統計分析再入門

（第５回）

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

太　田　浩　司

目　　　　次

１．はじめに
２．固定効果モデルとランダム効果モデル
３．パネル分析におけるクラスター頑健手法の使用パターン
４．クラスター頑健手法
５．パネル分析におけるクラスター頑健手法使用に関する注意
６．おわりに

本稿では、パネル回帰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、近年の会計・ファイナンスの実証研究で頻繁に用いられている、クラスター頑健手法の理論的背景と、それを使用する際の注意点について述べている。クラスター頑健手法による標準誤差は、最新の統計ソフトを用いれば容易に求められるが、その概要を理解して適切に使用することが、正しい統計的推定ならびに結果の解釈につながると思われる。

１．はじめに

会計･ファイナンス分野における実証研究では、企業と年度のパネル・データを使用した分析が行われることが多い。その際に、推定モデルの誤差項が標準的仮定（均一分散で互いに無相関）を満たしていない場合には、係数推定値の分散が正しく求められず、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。したがって、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、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、近年注目を集めているのが、クラスター頑健手法（cluster-robust method）である。

クラスター頑健手法は、Petersen［2009］等で統計ソフトのプログラムが公開されたことにより、近年その使用が急速に拡大しているが、その理論的背景についてはまだあまり知られていない。そこで、本稿では、クラスター頑健手法の概要について叙述し、その使用に関する注意点を指摘している。

太田　浩司（おおた　こうじ）
関西大学商学部教授。1994年京都大学文学部卒業。2003年関西大学大学院商学研究科博士後期課程単位取得。07年筑波大学大学院ビジネス科学研究科博士後期課程修了。博士(経営学、筑波大学)。㈱青木建設、関西CPA学院、武蔵大学経済学部等を経て、12年４月より現職。

なお、本稿の構成は次のようである。パネル分析では、クラスター頑健手法は、固定効果モデル、ランダム効果モデルあるいはPooled回帰モデルと併せて用いられるが、係数パラメータの推定自体は、固定効果モデル、ランダム効果モデルあるいはPooled回帰モデルで行われ、クラスター頑健手法は、あくまで係数推定値の分散をより適切に求めるために用いられる。そこで、第２章では、パネル推定における代表的手法である、固定効果モデルとランダム効果モデルの特徴について述べ、第３章で、クラスター頑健手法がこれらの手法と併せてどのような状況で用いられるのかを示す。その後、第４章で、クラスター頑健手法の理論的背景について論じ、第５章で、パネル分析におけるクラスター頑健手法の使用に関する注意点を指摘する。最後に第６章で、本稿を総括する。

２．固定効果モデルとランダム効果モデル

2.1.　固定効果モデルとランダム効果モデルの概要

最初に、i（$1, 2, \dots, N$）は個別企業、t（$1, 2, \dots, T$）は年度を表すものとして、企業と年度のパネル・データを用いた以下のような回帰モデルについて考える。

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + c_i + \varepsilon_{it} \tag{1}$$

ただし、y_{it}は被説明変数、$\mathbf{x}_{it}$は説明変数、c_iは観察不可能な各企業固有の効果、ε_{it}は通常の仮定を満たす誤差項である。

パネル回帰で注意しなければいけないのが、観察は不可能ではあるけれども各企業に固有に存在していると思われるc_iの取り扱いである。例えば、化粧品業界に属する企業の利益率（y_{it}）に与える要因として、研究開発費率と設備投資比率（$\mathbf{x}_{it}$）の影響を調査したいとする。このとき、各企業の利益率に与える他の要因として、観察は不可能であるけれども全ての年度において存在する、各企業固有のブランド力や独自の販売チャネルというようなものの影響（c_i）があると考えるのは自然である。つまりc_iは、データとして入手不可能ではあるが､各企業が有すると思われる時不変（time invariant）な効果であり、企業と年度のパネル・データを用いた分析では、c_iの存在を仮定することが理にかなうケースがままあるのである。

このとき、説明変数$\mathbf{x}_{it}$とc_iの間に相関がある場合に（$\mathrm{Cov}[\mathbf{x}_{it}, c_i] \neq \mathbf{0}$）（注１）、(1)の$c_i$を無視して以下の(2)を推定してしまうと、誤差項$u_{it}$に$c_i$が吸収されてしまう。

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + u_{it}, \quad u_{it} = c_i + \varepsilon_{it} \tag{2}$$

