会計学の実証研究における標準誤差の問題

山　本　達　司*

Ⅰ　はじめに

会計データは企業と期間から構成される膨大なパネルデータであり，現代の会計学の実証研究においてはパネルデータを用いた回帰分析の利用が一般的である。回帰分析において，標準回帰モデル（ordinary least square model）の仮定が満たされている場合，回帰係数（最小二乗推定量）の標準誤差にバイアスはない。しかし，それが満たされない場合，回帰係数の標準誤差は過大または過小に推定され，それが誤った分析結果を導く可能性がある。

このことについて，ファイナンス研究である Petersen（2009）は，次のように述べている[1]。

- 先行研究はパネルデータの標準誤差を推定するために，いろいろな方法を使っているが，選ばれた方法は多くの場合，間違っており，またどの方法を使うべきかについて，ほとんどガイダンスを示していない。
- Petersen（2009）の目的は，標準誤差の推定方法を比較し，どのようにすれば正しい標準誤差を選択できるのかを理解することである。

そして Petersen（2009）は誤差項の相関を次の 2 つに分類して，この問題を検討している[2]。

---

*　Yamamoto, Tatsushi
1　Petersen（2009）, p.436
2　Petersen（2009）, p.436

- 企業効果（firm effect）
  －同一企業内で，異なる時点の誤差項間に相関がある。
- 時間効果（time effect）
  －同一期間内で，異なる企業の誤差項間に相関がある。

会計データにおいても企業効果，時間効果は存在すると考えられるので，それらへの対処法は重要である。そして，現代の会計学の実証研究においても，Petersen（2009）が推奨するクラスター・ロバストな標準誤差の利用が一般的となっている[3]。

しかし，ファイナンスの実証研究のために提唱されたクラスター・ロバストな標準誤差を，会計学の実証研究において無批判に受入れてもよいのであろうか。本論文の目的は，この問題について検討し，会計学の実証研究において，企業効果，時間効果への適切な対処法を提言することである。

Ⅱ　回帰係数の標準誤差問題の本質

平均値 μ の母集団 $\{x_1,\cdots,x_N\}$ から標本サイズ $n(n<N)$ の標本 $\{x_1,\cdots,x_n\}$ を抽出し，標本平均 $\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ による母集団の平均値 μ の統計的推論について考える。このとき，標本の各観察値 x_i は，次のように表すことができる。

$$x_i=\mu+\widetilde{\varepsilon_i}\quad(i=1,\cdots,n)$$

$\widetilde{\varepsilon_i}$ は観察値 x_i の母集団平均値 μ からの乖離度を表し（以下，本論文では $\widetilde{\varepsilon_i}$ を「**誤差項**」と記す），標本抽出に依存して決まる値であるから確率変数となる。ここで，次の3つの仮定をおく。

仮定①　各誤差項の期待値が0である。$(E(\widetilde{\varepsilon_i})=0\ (i=1,\cdots,n))$

仮定②　各誤差項の分散が等しい。$(Var(\widetilde{\varepsilon_i})=\sigma_{\varepsilon}^2\ (i=1,\cdots,n))$

仮定③　異なる誤差項が独立である。$(Cov(\widetilde{\varepsilon_i},\widetilde{\varepsilon_j})=0)$，$i=1,\cdots,n,\ j=1,\cdots,n,\ i<j)$

３　例えば，会計学のトップ・ジャーナルの１つである *Journal of Accounting and Economics,* Volume 76, Issue 1, August 2023 において，11 篇の実証研究が掲載されている中で，one way cluster の標準誤差を採用している論文が５篇（Amiraslani, Donovan, Phillips, and Wittenberg-Moerman（2023）; Choi, Choi, and Malik（2023）; Colonnello, Koetter, and Wagner（2023）; Goldman and Ozel（2023）; Lennox, Wang, and Wu（2023）), two way cluster の標準誤差を採用している論文が２篇ある（deHaan, Li, and Watts（2023）; Kim and Valentine（2023））。

仮定①は，観察値 x_i の誤差項が平均的に0になることを意味する。仮定②は，観察値 x_i が母集団のどのクラスター（例えば，母集団が企業である場合，クラスターは業種，年度など）に属していても，誤差項 ε_i の分散は一定であることを意味する。仮定③は，異なる観察値の誤差項が独立であることを意味する。

このとき，標本平均 $\overline{x}$ の期待値と分散は次のようになる。

$$E(\overline{x}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n} E(\mu + \widetilde{\varepsilon_i}) = \frac{1}{n}\sum_{i=1}^{n}[E(\mu) + E(\widetilde{\varepsilon_i})] = \mu$$

（仮定①より）

$$\begin{aligned}
Var(\overline{x}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(x_i) + 2\sum_{i<j} Cov(x_i, x_j)\right] \\
&= \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(\mu + \widetilde{\varepsilon_i}) + 2\sum_{i<j} Cov(\mu + \widetilde{\varepsilon_i}, \mu + \widetilde{\varepsilon_j})\right] = \frac{1}{n^2}\left[\sum_{i=1}^{n} Var(\widetilde{\varepsilon_i}) + 2\sum_{i<j} Cov(\widetilde{\varepsilon_i}, \widetilde{\varepsilon_j})\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(\widetilde{\varepsilon_i}) = \frac{\sigma_\varepsilon^2}{n} \quad \text{（仮定②③より）}
\end{aligned}$$

そして中心極限定理より，n が十分大きければ，

$$\overline{x} \sim N\left(\mu, \frac{\sigma_\varepsilon^2}{n}\right)$$

$$\frac{\overline{x} - \mu}{\sigma_\varepsilon/\sqrt{n}} \sim N(0, 1) \qquad (1)$$

となり，（1）式に基づいて，母集団の平均値 μ の統計的推論が行われる[4]。

しかし現実の会計データにおいて，仮定②が満たされないことがある[5]。例えば，日本企業の利益率を調査対象とする場合，業種ごとに誤差項の分散が異なることは十分考えられる。また現実の会計データにおいて，仮定③が満たされないことがある。例えば，クロスセクションで日本企業の利益率を調査する場合，景気変動等のマクロ経済的要因が多くの企業の利益率に影響を与える。このような状況を考慮せずに，仮定②③に基づいて母集団の平均値の統計的推論を行うと，$Var(\overline{x}) = \sigma_\varepsilon^2/n$ により $Var(\overline{x})$ が過大または過小に推定され，その結果，（1）式の左辺の $\sigma_\varepsilon/\sqrt{n}$ が過大または過小となるため，バイアスのある分析結果が得られる可能性がある。

上述した平均値の統計的推論は，定数項のみを独立変数とする回帰モデル（2）にお

4　一般に，母集団の分散 σ_ε^2 は未知であるため，標本分散 $\widehat{\sigma_\varepsilon}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ を用いて統計的推論が行われる。このことは，本論文のこれ以後の議論に影響を与えない。

5　仮定①は満たされると考えてよい。その理由は，脚注6を参照してほしい。

ける回帰係数 β_0 の統計的推論とみなすことができる。そして回帰モデル（2）において，（1）式の $\sigma_\varepsilon/\sqrt{n}$ は**回帰係数の標準誤差**と呼ばれる。

$$x_i = \beta_0 \cdot 1 + \varepsilon_i \quad (i = 1, \cdots, n) \tag{2}$$

このように単純なクロスセクションの単回帰モデル（2）でさえ，仮定②③が満たされないことにより，上述のようにバイアスのある分析結果が得られる可能性がある。これがパネルデータを用いた重回帰モデルになると，この問題はより複雑かつ深刻になると予想される。これが，回帰係数の標準誤差問題の本質である。

