year），Two-way CR，そしてFMの合計５種類の t 値を載せている。なおFMの係数は，1990-2011年の各年で推定した係数推定値の平

表３　各種の手法を用いた場合の保守主義モデルの推定結果

Panel A: 説明変数の各年の変動

説明変数	R	R_{t-1}	DR	DR^*R
各年平均値の平均	0.0071	0.0268	0.5671	−0.1387
各年平均値の標準偏差	0.2441	0.2782	0.3044	0.1126

Panel B: 推定結果

説明変数	*Intercept*	DR	R	DR^*R	R_{t-1}
(16) 係数	0.034	−0.005	0.017	0.075	
通常のOLSの t 値	(29.10)**	(−2.98)**	(6.59)**	(14.44)**	
One-way CR（firm）の t 値	(26.73)**	(−2.96)**	(5.75)**	(12.88)**	
One-way CR（year）の t 値	(5.44)**	(−1.27)	(1.24)	(2.69)*	
Two-way CRの t 値	(5.41)**	(−1.26)	(1.24)	(2.68)*	
(16) FMの係数	0.033	−0.000	0.000	0.138	
FMの t 値	(7.86)**	(−0.12)	(0.06)	(4.95)**	
(17) 係数	0.033	−0.005	0.018	0.081	0.032
通常のOLSの t 値	(27.47)**	(−2.94)**	(6.65)**	(14.86)**	(29.03)**
One-way CR（firm）の t 値	(25.11)**	(−2.90)*	(5.56)**	(12.98)**	(14.10)**
One-way CR（year）の t 値	(5.62)**	(−1.35)	(1.36)	(2.62)*	(4.34)**
Two-way CRの t 値	(5.59)**	(−1.33)	(1.36)	(2.62)*	(4.34)**
(17) FMの係数	0.040	0.000	−0.001	0.159	0.052
FMの t 値	(11.14)**	(0.07)	(−0.10)	(5.13)**	(5.46)**

（注）推定モデルは以下のようである。

$$NI_{it} = \gamma_0 + \gamma_1 DR_{it} + \gamma_2 R_{it} + \gamma_3 DR^*R_{it} + \varepsilon_{it}, \quad (16)$$

$$NI_{it} = \gamma_0 + \gamma_1 DR_{it} + \gamma_2 R_{it} + \gamma_3 DR^*R_{it} + \gamma_4 R_{it-1} + \varepsilon_{it}, \quad (17)$$

ただし，NI_{it}：当期純利益，R_{it}：期首３ヵ月後から期末３ヵ月後までの年次リターン，DR_{it}：R_{it}が負の場合は１，それ以外はゼロのダミー変数。なお，下添字 i と t は，それぞれ企業と年度を表しており，NIは，期首の時価総額でデフレートされている。

*　５％水準で有意　**　１％水準で有意。

均値である.

表３ Panel Bの結果は，One-way CR（year）とTwo-way CRの t 値がほぼ同一であるという点と，One-way CR（firm）の t 値が，One-way CR（year），Two-way CRおよびFMの t 値と比較すると非常に大きいという点では，表２ Panel Bの結果と類似している。この結果は，ショート・パネルを用いた場合における(11)の考察を支持するものである。

一方，FMの t 値は，表２ Panel Bの結果と異なり，One-way CR（year）およびTwo-way CRの t 値とかなり差異がある。例えば，(16)の R に関する t 値は，One-way CR（year）とTwo-way CRが共に1.24であるのに対して，FMでは0.06であり，DR^*R に関する t 値は，One-way CR（year）が2.69でTwo-way CRが2.68であるのに対して，FMでは4.95である。この結果は，(12)が成立していないということを意味しているが，これは，説明変数の年度間の変動が大きくない場合にはという，(12)成立のための前提条件が，株式リターンを説明変数とする(16)(17)では成り立っていないということに起因していると考えられる。

最後に，表３の推定結果は，誤差項のクロスセクショナルな相関を考慮するOne-way CR（year），Two-way CR，FMの手法を用いると，係数の有意水準は低下するものの，保守主義の影響を表す DR^*R の係数は依然５％水準で有意に正であるので，会計の保守主義を支持する結果となっている。

６．おわりに

会計・ファイナンス領域における実証研究では，企業と年度のパネル・データを用いた推定がしばしば行われるが，その際に，誤差項の標準的仮定（均一分散で互いに無相関）が満たされない場合の対処法として頻繁に用いられているのが，Fama-MacBeth（FM）の手法ならびにCluster-robust（CR）の手法である。

しかしながら，FMの手法は，その広範な普及にもかかわらず，現在までその理論的考察が殆ど行われておらず，CRの手法についても，最近の研究が多いこともあってか，未だその適切な使用方法が十分に認知されていない。そこで本稿では，最初に，FMの手法に関する理論的分析を行い，次に，CRの手法についてその理論的概要および使用に関する実務的指針について述べた後に，FMの手法とCRの手法の特徴を比較している。また，最後に，各種の手法が係数の有意性検定の結果にどのような影響を与えるのかを，実際のデータを用いて調査している。

第１に，FMの手法に関する理論的分析からは，年度に亘って説明変数にそれほど大きな変化がない場合には，FMの手法を用いることによって，誤差項の未知のクロスセクショナルな相関に対処できるという考察を得ている。また，誤差項の異なる年度間の相関に対処するために提案されている各種の調整済みFMの手法については，何れもアドホックな対処方法で理論的根拠を欠いているので，安易に使用するべきではないといえる。