そしてこの場合には、説明変数$\mathbf{x}_{it}$と誤差項u_{it}の間に相関が生じることになるので（$\mathrm{Cov}[\mathbf{x}_{it}, u_{it}] \neq \mathbf{0}$）、内生性と呼ばれる深刻な問題が発生してしまい、パラメータ$\boldsymbol{\beta}$の推定量$\mathbf{b}$は、不偏性も一致性も有さなくなってしまう。このような、観察は不可能であるけれども、説明変数と相関があると思われる変数を無視して推定してしまうことから生じる内生性の問題は、しばしば、除外変数バイアス（omitted variables bias）と呼ばれる。

そこで、説明変数$\mathbf{x}_{it}$と観察不可能な各企業固有の効果c_iの間に相関があると思われる場合には、c_iをモデルに含む(1)を推定する必要があり、そのために用いられるのが固定効果モデルと呼ばれる

（注１）説明変数が複数個ある場合には、その内の一つの説明変数とでもc_iの間に相関があれば、除外変数バイアスが生じることになる。

図表１　固定効果モデルとランダム効果モデルの長所と短所

	固定効果モデル	ランダム効果モデル
①説明変数$\mathbf{x}_{it}$と観察不可能な各企業固有の効果c_iの間の相関。	長所：相関があっても良い。	短所：相関があってはならない。
②所属産業や上場市場等の年度によって変化しない各企業の特性を、ダミー変数（$dummy_i$）として説明変数に加えたい。	短所：説明変数に加えることができない。	長所：説明変数に加えることができる。
③モデルの定数項の符号や大きさを推定したい。	短所：定数項がN個存在するので特定できない。	長所：定数項は１つなので特定可能である。
④パラメータ$\boldsymbol{\beta}$推定のための自由度。	短所：自由度がN個減少する。	長所：自由度は減らない。

手法である。

一方、説明変数$\mathbf{x}_{it}$とc_iの間に相関がない場合には（$\mathrm{Cov}[\mathbf{x}_{it}, c_i] = \mathbf{0}$）、(1)の$c_i$を無視して(2)を推定しても、説明変数$\mathbf{x}_{it}$と誤差項$u_{it}$の間に相関が生じないので（$\mathrm{Cov}[\mathbf{x}_{it}, u_{it}] = \mathbf{0}$）、内生性の問題は生じず、$\boldsymbol{\beta}$の推定量$\mathbf{b}$は不偏性および一致性を有している。ただし、同一企業の異なる時点の誤差項u_{it}，u_{is}には、共通の効果c_iが含まれているので、誤差項の標準的仮定（均一分散で互いに無相関）は満たされず、$\mathbf{b}$は有効推定量とはならない。そこで、ランダム効果モデルと呼ばれる一般化最小二乗法（GLS：Generalized Least Squares）の一手法を用いて、誤差項が標準的仮定を満たすように変数変換を行った後に、OLS推定を行うのである。

2.2.　固定効果モデルとランダム効果モデルの長所と短所

固定効果モデルとランダム効果モデルの長所と短所は、ちょうど表と裏のような関係で、固定効果モデルにとっての長所（短所）はランダム効果モデルにとっては短所（長所）となっている。それをまとめたものが、**図表１**である。

第１に、固定効果モデルでは、説明変数$\mathbf{x}_{it}$と観察不可能な各企業固有の効果c_iの間に相関があっても良いが、ランダム効果モデルでは、両者の間に相関があってはならない。第２に、固定効果モデルでは、所属産業や上場市場等の年度によって変化しない個々の企業の特性を、ダミー変数（$dummy_i$）として説明変数$\mathbf{x}_{it}$に含めることができないが、ランダム効果モデルでは、これらのダミー変数を説明変数に含めることができる。

第３に、定数項を含む理論モデルの妥当性を実証的に検証する目的でパネル推定を行うような場合には、定数項の符号や大きさを特定したい場合がある。このような場合に、固定効果モデルでは、定数項が企業数のN個推定されてしまうので定数項を特定できないが、ランダム効果モデルでは、定数項が１つに特定されるので理論モデルとの整合性が検証可能となる。

第４に、固定効果モデルでは、調査対象のパラメータ$\boldsymbol{\beta}$推定のための自由度が企業数であるN個減少してしまうので、年度数Tが小さいような場合には、パラメータの推定精度が大きく低下してしまう恐れがある。一方、ランダム効果モデルでは、パラメータ推定のための自由度は低下しない。

以上、固定効果モデルとランダム効果モデルの長所と短所の比較からは、１番目の、説明変数$\mathbf{x}_{it}$と観察不可能な各企業固有の効果c_iの間に相関があってはならないという短所以外は、ランダム効果モデルの方が固定効果モデルよりも長所は多いといえる。しかしながら、会計・ファイナンスの研究でしばしば用いられる、企業と年度のパネル・データを使用する分析では、この１番目の、$\mathbf{x}_{it}$とc_iの間に相関があってはならないという条件が満