Ⅲ　パネルデータを用いた標準回帰モデル

（1）標準回帰モデルの仮定

N 個の企業，T 期間のパネルデータに対して，K 個の独立変数を用いた回帰モデルは，次のように表される。

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_{it} \\ \vdots \\ Y_{NT} \end{bmatrix} = \begin{bmatrix} X_{1,1} & & \cdots & & X_{1,K} \\ \vdots & & & & \vdots \\ X_{it,1} & \cdots & X_{it,k} & \cdots & X_{it,K} \\ \vdots & & & & \vdots \\ X_{NT,1} & & \cdots & & X_{NT,K} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{it} \\ \vdots \\ \varepsilon_{NT} \end{bmatrix}$$

$$(K < NT; i = 1, \cdots, N; t = 1, \cdots, T)$$

ここにおいて，Y_{it} は企業 i の t 期の従属変数 Y の観察値，$X_{it,k}$ は i 企業の t 期の独立変数 X_k の観察値，β_k は独立変数 X_k の回帰係数，ε_{it} は企業 i の t 期の誤差項である。本論文では，これを（3）式のように表すことにする。

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{3}$$

標準回帰モデルでは，（4）式，（5）式が仮定されている。

$$E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \tag{4}$$

$$Var(\boldsymbol{\varepsilon}) = \begin{bmatrix} \sigma_\varepsilon^2 & & & 0 \\ & \sigma_\varepsilon^2 & & \\ & & \ddots & \\ 0 & & & \sigma_\varepsilon^2 \end{bmatrix} = \sigma_\varepsilon^2 \boldsymbol{E} (\equiv \boldsymbol{\Sigma}_{OLS}) \tag{5}$$

（$\boldsymbol{E}$ は $[NT \times NT]$ の単位行列
以下では，このように行列 X が N 行 T 列の行列
であることを $X = [N \times T]$ などと書く。）

(4) 式は誤差項の期待値が 0 であることを意味し（$E(\varepsilon_{it}) = 0,\ it = 1, \cdots, NT$），II 節の仮定①に対応する。(5) 式は誤差項の分散が一定であり（$Var(\varepsilon_{it}) = \sigma_\varepsilon^2, it = 1, \cdots, NT$），かつ異なる誤差項の共分散が 0 であることを意味し（$Cov(\varepsilon_{it}, \varepsilon_{js}) = 0,\ it \neq js,\ it = 1, \cdots, NT;\ js = 1, \cdots, NT$），それぞれ II 節の仮定②，仮定③に対応する。

(2) 回帰係数の一致性

(a) 標準回帰モデルにおける回帰係数の一致性

回帰モデル（3）の回帰係数ベクトル $\hat{\boldsymbol{\beta}}$ は次のように表され，

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X\beta}+\boldsymbol{\varepsilon}) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} \tag{6}$$

$\hat{\boldsymbol{\beta}}$ の期待値と分散は，次のようになる。

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E(\boldsymbol{\varepsilon})$$
$$Var(\hat{\boldsymbol{\beta}}) = E(\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}}))(\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}}))'$$

一般に仮定（4）式は満たされると考えてよいので[6]，（7）式のように推定値 $\hat{\boldsymbol{\beta}}$ は真の $\boldsymbol{\beta}$ の不偏推定量となる。

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E(\boldsymbol{\varepsilon}) = \boldsymbol{\beta} \tag{7}$$

しかし，非現実的な推定量でも不偏性をもつ場合がある。例えば，平均値 μ の母集

6　一般に回帰モデル $Y = \alpha + \beta X + \varepsilon\ (E(\varepsilon) = c (\neq 0))$ において，$\varepsilon' \equiv \varepsilon - c$ と定義すれば，$Y = (\alpha + c) + \beta X + \varepsilon'\ (E(\varepsilon') = 0)$ となるので，定数項のある回帰モデルであれば，誤差項の期待値が 0 と仮定して回帰分析を行っても問題はない（浅野・中村（2009），p.26）

団から得られたサンプルサイズ N の無作為標本において，II 節の仮定①（誤差項の期待値＝0）が満たされれば，最初と最後の観察値の平均値 $(x_1+x_N)/2$ や，N 番目の観察値 x_N も μ の不偏推定量となる[7]。そこで一致性が，推定量を評価する重要な基準となる。

一般に，真のパラメータ θ に対して，サンプルサイズ N のときの推定量 $\widehat{\theta_N}$ とすると，$\widehat{\theta_N}$ が真のパラメータ θ の一致推定量であるための条件は，$\mathrm{plim}_{N\to\infty}\widehat{\theta_N}=\theta$ であり，これは $\lim_{N\to\infty}AVar\widehat{\theta_N}=0$ と書き直せる。

そこで標準回帰モデルは，仮定（4）式に加えて仮定（5）式を課し，推定値 $\hat{\boldsymbol{\beta}}$ の一致性を確保している。そのメカニズムは，次のようである。（4）式の仮定が満たされたとすると，（7）式が成り立つから，

$$
\begin{aligned}
Var(\hat{\boldsymbol{\beta}}) &= E(\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}}))(\hat{\boldsymbol{\beta}}-E(\hat{\boldsymbol{\beta}}))' \\
&= E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})' \qquad \text{（(7) 式より）} \\
&= E\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\ ((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon})'\right] \qquad \text{（(6) 式より）} \\
&= E\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')'\right] \\
&= E\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{X}((\boldsymbol{X}'\boldsymbol{X})^{-1})'\right] \\
&= E\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\right] \qquad (\boldsymbol{X}'\boldsymbol{X}, (\boldsymbol{X}'\boldsymbol{X})^{-1}\ \text{は対称行列}) \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \quad (E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\equiv\boldsymbol{\Sigma}) \\
&= \frac{1}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}\frac{\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}
\end{aligned}
\tag{8}
$$

$$
\begin{pmatrix}
Var(\hat{\boldsymbol{\beta}})=[K\times K], \boldsymbol{X}'\boldsymbol{X}=[K\times NT][NT\times K]=[K\times K], \\
(\boldsymbol{X}'\boldsymbol{X})^{-1}=[K\times K], \boldsymbol{\Sigma}=[NT\times NT], \boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}=[K\times K]
\end{pmatrix}
$$

そして，（5）式の仮定が満たされるとすると，

$$
\begin{aligned}
Var(\hat{\boldsymbol{\beta}}) &= \frac{1}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}\frac{\boldsymbol{X}'\boldsymbol{\Sigma}_{OLS}\boldsymbol{X}}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1} = \frac{1}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}\frac{\boldsymbol{X}'\sigma_\varepsilon^2\boldsymbol{E}\boldsymbol{X}}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1} \\
&= \frac{\sigma_\varepsilon^2}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1} = \frac{\sigma_\varepsilon^2}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1} \\
&= \frac{\sigma_\varepsilon^2}{NT}\ \text{［有限な行列］}
\end{aligned}
\tag{9}
$$

7　浅野・中村（2009），p.275

ここで $NT \to \infty$ のとき，［$Var(\hat{\boldsymbol{\beta}})$ のすべての主対角要素 $\widehat{\beta_k}(k=1,\cdots,K)$］$\xrightarrow{p} 0$ となり，推定値 $\hat{\boldsymbol{\beta}}$ は一致推定量となる。

（b）Petersen（2009）の表記法

Petersen（2009）は，パネルデータにおいて標準回帰モデルを仮定したとき，回帰係数の一致性について，次のように記述している。[8]

パネル回帰モデル $Y_{it}=X_{it}\beta+\varepsilon_{it}(i=1,\cdots,N;t=1,\cdots,T)$ において，回帰係数 $\widehat{\beta}_{OLS}$ は，

$$\widehat{\beta}_{OLS}=\beta+\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}^2}$$