第２に，CRの手法を実際に使用する際の注意事項として，（ⅰ）クラスター数が20～30個以下である場合には分散の過小推定バイアスに対処するための調整が必要である，（ⅱ）Two-way CRの手法から得られる分散推定量には正定値性が保証されておらず，現段階ではその根本的な解決策は見つかっていない，（ⅲ）ショート・パネル・データを用いた分析では，クラスター数の少ないディメンションでCRの手法を用いることが重要である，という３点について指摘している。

第３に，両手法の特徴の比較としては，CRの手法に基づく分散推定量が，基本的にHCやHAC推定量と同じ発想で，残差が誤差項の一致推定量となるということを利用して誤差項の共分散行列の一致推定量を求めているのに対して，FMの手法は，誤差項の共分散行列を推定することなしに分散推定量を求めようとするものであるので，CRの手法に対する一種の簡便法であるといえる。さらに，会計・ファイナンスのパネル推定では，企業数が大きく年度数が小さいショート･パネル･データが用いられるケースが非常に多く，そのような場合には，（ⅰ）Two-way CRと年度に関するOne-way CRの両手法から得られる標準誤差は近似し，さらに説明変数の年度間による変動が大きくない場合には，（ⅱ）FMの手法，Two-way CRの手法，年度に関するOne-way CRの手法の何れの手法を用いても，得られる標準誤差には大差がないという理論的帰結を得ている。

最後に，実際のデータを用いて，上記の考察を検証した所，（ⅰ）についてはショート・パネル・データであれば成立し，（ⅱ）については，年度間の変動が小さいと思われる会計利益を説明変数とするモデルでは成立していたが，変動が大きいと思われる株式リターンを説明変数とするモデルでは成立しておらず，本稿の理論的考察と一致する証拠が得られている。

以上の本稿の結果は，その広範な普及にも関わらずこれまで理論的な分析が殆ど行われてこなかったFMの手法を理論的に解明し，CRの手法との関連性を導いている点で，今後の両手法の使用に関して一定の指針を与えるものと考えられる。

本研究は，平成28年度関西大学研修員研修費およびJSPS科研費16K03762の助成を受けて行ったものである。

引用文献

井手健二（2006），「わが国証券市場における純資産直入項目の情報価値」『武蔵大学論集』第54巻第２号，139-154。

久保田敬一・須田一幸・竹原均（2006），「株式収益率と経営者報酬における包括利益の情報内容」『経営財務研究』第26巻第1-2号，53-69。

高田知実（2006），「利益／株価比率を利用した保守主義の定量化」『経済経営研究所年報』第56号，1-38。

田澤宗裕（2004），「会計利益と発生項目の適時性－保守主義に焦点を当てて－」『産業経理』第64巻第２号，94-107。

若林公美（2009），『包括利益の実証研究』中央経済社。

若林公美（2010），「包括利益と純利益の特性比較」『証券アナリストジャーナル』第48巻第５号，17-25。

Angrist, J. D., and J-S. Pischke. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Arellano, M. (1987), "Computing Robust Standard Errors for Within-groups Estimators," *Oxford Bulletin of Economics and Statistics* 49 (4): 431-434.

Ball, R., S. P. Kothari, and A. Robin. (2000), "The Effect of International Institutional Factors on Properties of Accounting Earnings," *Journal of Accounting and Economics* 29 (1): 1-51.

Basu, S. (1997), "The Conservatism Principle and the Asymmetric Timeliness of Earnings," *Journal of Accounting and Economics* 24 (1): 3-37.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics* 90 (3): 414-427.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. (2011), "Robust Inference With Multiway Clustering," *Journal of Business and Economic Statistics* 29 (2): 238-249.

Cameron, A. C., and D. L. Miller. (2010), "Robust Inference with Clustered Data," In *Handbook of Empirical Economics and Finance*, edited by A. Ullah and D. Giles, Florida: Chapman & Hall/CRC: 1-28.

Chakravarty, S., H. Gulen, and S. Mayhew. (2004), "Informed Trading in Stock and Option Markets," *The Journal of Finance* 59 (3): 1235-1258.

Chambers, D., T. J. Linsmeier, C. Shakespeare, and T. Sougiannis. (2007), "An Evaluation of SFAS No. 130 Comprehensive Income Disclosures," *Review of Accounting Studies* 12 (4): 557-593.

Cochrane, J. H. (2005), *Asset Pricing, Revised ed.*, Princeton, NJ: Princeton University Press.

Fama, E. F., and J. D. MacBeth. (1973), "Risk, Return, and Equilibrium: Empirical Tests," *The Journal of Political Economy* 81 (3): 607-636.

Fama, E. F., and K. R. French. (2002), "Testing Trade-Off and Pecking Order Predictions About Dividends and Debt," *The Review of Financial Studies* 15 (1): 1-33.

Gow, I. D., G. Ormazabal, and D. J. Taylor. (2010), "Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research," *The Accounting Review* 85 (2): 483-512.

Liang, K-Y., and S. L. Zeger. (1986), "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika* 73 (1): 13-22.

Newey, W. K., and K. D. West. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55 (3): 703-708.

Petersen, M. A. (2009), "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *The Review of Financial Studies* 22 (1): 435-480.

Pope, P. F., and M. Walker. (1999), "International Differences in the Timeliness, Conservatism, and Classification of Earnings," *Journal of Accounting Research* 37 (Supplement): 53-87.

Shuto, A., and T. Takada. (2010), "Managerial Ownership and Accounting Conservatism in Japan: A Test of Management Entrenchment Effect," *Journal of Business Finance & Accounting* 37 (7-8): 815-840.

Thompson, S. B. (2011), "Simple Formulas for Standard Errors that Cluster by Both Firm and Time," *Journal of Financial Economics* 99 (1): 1-10.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48 (4): 817-838.