たされないことが多く、それ故に、固定効果モデルが用いられることが多いのである。

例えば、(1)の場合では、c_iは各企業固有のブランド力のようなものであると想定でき、ブランド力のある企業ほど、それを維持するために積極的に研究開発を行ったり、あるいは、最新の設備を導入するのではないかと考えられるので、説明変数$\mathbf{x}_{it}$とc_iの間には相関があることが予想される。したがって、ランダム効果モデルではなく、固定効果モデルを用いて推定を行うことが適切であると思われるのである（注2）。

3．パネル分析におけるクラスター頑健手法の使用パターン

前章では、パネル分析における代表的な推定方法として、固定効果モデルとランダム効果モデルについて概説した。最初に、固定効果モデルは、説明変数$\mathbf{x}_{it}$と、観察は不可能であるけれども各企業が有していると思われる時不変の効果c_iの間に相関があると思われる場合に用いられる手法である。しかしながら、固定効果モデルを用いてc_iの影響を取り除いたとしても、誤差項が均一分散で互いに無相関であるという標準的な仮定を満たしているという保証はない。

例えば、(1)の説明変数$\mathbf{x}_{it}$の１つである研究開発費率は、その比率が高いほど利益率の高い新製品の導入を可能にするので、企業の利益率y_{it}とは正の相関が予想される。しかしながら、研究開発は必ずしも成功するとは限らず、失敗した場合には逆に企業の利益を圧迫する要因に成りかねない。したがって、研究開発費率が高くなるほど企業の利益率のバラつきも大きくなり、誤差項の分散も増加するのではないかと考えられるのである。

このような場合には、たとえ固定効果モデルでc_iの影響を取り除いたとしても、誤差項には不均一分散が発生してしまい、誤差項の標準的な仮定は満たされないことになる。

次に、ランダム効果モデルは、説明変数$\mathbf{x}_{it}$と、観察不可能な各企業の効果c_iの間に相関がないと思われる場合に用いられる手法である。ランダム効果モデルでは、誤差項にc_iが吸収されてしまうことから生じる誤差項間の相関に対処するために、誤差項の共分散構造をある特定の形であると仮定して、GLSで推定を行っている。

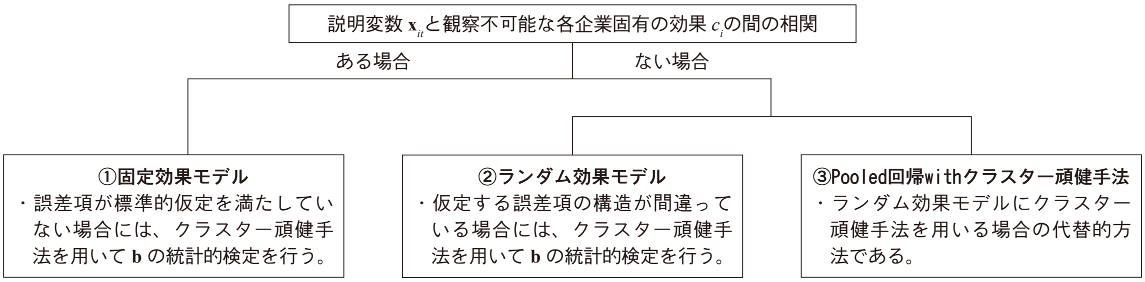
このとき、もし誤差項の真の構造が、ランダム効果モデルが想定する特定の形であった場合には、$\boldsymbol{\beta}$の推定量$\mathbf{b}$は最良線形不偏推定量（BLUE）になるという優れた長所がある。その一方で、もし誤差項の真の構造が、ランダム効果モデルが想定する特定の形でなかった場合には、ランダム効果推定から得られる$\mathbf{b}$の有効性がかえって低下してしまうという危険性がある（注3）。

例えば、ランダム効果モデルでは、誤差項

（注２）固定効果モデルとランダム効果モデルの選択には、Hausman検定が用いられることが多いが、Hausman検定は厳しい制約条件の下でのみ成立する検定であるので、その実際的有用性には疑問が残る。それよりも、両モデルの選択は、状況が許すのであれば基本的に固定効果モデルを使用することが好ましいと思われる。なぜなら、固定効果モデルでは、説明変数と観察不可能な各企業固有の効果の間に相関がある場合のみならず、相関がない場合にも使用可能であるが、ランダム効果モデルでは、相関がない場合にしか使用できないからである。しかしながら、現実には、**図表１**②③④の事情により固定効果モデルが使用不可能な場合も多く、そのような場合には、ランダム効果モデルを用いることによって、何らかの考察を得ざるをえないのである。