となり，標準回帰モデルの仮定（5）式より，$Var(\varepsilon_{it})=\sigma_\varepsilon^2$ かつ $Cov(\varepsilon_{it},\varepsilon_{js})=0$ であるので（$it\neq js, it=1,\cdots,NT; js=1,\cdots,NT$），回帰係数の漸近分散 $AVar[\widehat{\beta}_{OLS}]$ は次のようになる。[9]

$$\begin{aligned}
AVar[\widehat{\beta}_{OLS}] &= AVar\left[\beta+\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}^2}\right]\\
&= AVar\left[\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}^2}\right]\\
&= \operatorname*{plim}_{\substack{N\to\infty\\ T\ \ fixed}}\left[\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}^2}\right]^2\\
&= \operatorname*{plim}_{\substack{N\to\infty\\ T\ \ fixed}}\left[\frac{1}{N^2}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}\varepsilon_{it}\right)^2\left(\frac{\sum_{t=1}^{T}\sum_{i=1}^{N}X_{it}^2}{N}\right)^{-2}\right]\\
&= \operatorname*{plim}_{\substack{N\to\infty\\ T\ \ fixed}}\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{t=1}^{T}X_{it}^2\varepsilon_{it}^2\left(\sum_{t=1}^{T}\frac{1}{N}\left(\sum_{i=1}^{N}X_{it}^2\right)\right)^{-2}\right]\\
&= \frac{1}{N^2}\cdot NT\sigma_X^2\sigma_\varepsilon^2\cdot\frac{1}{T^2\sigma_X^2\sigma_X^2}\\
&= \frac{\sigma_\varepsilon^2}{NT\sigma_X^2}
\end{aligned} \tag{10}$$

8　本文は Petersen（2009），p.438 について，筆者が加筆・要約している。

9　Petersen（2009），p.438 では，$AVar[\widehat{\beta}_{OLS}-\beta]$ を求めているが，$AVar[\widehat{\beta}_{OLS}]$ を求めて論理展開する方がわかりやすいので，本論文では $AVar[\widehat{\beta}_{OLS}]$ を計算している。

（10）式は（9）式に対応している。ここにおいて，$T \to \infty$ のとき $AVar[\hat{\beta}_{OLS}] \xrightarrow{p} 0$ となるので，$\hat{\beta}_{OLS}$ は一致推定量となる。

Ⅳ　パネル回帰モデルにおいて解決すべき問題

(1) 標準回帰モデルの標準誤差のバイアス

現実のパネルデータでは，標準回帰モデルの仮定（5）式（$Var(\boldsymbol{\varepsilon}) = \sigma_{\varepsilon}^2 \boldsymbol{E}$）が成り立たないことが多い。それにもかかわらず（5）式を用いたときに生じるバイアスを，Petersen（2009）は次のような例を用いて示している[10]。

企業が異なれば誤差項は独立であるが（$Cov(\varepsilon_{it}, \varepsilon_{js}) = 0,\ i \neq j$），同一企業内では異なる時点の誤差項に時系列相関（企業効果）がある次のような状況を想定する（$Cov(\varepsilon_{it}, \varepsilon_{is}) \neq 0,\ t \neq s$）。

$\varepsilon_{it} = \gamma_i + \eta_{it}$　（$E(\gamma_i) = E(\eta_{it}) = 0,\ Var(\gamma_i) \equiv \sigma_{\gamma}^2,\ Var(\eta_{it}) \equiv \sigma_{\eta}^2$ とする）

γ_i = 誤差項の企業効果

η_{it} = 誤差項の企業効果以外の部分 ($Cov(\eta_{it}, \eta_{is}) = 0, t \neq s$)

つまり，企業は過去を部分的に引きずるが，引きずり方 γ_i は企業ごとに異なっている。そして，同一企業内の誤差項に企業効果を仮定したので，独立変数にも企業効果を仮定する。

$X_{it} = \mu_i + \upsilon_{it}$　　（$E(\mu_i) = E(\upsilon_{it}) = 0,\ Var(\mu_i) \equiv \sigma_{\mu}^2,\ Var(\upsilon_{it}) \equiv \sigma_{\upsilon}^2$ とする）

μ_i = 独立変数の企業効果

υ_{it} = 独立変数の企業効果以外の部分（$Cov(\upsilon_{it}, \upsilon_{is}) = 0, t \neq s$）

ここにおいて，$\gamma_i, \eta_{it}, \mu_i, \upsilon_{it}$ はすべて独立であると仮定すると，ε_{it} と ε_{js} の相関係数 ρ_{ε}, X_{it} と X_{js} の相関係数 ρ_X は次のようになる。

$$\rho_{\varepsilon} \equiv corr(\varepsilon_{it}, \varepsilon_{js}) = \frac{\sigma_{\gamma}^2}{\sigma_{\varepsilon}^2} = \frac{\sigma_{\gamma}^2}{\sigma_{\gamma}^2 + \sigma_{\eta}^2} > 0 \text{（} i = j, t \neq s \text{ のとき）}$$

$$\rho_X \equiv corr(X_{it}, X_{js}) = \frac{\sigma_{\mu}^2}{\sigma_X^2} = \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2 + \sigma_{\upsilon}^2} > 0 \text{（} i = j, t \neq s \text{ のとき）}$$

10　本文は Petersen（2009），pp.438-440 について，筆者が加筆・要約している。

ここで，回帰係数の推定量の漸近分散を求めると，次のようになる[11]。

$$AVar[\hat{\beta}_{OLS}] = \underset{\substack{N\to\infty \\ T\ \ fixed}}{\text{plim}} \left[\frac{1}{N^2} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left(\frac{\sum_{t=1}^{T}\sum_{i=1}^{N} X_{it}^2}{N} \right)^{-2} \right]$$

（ここまでは（10）式 4 行目と同じ）

（同一企業について，X_{it} ならびに ε_{it} に時系列相関がある）

$$= \underset{\substack{N\to\infty \\ T\ \ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}\varepsilon_{it} \right)^2 \left(\frac{\sum_{t=1}^{T}\sum_{i=1}^{N} X_{it}^2}{N} \right)^{-2} \right]$$

$$= \underset{\substack{N\to\infty \\ T\ \ fixed}}{\text{plim}} \left[\frac{1}{N^2} \sum_{i=1}^{N} \left(\sum_{t=1}^{T} X_{it}^2\varepsilon_{it}^2 + 2\sum_{s=t+1}^{T}\sum_{t=1}^{T-1} X_{it}X_{is}\varepsilon_{it}\varepsilon_{is} \right) \left(\sum_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{N} X_{it}^2 \right)^{-2} \right]$$

$$= \frac{1}{N^2} \left[\sum_{i=1}^{N} \left(\sum_{t=1}^{T} \sigma_X^2\sigma_\varepsilon^2 + 2\sum_{s=t+1}^{T}\sum_{t=1}^{T-1} \rho_X\sigma_X^2\rho_\varepsilon\sigma_\varepsilon^2 \right) \right] \left(\sum_{t=1}^{T} \sigma_X^2 \right)^{-2}$$

$$= \frac{1}{N^2} N \left(T\sigma_X^2\sigma_\varepsilon^2 + (T^2 - T)\rho_X\sigma_X^2\rho_\varepsilon\sigma_\varepsilon^2 \right) \frac{1}{T^2\sigma_X^2\sigma_X^2}$$

$$= \frac{\sigma_\varepsilon^2}{NT\sigma_X^2}(1 + (T-1)\rho_X\rho_\varepsilon) \tag{11}$$

（11）式から，標準回帰モデルの標準誤差のバイアスをまとめると，次のようである。

〈標準回帰モデルの標準誤差のバイアス〉

$\rho_X > 0$ かつ $\rho_\varepsilon > 0$ のとき（プラスの企業効果があるとき），標準回帰モデルの回帰係数の標準誤差 $AVar[\hat{\beta}_{OLS}]$ は，真の標準誤差を

$$\frac{\sigma_\varepsilon^2(T-1)\rho_X\rho_\varepsilon}{NT\sigma_X^2} = \frac{\sigma_\varepsilon^2\rho_X\rho_\varepsilon}{N\sigma_X^2}\cdot\frac{T-1}{T} = \frac{\sigma_\varepsilon^2\rho_X\rho_\varepsilon}{N\sigma_X^2}\left(1 - \frac{1}{T}\right)$$

だけ過小推定しており，過小推定の程度はサンプルにおける企業数 N の減少関数であり，同一企業内の独立変数の時系列相関 ρ_X，同一企業内の誤差項の時系列相関 ρ_ε，サンプルにおける期間の長さ T の増加関数である。そして $\rho_X\rho_\varepsilon > 0$ であるので，$T \to \infty$ としても，過小推定された部分は 0 に収束しないため，$\hat{\beta}_{OLS}$ は一致推定量ではなくなる。

11　Petersen（2009），p.439-440 では，$AVar[\hat{\beta}_{OLS} - \beta]$ を求めているが，$AVar[\hat{\beta}_{OLS}]$ を求めて論理展開する方がわかりやすいので，本論文では $AVar[\hat{\beta}_{OLS}]$ を計算している。

（2）Fama-MacBeth 回帰モデルの標準誤差のバイアス

Fama-MacBeth 回帰は，パネルデータにおいてまず期間ごとのクロスセクション回帰を行い，回帰係数の平均値をパネル回帰の回帰係数として推定する方法である。Petersen（2009）は，標準回帰モデルの仮定（5）式が満たされないとき，Fama-MacBeth 回帰の回帰係数の標準誤差に生じるバイアスについて，次のように記述している[12]。

クロスセクション回帰 $Y_{it} = \alpha_t + X_{it}\beta_t + \varepsilon_{it}$ を $t = 1, \cdots, T$ について行い，回帰係数 $\widehat{\beta_t}$ の平均値を Fama-MacBeth の回帰係数 $\widehat{\beta_{FM}}$ として推定する。

$$\widehat{\beta_{FM}} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\beta_t}$$

$Cov(\widehat{\beta_t}, \widehat{\beta_s}) = 0(t \neq s)$, $Var(\widehat{\beta_t})(\equiv \sigma_\beta^2) =$ 一定を仮定すると，

$$Var(\widehat{\beta_{FM}}) = Var\left(\frac{1}{T}\sum_{t=1}^{T}\hat{\beta}_t\right) = \frac{1}{T^2}\sum_{t=1}^{T}Var(\hat{\beta}_t) = \frac{1}{T}\sigma_\beta^2$$

ここで σ_β^2 は未知なので，$S^2(\widehat{\beta_t}) = \frac{1}{T-1}\sum_{t=1}^{T}(\widehat{\beta_t} - \widehat{\beta_{FM}})^2$ で代用すると，$\widehat{\beta_{FM}}$ の標本分散 $S^2(\widehat{\beta_{FM}})$ は次のようになる。

$$S^2(\widehat{\beta_{FM}}) = \frac{1}{T}\cdot\frac{1}{T-1}\sum_{t=1}^{T}(\widehat{\beta_t} - \widehat{\beta_{FM}})^2$$

つまり，標本分散 $S^2(\widehat{\beta_{FM}})$ も $Cov(\widehat{\beta_t}, \widehat{\beta_s}) = 0\ (t \neq s)$ の仮定に基づいて計算されている。

一方，期間ごとのクロスセクション回帰の回帰係数 $\widehat{\beta_t}$ は，

$$\widehat{\beta_t} = \beta + \frac{\sum_{i=1}^{N}X_{it}\varepsilon_{it}}{\sum_{i=1}^{N}X_{it}^2} \quad (t = 1, \cdots, T)$$

であり，$Cov(\widehat{\beta_t}, \widehat{\beta_s}) = 0(t \neq s)$ が成り立つためには，$Cov(X_{it}\varepsilon_{it}, X_{is}\varepsilon_{is}) = 0(t \neq s)$ が成り立たなければならない。しかし企業効果がある場合，$Cov(\varepsilon_{it}, \varepsilon_{is}) \neq 0(t \neq s)$ であるので，$Cov(X_{it}\varepsilon_{it}, X_{is}\varepsilon_{is}) \neq 0(t \neq s)$ である。つまり，Fama-MacBeth 回帰の標準誤差にはバイアスがあり，$\widehat{\beta_{FM}}$ の漸近分散 $AVar(\widehat{\beta_{FM}})$ は次のようになる[13]。

$$AVar(\widehat{\beta_{FM}}) = \frac{\sigma_\varepsilon^2}{NT\sigma_X^2}[1 + (T-1)\rho_X\rho_\varepsilon]$$

この結果は，企業効果がある場合の標準回帰モデルの漸近分散（(11）式）と同じで

12 本文は Petersen（2009），pp.446-447 について，筆者が加筆・要約している。
13 詳細は，Petersen（2009），p.447 を参照してほしい。

あるため，Fama-MacBeth 回帰の標準誤差には，標準回帰モデルの標準誤差と同じバイアスが存在する。

（3）パネル回帰モデルの標準誤差の課題

回帰係数ベクトル $\hat{\boldsymbol{\beta}}$ の分散共分散行列 $Var(\hat{\boldsymbol{\beta}})$ を示した（8）式を再掲すると，次のようである。

$$Var(\hat{\boldsymbol{\beta}})=\frac{1}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1}\frac{\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}}{NT}\left(\frac{\boldsymbol{X}'\boldsymbol{X}}{NT}\right)^{-1} \tag{8}$$

ここで会計データにおいては，標準回帰モデルの仮定（5）式が満たされないことが多い。このとき IV 節（1）で述べたように，推定された回帰係数の一致性は保証されない。そこで，仮定（5）式を緩和したとき，最小二乗推定量 $\hat{\boldsymbol{\beta}}$ が一致推定量となるように，誤差項の分散共分散行列 $\boldsymbol{\Sigma}$ が満たすべき条件を考えることにする。

（8）式より，$\hat{\boldsymbol{\beta}}$ が一致推定量となるための条件は，次の 2 つの仮定が成り立つことである。

仮定（i）　$NT \to \infty$ のとき，$\dfrac{\boldsymbol{X}'\boldsymbol{X}}{NT} \xrightarrow{p}$ 有限な正定値行列となる[14]。

仮定（ii）　$NT \to \infty$ のとき，$\dfrac{\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}}{NT} \xrightarrow{p}$ 有限な正定値行列となる。

ここにおいて，**仮定（i）**は一般に成り立つ。**仮定（ii）**については，標準回帰モデルの仮定（4）式，（5）式が満たされれば成り立つが，一般には成り立たない。その理由は，標準回帰モデルにおける誤差項の分散共分散行列 $\boldsymbol{\Sigma}_{OLS}$（（5）式）と一般的な回帰モデルにおける誤差項の分散共分散行列 $\boldsymbol{\Sigma}$（（12）式）を比較すればわかるように，$\boldsymbol{\Sigma}$ において，0 でない共分散要素 $\sigma_{ij}(\neq 0)$ の数が多いためである[15]。

$$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,NT}\\ \sigma_{2,1} & \sigma_2^2 & \cdots & \sigma_{2,NT}\\ \vdots & \vdots & & \vdots\\ \sigma_{NT,1} & \sigma_{NT,2} & \cdots & \sigma_{NT}^2\end{bmatrix} \tag{12}$$

そこで，（5）式の仮定を緩和しつつも，$\boldsymbol{\Sigma}$（（12）式）のいくつかの共分散要素

14　実対称行列 A について，「A が正定値行列であること」，「任意の実ベクトル $\mathbf{x}\,(\mathbf{x}\neq\mathbf{0})$ について $\mathbf{x}'A\,\mathbf{x}>0$ であること」，「A のすべての固有値は正であること」は，論理的に同値である。

15　厳密な数学的証明は，太田（2012），pp.5-8 などを参照してほしい。

$\sigma_{ij}(i \neq j)$ を 0 と仮定することによって，推定量 $\hat{\boldsymbol{\beta}}$ の一致性を確保することを考える。このとき，（8）式で表される $Var(\hat{\boldsymbol{\beta}})$ の主対角要素 $Var(\widehat{\beta_k})$ ついて，$Var(\widehat{\beta_k}) > 0\ (k = 1, \cdots, K)$ となる必要がある。