（注３）たとえ、誤差項の真の構造がランダム効果モデルが想定する特定の形でなかったとしても、$\mathbf{b}$の一致性は通常担保されている。

図表2　パネル分析におけるクラスター頑健手法の使用パターン



u_{it}（$u_{it}=c_i+\varepsilon_{it}$）の分散は全て同一である（$\sigma_c^2+\sigma_\varepsilon^2$）と仮定しているが、上述のように、研究開発費率が高くなるほど企業の利益率のバラつきも大きくなるような場合には、誤差項は不均一分散となると考えられる。

また、誤差項の不均一分散に限らず、系列相関に関しても、固定効果モデルでは系列相関はないと仮定されており、ランダム効果モデルでは企業内の系列相関は全て等しい（σ_c^2）と仮定しているが、企業と年度のパネル・データ分析では、これらの系列相関の仮定が満たされないことが多い。

このように、(i)固定効果モデルを用いても誤差項が標準的な仮定を満たしていないと思われる場合や、(ii)ランダム効果モデルの仮定する誤差項の構造が正しくないと思われる場合に用いられるのが、クラスター頑健手法（cluster-robust method）である。

クラスター頑健手法の使い方としては、固定効果モデルに関しては、パラメータ**β**の推定に関しては固定効果モデルで行い、得られた推定量**b**の統計的検定を行う際にはクラスター頑健手法で得られた標準誤差を用いるという手順である。

一方、ランダム効果モデルに関しても、**β**の推定はランダム効果モデルで行い、**b**の統計的検定にはクラスター頑健手法で得られた標準誤差を用いるという方法が可能である。しかしながら、この場合には、ランダム効果モデルが本来想定する誤差項の共分散構造の形が間違っていたということになる。そして、誤差項の共分散構造が誤って識別されたランダム効果モデルにクラスター頑健手法を用いる場合と、通常のPooled回帰にクラスター頑健手法を用いる場合の有効性の優劣については定かではないので、その両方を用いることが考えられる。

なお、**図表2**は、パネル分析におけるクラスター頑健手法の使用パターンを図示したものである。

4．クラスター頑健手法

4.1．クラスター頑健手法とは

本章では、クラスター頑健手法の概要について述べるが、注意しなければいけないのが、クラスター頑健手法はあくまで係数推定値の有意性検定をより適切に行うために用いる手法であって、係数推定値自体に影響を与えるものではないということである。前章で述べたように、パネル・データを用いる分析では、クラスター頑健手法は、固定効果モデル、ランダム効果モデルおよびPooled回帰と併せて用いられる。その際、係数の推定に関してはそれぞれの方法で行うのだが、その誤差項が標準的仮定（均一分散で互いに無相関）を満たしていない場合には、係数の標準誤差が正しく求められない。したがって、係数推定値の有意性

図表３　誤差項の共分散構造（不均一分散と企業のクラスタリング）

(i)　不均一分散Σ_{hetro}

		firm a			firm b		
		t=1	t=2	t=3	t=1	t=2	t=3
firm a	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3
	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3
firm b	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3
	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2

(ii)　企業のクラスタリングΣ_{firm}

		firm a			firm b		
		t=1	t=2	t=3	t=1	t=2	t=3
firm a	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3
	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3
firm b	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3
	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2

（図表注）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。

検定が適切に行われているという保証がなく、本来は有意ではない係数推定値を誤って有意であると判断してしまい、結果として間違った結論を導いてしまう可能性があるのである。

クラスター頑健手法は、固定効果モデル、ランダム効果モデルおよびPooled回帰の誤差項に不均一分散や系列相関が存在する場合に、それらに対して頑健な標準誤差を求める方法であり、それによって、係数推定値の有意性検定がより適切に行えるのである。

4.2.　クラスター頑健手法の概要

最初に、観測値数がn個の回帰モデルは、$\mathbf{y}$を被説明変数ベクトル、$\mathbf{X}$を説明変数行列、$\boldsymbol{\varepsilon}$を誤差項ベクトル、$\boldsymbol{\beta}$を回帰係数ベクトルとすると、

$$\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$$

と表される。このとき、パラメータ$\boldsymbol{\beta}$の推定量$\mathbf{b}$の分散推定量は、

$$\mathrm{Var}[\mathbf{b}]=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$\text{ただし、}\quad \boldsymbol{\Sigma}=E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']=\begin{pmatrix}\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n}\\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2\end{pmatrix}$$