まとめると，標準回帰モデルの仮定（5）式が満たされないとき，望ましい分散共分散行列 $\boldsymbol{\Sigma}$ の条件は次の 2 つである。

［条件 1］ 標準回帰モデルの仮定（5）式を緩和しつつも，$\boldsymbol{\Sigma}$（（12）式）のいくつかの共分散要素 $\sigma_{i,j}(i \neq j)$ を 0 と仮定することによって，推定量 $\hat{\boldsymbol{\beta}}$ の一致性を保証する $\boldsymbol{\Sigma}$ である。

［条件 2］ $\hat{\boldsymbol{\beta}}$ の分散共分散行列 $Var(\hat{\boldsymbol{\beta}})$ の主対角要素が，すべて正となるような $\boldsymbol{\Sigma}$ である $(Var(\widehat{\beta_k}) > 0\ (k = 1, \cdots, K))$。

これが，パネル回帰モデルの標準誤差に関して解決すべき問題である。

Ⅴ　パネル回帰における標準誤差の改善

上記の［条件 1］，［条件 2］を満たす分散共分散行列 $\boldsymbol{\Sigma}$ を考案する様々な努力が，これまでに行われてきた。本節ではまず，クロスセクション回帰（$T = 1$）における White の分散共分散行列，もともと時系列回帰（$N = 1$）のために考案された Newey-West の分散共分散行列を検討した後，パネル回帰における one-way cluster の分散共分散行列，two-way cluster の分散共分散行列を検討する。

（1）White の分散共分散行列

White（1980）は（13）式に示す分散共分散行列 $\boldsymbol{\Sigma}_{White}$ を用いて，クロスセクション回帰（$T = 1$）において，誤差項の均一分散の仮定を緩和している。そのため，（5）式は（13）式のようになる。

$$\boldsymbol{\Sigma}_{White} = \begin{bmatrix} \sigma_1^2 & & & \mathbf{0} \\ & \sigma_2^2 & & \\ & & \ddots & \\ \mathbf{0} & & & \sigma_N^2 \end{bmatrix} \tag{13}$$

$\boldsymbol{\Sigma}_{White}$ においては，主対角要素 $\sigma_i^2(i = 1, \cdots, N)$ 以外の共分散要素はすべて 0 であるので，$\hat{\boldsymbol{\beta}}$ は一致推定量となり，［条件 1］は満たされる。そして，$\boldsymbol{\Sigma}_{White}$ において 0 でない要素は分散 $\sigma_i^2(> 0; i = 1, \cdots, N)$ だけであるので，$Var(\widehat{\beta_k}) > 0(k = 1, \cdots, K)$ とな

り，［条件 **2**］も満たされる。

$\boldsymbol{\Sigma}_{White}$ は，もともとクロスセクション回帰のために考案されたが，パネル回帰モデルにも適用可能である。

（2）Newey-West の分散共分散行列

（a）時系列データに対する Newey-West の分散共分散行列

Newey and West（1987）は，時系列回帰において（$N = 1$），同一企業における誤差項の時系列相関への対処法を，次の 2 つのステップで行っている。

〈ステップ **1**〉

一般的な回帰モデルにおける誤差項の分散共分散行列 $\boldsymbol{\Sigma}$（(12）式）においては，0 でない要素が多すぎるので，［条件 **1**］が成り立たず推定量 $\hat{\boldsymbol{\beta}}$ の一致性が得られない。そのため，M 期以内の誤差項に相関はあるが，$(M+1)$ 期以上離れている誤差項に相関はないと考えて（$1 \leq M \leq T-2$），$\boldsymbol{\Sigma}$ の右上と左下の共分散要素を 0 として（以下，この操作を「両端捨て」と記す）$\boldsymbol{\Sigma}'$ を作り，一致性を確保している。

$$\begin{cases} \sigma_{i,M+1+i} = \cdots = \sigma_{i,T} = 0 & (i = 1, \cdots, T-(M+1)) \\ \sigma_{M+1+j,\,j} = \cdots = \sigma_{T,\,j} = 0 & (j = 1, \cdots, T-(M+1)) \end{cases} \tag{14}$$

この段階で，分散共分散行列は次のようになる。

$$\boldsymbol{\Sigma}' = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,M+1} & & & \mathbf{0} \\ \sigma_{2,1} & \sigma_2^2 & \ddots & & \ddots & & \\ \vdots & \ddots & \ddots & & & \ddots & \\ \vdots & & & & & & \sigma_{T-M,T} \\ \sigma_{M+1,1} & & & & & & \vdots \\ & \ddots & & & \ddots & \ddots & \vdots \\ & & \ddots & & \ddots & \sigma_{T-1}^2 & \sigma_{T-1,T} \\ \mathbf{0} & & & \sigma_{T,T-M} & \cdots & \sigma_{T,T-1} & \sigma_T^2 \end{bmatrix}$$

〈ステップ **2**〉

（14）式で示すように，ステップ 1 で恣意的に共分散要素を 0 と仮定したため，［条件 **2**］が保証されない。そのため，0 でない共分散要素に 1 より小さいウェイト $w(l)$ を与えることにより，$Var(\beta_k) > 0 (k = 1, \cdots, K)$ を確保し，［条件 **2**］を保証している。

$$w(l)=1-\frac{l}{M+1}$$

$$\left(\begin{array}{l} M\text{：異なる誤差項に相関があると考える最大ラグ} \\ l\text{：異なる誤差項の期間差（}l=1,\cdots,M\text{）} \end{array}\right)$$

このウェイトは「時間の経過につれて誤差項の相関が弱まる」という経済的直観と整合的である。その結果，Newey-West の分散共分散行列 $\boldsymbol{\Sigma}_{NW}$ は次のようになる。

$$\boldsymbol{\Sigma}_{NW}=\begin{bmatrix} \sigma_1^2 & w(1)\sigma_{1,2} & \cdots & w(M)\sigma_{1,M+1} & & & \mathbf{0} \\ w(1)\sigma_{2,1} & \sigma_2^2 & \ddots & & \ddots & & \\ \vdots & \ddots & \ddots & & & \ddots & \\ w(M)\sigma_{M+1,1} & & & & & & w(M)\sigma_{T-M,T} \\ & \ddots & & & \ddots & \ddots & \vdots \\ & & \ddots & & \ddots & \sigma_{T-1}^2 & w(1)\sigma_{T-1,T} \\ \mathbf{0} & & & w(M)\sigma_{T,T-M} & \cdots & w(1)\sigma_{T,T-1} & \sigma_T^2 \end{bmatrix}$$

まとめると，［**条件 1**］，［**条件 2**］が成り立つためには，ステップ 1，ステップ 2 で示したように，両端捨てとウェイトづけが必要である。

(b) パネルデータに対する Newey-West の分散共分散行列

もともと，時系列回帰における誤差項の系列相関への対処法として考えられた Newey-West の手法であったが，パネルデータにおいて同一企業内で誤差項に時系列相関（企業効果）があるときの対処法としても用いられるようになった[16]。ここにおいて，誤差項に相関があると考える最大ラグは $(T-2)$ 期としている。その理由は，$(T-1)$ 期にすると，両端捨てができないからである。そのため，企業 i の誤差項の分散共分散行列 $\sum_{NW\_i}$ は次のようになる。

16　この手法を用いた研究として，Brockman and Chung（2001），MacKay（2003），Bertrand, Duflo, and Mullainathan（2004），Doidge（2004）などがある。

$$\sum_{NW\_i} = \begin{bmatrix} \sigma^2_{(i)1} & w(1)\sigma_{(i)1,2} & \cdots & w(T-2)\sigma_{(i)1,T-1} & 0 \\ w(1)\sigma_{(i)2,1} & \sigma^2_{(i)2} & \ddots & & w(T-2)\sigma_{(i)2,T} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ w(T-2)\sigma_{(i)T-1,1} & & \ddots & \sigma^2_{(i)T-1} & w(1)\sigma_{(i),T-1,T} \\ 0 & w(T-2)\sigma_{(i)T,2} & \cdots & w(1)\sigma_{(i)T,,T-1} & \sigma^2_{(i)T} \end{bmatrix}$$

$$\left(w(l) = 1 - \frac{l}{(T-2)+1} = 1 - \frac{l}{T-1} \quad (l = 1,\cdots,T-2), \quad \sum_{NW\_i} = [T \times T] \right)$$

そして，N 企業，T 期間のパネル回帰の誤差項の分散共分散行列を，（15）式のように仮定する。

$$\sum_{NW\_panel} = \begin{bmatrix} \sum_{NW\_1} & & & \mathbf{0} \\ & \sum_{NW\_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \sum_{NW\_N} \end{bmatrix} \left(\sum_{NW\_panel} = [NT \times NT] \right) \qquad (15)$$

$\sum_{NW\_panel}$ においては，異なる企業の誤差項に相関はないことが仮定されているので，$\sum_{NW\_panel}$ の右上と左下で十分な両端捨てがされており，たとえ $\sum_{NW\_panel}$ の小行列 $\sum_{NW\_i}(i = 1,\cdots,NT)$ において，誤差項に相関があると考えるラグをその最大値である $(T-2)$ 期としても，回帰係数の一致性は確保される（**［条件 1］**は満たされる)。そして，各企業の分散共分散行列 $\sum_{NW\_i}$ において上記のようにウェイトづけがされているので，$Var(\widehat{\beta_k}) > 0(k = 1,\cdots,K)$ となり，**［条件 2］**は満たされる。

（3）one way cluster の分散共分散行列

one way cluster の標準誤差は，次の 2 つの仮定に基づいている。

［one way－仮定 1］同一企業内では，すべての期間の誤差項に時系列相関がある。

$$(Cov(\varepsilon_{it}, \varepsilon_{is}) \neq 0,\ t \neq s)$$

［one way－仮定 2］異なる企業の誤差項に相関はない。

$$(Cov(\varepsilon_{it}, \varepsilon_{js}) = 0, i \neq j)$$

このように，同一企業内で誤差項に時系列相関がある一方で，異なる企業の誤差項に相

関がないと仮定するので，one way cluster の標準誤差と呼ばれている[17]。

［**one way－仮定 1**］より，企業 i の誤差項の分散共分散行列 $\boldsymbol{\Sigma}_{one\_way\_i}$ は（16）式のように書ける。

$$\boldsymbol{\Sigma}_{one\_way\_i}=\begin{bmatrix} \sigma^2_{(i)1} & \sigma_{(i)1,2} & \cdots & \sigma_{(i)1,T} \\ \sigma_{(i)2,1} & \sigma^2_{(i)2} & \cdots & \sigma_{(i)2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{(i)T,1} & \sigma_{(i)T,2} & \cdots & \sigma^2_{(i)T} \end{bmatrix} \tag{16}$$

そして［**one way－仮定 2**］より，異なる企業間で誤差項に相関がないので，サンプルの全企業，全期間についての分散共分散行列 $\boldsymbol{\Sigma}_{one\_way\_firm}$ は（17）式のようになる。

$$\boldsymbol{\Sigma}_{one\_way\_firm}=\begin{bmatrix} \boldsymbol{\Sigma}_{one\_way\_1} & & & & \mathbf{0} \\ & \ddots & & & \\ & & \boldsymbol{\Sigma}_{one\_way\_i} & & \\ & & & \ddots & \\ \mathbf{0} & & & & \boldsymbol{\Sigma}_{one\_way\_N} \end{bmatrix} \tag{17}$$

$\boldsymbol{\Sigma}_{one\_way\_firm}$ の右上と左下において 0 の要素が十分あるので，$\hat{\boldsymbol{\beta}}$ は一致推定量となり[18]，［**条件 1**］は満たされる。そして，Newey-West と異なり，企業 i の部分行列 $\boldsymbol{\Sigma}_{one\_way\_i}$ において恣意的に両端捨てを行っていないので，$Var(\widehat{\beta_k}) > 0(k = 1,\cdots,K)$ となり，［**条件 2**］は満たされる。

（4）two way cluster の分散共分散行列

two way cluster の標準誤差は，次の 3 つの仮定に基づいている。

［**two way－仮定 1**］同一企業内では，すべての期間の誤差項に時系列相関がある。

$$(Cov(\varepsilon_{it},\varepsilon_{is}) \neq 0,\ t \neq s)$$

［**two way－仮定 2**］同一期間内では，すべての企業の誤差項に相関がある。

$$(Cov(\varepsilon_{it},\varepsilon_{jt}) \neq 0, i \neq j)$$

［**two way－仮定 3**］期間も企業も異なれば，誤差項に相関はない。

$$(Cov(\varepsilon_{it},\varepsilon_{js}) = 0, i \neq j \text{ かつ } t \neq s)$$

17　同一期間内ではすべての企業の誤差項に相関があり，期間が異なれば誤差項に相関がない one way cluster の標準誤差も定義可能である。その具体例は，V 節（4）の（19）式である。

18　厳密な数学的証明は，太田（2017），p.51-53 などを参照してほしい。

このように，同一企業内で誤差項に時系列相関があり，かつ同一期間内ですべての企業の誤差項に相関があると仮定するので，two way cluster の標準誤差と呼ばれている。

two way cluster の分散共分散行列には，［**two way－仮定 1**］より $\boldsymbol{\Sigma}_{one\_way\_firm}$（(17) 式）の要素が必要であり，［**two way－仮定 2**］より $\boldsymbol{\Sigma}_{one\_way\_time}$ の要素が必要である。[19] そのために両者を加算するが，このままでは主対角要素である分散要素が重複して加算されてしまう。そのため，(18) 式のように $[NT \times NT]$ の White の分散共分散行列 $\boldsymbol{\Sigma}_{White[NT]}$ を減算する必要がある。[20]

$$\boldsymbol{\Sigma}_{two\_way} = \boldsymbol{\Sigma}_{one\_way\_firm} + \boldsymbol{\Sigma}_{one\_way\_time} - \boldsymbol{\Sigma}_{White[NT]} \tag{18}$$

(18) 式の計算過程を $N = 3, T = 3$ として，$[NT \times NT] = [9 \times 9]$ の行列で示すと，下記のようになる。[21]

$$\boldsymbol{\Sigma}_{one\_way\_firm} = \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,12} & \sigma_{1,13} & 0 & 0 & 0 & 0 & 0 & 0 \\ \sigma_{1,21} & \sigma_{1,2}^2 & \sigma_{1,23} & 0 & 0 & 0 & 0 & 0 & 0 \\ \sigma_{1,31} & \sigma_{1,32} & \sigma_{1,3}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{2,1}^2 & \sigma_{2,12} & \sigma_{2,13} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{2,21} & \sigma_{2,2}^2 & \sigma_{2,23} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{2,31} & \sigma_{2,32} & \sigma_{2,3}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{3,1}^2 & \sigma_{3,12} & \sigma_{3,13} \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{3,21} & \sigma_{3,2}^2 & \sigma_{3,23} \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{3,31} & \sigma_{3,32} & \sigma_{3,3}^2 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_{one\_way\_time} = \begin{bmatrix} \sigma_{1,1}^2 & 0 & 0 & \sigma_{12,1} & 0 & 0 & \sigma_{13,1} & 0 & 0 \\ 0 & \sigma_{1,2}^2 & 0 & 0 & \sigma_{12,2} & 0 & 0 & \sigma_{13,2} & 0 \\ 0 & 0 & \sigma_{1,3}^2 & 0 & 0 & \sigma_{12,3} & 0 & 0 & \sigma_{13,3} \\ \sigma_{21,1} & 0 & 0 & \sigma_{2,1}^2 & 0 & 0 & \sigma_{23,1} & 0 & 0 \\ 0 & \sigma_{21,2} & 0 & 0 & \sigma_{2,2}^2 & 0 & 0 & \sigma_{23,2} & 0 \\ 0 & 0 & \sigma_{21,3} & 0 & 0 & \sigma_{2,3}^2 & 0 & 0 & \sigma_{23,3} \\ \sigma_{31,1} & 0 & 0 & \sigma_{32,1} & 0 & 0 & \sigma_{3,1}^2 & 0 & 0 \\ 0 & \sigma_{31,2} & 0 & 0 & \sigma_{32,2} & 0 & 0 & \sigma_{3,2}^2 & 0 \\ 0 & 0 & \sigma_{31,3} & 0 & 0 & \sigma_{32,3} & 0 & 0 & \sigma_{3,3}^2 \end{bmatrix} \tag{19}$$

19　$\boldsymbol{\Sigma}_{one\_way\_time}$ の $[9 \times 9]$ の具体例は，(19) 式である。

20　Pertersen (2009)，p.458

21　$\boldsymbol{\Sigma}_{one\_way\_firm}$ において，例えば 1 行 2 列目の要素は，(16) 式にならって，$\sigma_{(1)1,2}$ と書くべきであるが，記号が煩雑になるので，$\sigma_{1,12}$ と記している（以下同様）。$\boldsymbol{\Sigma}_{one\_way\_ftime}$ において，例えば 1 行 4 列目の要素は，(16) 式にならえば，$\sigma_{1,2(1)}$ と書くべきであるが，記号が煩雑になるので，$\sigma_{12,1}$ と記している（以下同様）。

$$
\boldsymbol{\Sigma}_{White[NT]} = \begin{bmatrix}
\sigma^2_{1,1} & & & & & & & & \\
& \sigma^2_{1,2} & \ddots & & & & & \mathbf{0} & \\
& \ddots & \sigma^2_{1,3} & \ddots & & & & & \\
& & \ddots & \sigma^2_{2,1} & \ddots & & & & \\
& & & \ddots & \sigma^2_{2,2} & \ddots & & & \\
& & & & \ddots & \sigma^2_{2,3} & \ddots & & \\
& & & & & \ddots & \sigma^2_{3,1} & \ddots & \\
& \mathbf{0} & & & & & \ddots & \sigma^2_{3,2} & \\
& & & & & & & & \sigma^2_{3,3}
\end{bmatrix}
$$

$$
\boldsymbol{\Sigma}_{two\_way} = \boldsymbol{\Sigma}_{one\_way\_firm} + \boldsymbol{\Sigma}_{one\_way\_time} - \boldsymbol{\Sigma}_{White[NT]}
$$

$$
= \begin{bmatrix}
\sigma^2_{1,1} & \sigma_{1,12} & \sigma_{1,13} & \sigma_{12,1} & 0 & 0 & \sigma_{13,1} & 0 & 0 \\
\sigma_{1,21} & \sigma^2_{1,2} & \sigma_{1,23} & 0 & \sigma_{12,2} & 0 & 0 & \sigma_{13,2} & 0 \\
\sigma_{1,31} & \sigma_{1,32} & \sigma^2_{1,3} & 0 & 0 & \sigma_{12,3} & 0 & 0 & \sigma_{13,3} \\
\sigma_{21,1} & 0 & 0 & \sigma^2_{2,1} & \sigma_{2,12} & \sigma_{2,13} & \sigma_{23,1} & 0 & 0 \\
0 & \sigma_{21,2} & 0 & \sigma_{2,21} & \sigma^2_{2,2} & \sigma_{2,23} & 0 & \sigma_{23,2} & 0 \\
0 & 0 & \sigma_{21,3} & \sigma_{2,31} & \sigma_{2,32} & \sigma^2_{2,3} & 0 & 0 & \sigma_{23,3} \\
\sigma_{31,1} & 0 & 0 & \sigma_{32,1} & 0 & 0 & \sigma^2_{3,1} & \sigma_{3,12} & \sigma_{3,13} \\
0 & \sigma_{31,2} & 0 & 0 & \sigma_{32,2} & 0 & \sigma_{3,21} & \sigma^2_{3,2} & \sigma_{3,23} \\
0 & 0 & \sigma_{31,3} & 0 & 0 & \sigma_{32,3} & \sigma_{3,31} & \sigma_{3,32} & \sigma^2_{3,3}
\end{bmatrix}
$$

上述のように $\boldsymbol{\Sigma}_{two\_way}$ においては $\boldsymbol{\Sigma}_{one\_way\_firm}$ と比べて 0 でない要素は少ないが，$N \to \infty, T \to \infty$ であれば 0 の要素が十分増えるので，$\hat{\boldsymbol{\beta}}$ は一致推定量となり，［**条件1**］は満たされる。但し（18）式で示すように，$\boldsymbol{\Sigma}_{two\_way}$ は 3 つの分散共分散行列の加算と減算によって作成されているため，$Var(\widehat{\beta_k}) > 0(k = 1, \cdots, K)$ は保証されない。

そこで共分散行列 $\boldsymbol{\Sigma}_{two\_way}$ の正定値性が問題となる。$\boldsymbol{\Sigma}_{two\_way}$ が正定値行列なら[22]，$Var(\widehat{\beta_k}) > 0(k = 1, \cdots, K)$ であることが知られている。ここにおいて，$\boldsymbol{\Sigma}_{two\_way}$ の固有値をがすべて正なら，$\boldsymbol{\Sigma}_{two\_way}$ は正定値行列であり，$Var(\widehat{\beta_k}) > 0(k = 1, \cdots, K)$ となる。問題は，負の固有値があるときである。その場合は，負の固有値を 0 とみなして，新たに $\boldsymbol{\Sigma}'_{two\_way}$ を作るとしている[23]。$\boldsymbol{\Sigma}'_{two\_way}$ は半正定値行列であるので，$Var(\widehat{\beta_k})$

22　実対称行列 A について，「A が正定値行列であること」，「任意の実ベクトル $\mathbf{x}\,(\mathbf{x} \neq \mathbf{0})$ について $\mathbf{x}' A\, \mathbf{x} > 0$ であること」，「A のすべての固有値は正であること」は，論理的に同値である。

23　Cameron, Gelbach, and Miller（2011）, p.241, Politis（2011）, p.715

$\geq 0\ (k = 1, \cdots, K)$ となり，［条件 2］はほぼ満たされることになる。

Ⅵ　ディスカッション

（1）標準回帰モデルの標準誤差

標準回帰モデルの回帰係数の標準誤差は，（11）式から明らかなように，誤差項に企業効果が存在するとき，バイアスをもつ。そして誤差項に時間効果が存在するときも，同様のことが言える。従って，誤差項に相関があるときは，Petersen（2009）の指摘どおり，標準回帰モデルの回帰係数の標準誤差の利用は適切ではない[24]。

（2）White の標準誤差

White の標準誤差は，もともとクロスセクション回帰のために考案された手法であるが，パネル回帰モデルにも適用可能である。但し（13）式からわかるように，$\boldsymbol{\Sigma}_{White}$ の非対角要素（共分散要素）がすべて 0 であるため，Petersen（2009）の指摘どおり，誤差項に相関のあるとき White の標準誤差の利用は適切ではない[25]。

（3）Fama-MacBeth の標準誤差

Fama-MacBeth 回帰は，パネルデータにおいてまず期間ごとのクロスセクション回帰を行い，回帰係数の平均値をパネル回帰の回帰係数として推定する方法である。Petersen（2009）は，時間効果のみ存在し企業効果がないとき，Fama-MacBeth 回帰を推奨している[26]。

しかし，IV 節（2）で示したように，企業効果がある状況においては，Fama-MacBeth 回帰の標準誤差は，標準回帰モデルの回帰係数の標準誤差と同様にバイアスをもつことになるので，Petersen（2009）の指摘どおり，Fama-MacBeth の標準誤差の利用は適切ではない[27]。

（4）Newey-West の標準誤差と one way cluster の標準誤差の比較

Newey-West の手法はもともと時系列回帰のために考案されたが，（15）式でも示したように，企業効果の存在を前提とするパネル回帰に適用できるように拡張されている。one way cluster の標準誤差（17）式も同様に，企業効果の存在を前提とするパネル回帰の手法として考案されている。

24　Pertersen（2009）, p.475
25　Pertersen（2009）, p.475
26　Pertersen（2009）, p.475
27　Pertersen（2009）, p.475

（15）式と（17）式を比較すればわかるように，Newey-West の分散共分散行列と one way cluster の分散共分散行列は類似している。相違点は，Newey-West の分散共分散行列の小行列 $\sum_{NW\_i}$ において，ウェイトづけと両端捨てが行われていることである。これらは時系列回帰において，回帰係数の一致性と回帰係数の分散の非負性を確保するために行った恣意的な操作である。一方，one way cluster の標準誤差を用いればこのような恣意的な操作は不要であるので，one way cluster の分散共分散行列は Newey-West の分散共分散行列より優れていると考えられる。

Petersen（2009）も，同一企業内ではすべての期間の誤差項に時系列相関がある場合（企業効果のある場合），企業による one way cluster の標準誤差の利用を，同一期間内ではすべての企業の誤差項に相関がある場合（時間効果のある場合）で十分なクラスター数があるとき[28]，期間による one way cluster の標準誤差の利用を推奨している[29]。

（5）two way cluster の標準誤差

two way cluster の分散共分散行列は，同一企業内の企業効果，同一期間内の時間効果を同時に想定しているという特徴がある。Petersen（2009）は，このように企業効果と時間効果の両方が存在する場合で，2 つのディメンジョンで十分なクラスター数があるとき[30]，two way cluster の分散共分散行列の利用を推奨している[31]。

しかし two way cluster の分散共分散行列においては，V 節（4）で述べたように，two way cluster の回帰係数の分散の非負性を保証するために，分散共分散行列の固有値が負の値であった場合，それらを 0 とみなして分散共分散行列が作り直される。これはかなり恣意的な操作と考えられるが，Petersen（2009）はこの問題について検討していない。この恣意的な操作は，分析結果に大きな影響を与えると考えられるので，two way cluster の分散共分散行列は，その固有値がすべて正の値のときのみ用いるべきであろう。

（6）提言

ここでファイナンス研究のデータと会計学研究のデータについて考えることにしたい。例えば，日本の上場企業をサンプルとするとき，両者のデータセットにおいて企業数は同じである。しかし，株価などのようなファイナンス・データは，月次，週次，日次，時間単位と細分化できるのに対して，会計データの期間は，一般に年度，半期，四半期までしか細分化できない。two way cluster の分散共分散行列において，あまり小さ

28　この場合のクラスターは期間数であり，クラスター・サイズではないことに注意してほしい。
29　Pertersen（2009），p.475
30　クラスター・サイズではないことに注意してほしい。
31　Pertersen（2009），p.475-476

なサイズのクラスターを作成すると，推定値の信頼性が低下すると考えられる。そのため，会計学の実証研究においては，期間でクラスタリングを行う one way cluster の標準誤差を利用し，時間効果については期間ダミー変数によってコントロールするのが適切であると考えられる。

Ⅶ　む　す　び

「誤差項に相関がある場合，どのような標準誤差を用いるべきか」は，会計学，ファイナンスの実証研究において，重要な問題である。Petersen（2009）の発表後，会計学の実証研究においても，クラスター・ロバスト標準誤差の利用が一般的となっている。しかし，会計学の実証研究において，クラスター・ロバスト標準誤差の統計的性質を理解せずにそれを用いること，ファイナンス研究である Petersen（2009）の主張を無批判に受け入れることは，バイアスのある検証結果を導くリスクがある。

そのような問題意識に基づいて，本論文では誤差項に相関がある場合の適切な標準誤差について検討した。その結果，企業数は十分であるが，期間数が少ない会計データにおいては，期間でクラスタリングを行う one way cluster の標準誤差を利用し，時間効果については期間ダミー変数によってコントロールするのが適切であると考えられる。

参考文献

［1］Amiraslani, H., J. Donovan, M. A. Phillips, and R. Wittenberg-Moerman（2023）, Contracting in the Dark: The Rise of Public-side Lenders in the Syndicated Loan Market, *Journal of Accounting and Economics* 76(1), pp.1-24.

［2］Bertrand, M., E. Duflo, and S. Mullainathan（2004）, How Much Should We Trust Differences-in-Differences Estimates?, *Quarterly Journal of Economics* 119(1), pp.249-275.

［3］Brockman, P., and D. Y. Chung（2001）, Managerial Timing and Corporate Liquidity: Evidence from Actual Share Repurchases, *Journal of Financial Economics* 61, pp.417-448.

［4］Cameron, A. C., J. B. Gelbach, and D. L. Miller（2011）, Robust Inference with Multiway Clustering, *Journal of Business and Economic Statistics,* 29(2), pp.238-249.

［5］Choi, B. G., J. H. Choi, and S. Malik（2023）, Not Just for Investors: The Role of Earnings Announcements in Guiding Job Seekers, *Journal of Accounting and Economics* 76(1), pp.1-23.

［6］Colonnello, S., M. Koetter, and K. Wagner（2023）, Compensation Regulation in Banking: Executive Director Behavior and Bank Performance after the EU Bonus Cap, *Journal of Accounting and Economics* 76(1), pp.1-22.

［7］deHaan, E., J. Li, E. M. Watts（2023）, Retail Bond Investors and Credit Ratings, *Journal of Accounting and Economics* 76(1), pp.1-29.

［8］Doidge, C.（2004）, U.S. Cross-Listings and the Private Benefits of Control: Evidence from Dual-Class Firms, *Journal of Financial Economics,* 72, pp.519-553.

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

［10］Goldman, N. C. and N. B. Ozel（2023）, Executive Compensation, Individual-level Tax Rates, and Insider Trading Profits, *Journal of Accounting and Economics* 76（1）, pp.1-21.

［11］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）, pp.483-512.

［12］Flannery, M. J. and K. W. Hankins（2013）, Estimating Dynamic Panel Models in Corporate Finance, *Journal of Corporate Finance* 19, pp.1-19.

［13］Kim, J. and K. Valentine（2023）, Public Firm Disclosures and the Market for Innovation, *Journal of Accounting and Economics* 76（1）, pp.1-23.

［14］Lennox, C., C. Wang, and X. Wu（2023）, Delegated Leadership at Public Accounting Firms, *Journal of Accounting and Economics* 76（1）, pp.1-23.

［15］MacKay, P.（2003）, Real Flexibility and Financial Structure: An Empirical Analysis, *Review of Financial Studies* 16（4）, pp.1131-1165.

［16］Newey, W. K. and K. D. West（1987）, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica,* 55（3）, pp.703-708.

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

［18］Politis, D. N.（2011）, Higher-Order Accurate, Positive Semidefinite Estimation of Large-Sample Covariance and Spectral Density Matrices, *Econometric Theory* 27, pp.703-744.

［19］White, H.（1980）, A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica,* 48（4）, pp.817-838.

［20］Wooldridge, J. M.（2020）, *Introductory Econometrics: A Modern Approach,* Cengage.

［21］浅野晳・中村二朗（2009）『計量経済学（第２版）』有斐閣.

［22］太田浩司（2012）「White, Newey-West, Cluster-robust, Fama-Macbeth の標準誤差の理論と応用」日本会計研究学会第 71 回大会報告論文，pp.1-41.

［23］太田浩司（2013）「パネル・データ分析におけるクラスター頑健手法の使用について」『証券アナリストジャーナル』第 51 巻第 11 号，pp.77-87.

［24］太田浩司（2017）「パネル分析における Fama-MacBeth と Cluster-robust の手法の理論と応用」『関西大学商学論集』第 62 巻第 2 号，pp.43-67.

［25］谷崎久志・溝渕健一（2023）『計量経済学』新世社.

［26］森田果（2014）『実証分析入門－データから「因果関係」を読み解く作法』日本評論社.

［27］山本拓（2022）『計量経済学（第２版）』新世社.