と表せる。

$\mathbf{X}$はデータとして与えられるものであるので、上式からは、$\mathbf{b}$の分散が、誤差項の共分散行列$\boldsymbol{\Sigma}$がどのような構造を持っていると仮定するかによって決定されることがわかる。

ここで、簡単化のために、個別企業数２社（$i=a$, b）で年度数３年（t=1，2，3）のパネル・データについて考えることとする。なお、a社の１、２、３年目の誤差項をそれぞれa_1，a_2，a_3，b社の１、２、３年目の誤差項をそれぞれb_1，b_2，b_3で表すこととし、期待オペレーターE[ ] は省略することとする。

最初に、誤差項が不均一分散をもつ場合の誤差項の共分散行列$\boldsymbol{\Sigma}_{\text{hetro}}$は、**図表３**(i)のような構造を有している。なお、図表の見方としては、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（すなわちゼロ）とみなしている。したがって、**図表３**(i)の不均一分散のケースでは、対角要素のa_1^2，a_2^2，a_3^2，b_1^2，b_2^2，b_3^2にだけ相関があって、それ以外の非対角要素は全てゼロと仮定しているのである。ちなみに、対角要素が全て等しいと仮定した場合（$a_1^2=a_2^2=a_3^2=b_1^2=b_2^2=b_3^2$）には、均一分散で系列相関がないという誤差項の標準的仮定となる。

ここで、$\mathbf{b}$が$\boldsymbol{\beta}$の一致推定量であることを利用してSlutskyの定理を用いれば、残差$\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$,

$\hat{b}_1$, $\hat{b}_2$, $\hat{b}_3$は、それぞれ、誤差項a_1, a_2, a_3, b_1, b_2, b_3の一致推定量であるといえる。そこで、**図表3**(i)の対角要素を、それぞれ、$\hat{a}_1^2$, $\hat{a}_2^2$, $\hat{a}_3^2$, $\hat{b}_1^2$, $\hat{b}_2^2$, $\hat{b}_3^2$で置き換えた、$\hat{\boldsymbol{\Sigma}}_{\mathrm{White}}$を用いると、Whiteの不均一分散に対して頑健な漸近分散推定量が得られる。

$$\mathrm{Est.Asy.Var}[\mathbf{b}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\Sigma}}_{\mathrm{White}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{3}$$

次に、誤差項が企業でクラスタリングしている場合の誤差項の共分散行列$\boldsymbol{\Sigma}_{\mathrm{firm}}$は、**図表3**(ii)のような構造を有しており、先と同様に、灰色のエリアにだけ相関があると仮定している。図からもわかるように、企業のクラスタリングでは、同一企業の誤差項間に相関があると仮定されている。例えば、a社に関しては、その誤差項a_1, a_2, a_3は互いに相関しており、分散も不均一であるという仮定がなされており、b社に関しても同じ仮定がなされている。つまり、同一企業内の誤差項の構造はどのような形でも良く、これが企業に関する誤差項のクラスタリングと呼ばれるゆえんである。ただし、企業のクラスタリングでは、異なる企業間の誤差項、すなわちa社とb社の誤差項には相関がないと仮定されていることには注意が必要である。

ここで、先のWhiteの不均一分散に対して頑健な漸近分散推定量のときと同様に、灰色のエリアの誤差項a_1, a_2, a_3, b_1, b_2, b_3を、その一致推定量である残差$\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$, $\hat{b}_1$, $\hat{b}_2$, $\hat{b}_3$で置き換えた、$\hat{\boldsymbol{\Sigma}}_{\mathrm{CR(firm)}}$を用いると、企業のクラスタリングに対して頑健な漸近分散推定量が得られる。

$$\mathrm{Est.Asy.Var}[\mathbf{b}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\boldsymbol{\Sigma}}_{\mathrm{CR(firm)}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{4}$$

クラスター頑健手法では、この分散推定量から得られる標準誤差を用いて、係数推定値の有意性検定を行うのである。

4.3. Two-wayクラスター頑健手法

第2章では、企業と年度のパネル・データを用いた回帰モデル(1)について述べたが、それに、観察不可能な各年度固有の効果γ_tを加えた以下のような回帰モデルも考えられる。

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

ただし、γ_tは、観察は不可能であるけれども、全ての企業が影響を受けると思われる各年度固有の効果で、例えば、リーマン・ショックやアベノミクスのようなマクロ経済的なインパクトが考えられる。

γ_tもc_iの場合と同様に、説明変数$\mathbf{x}_{it}$との間に相関があると思われる場合には固定効果モデル、相関がないと考えられる場合にはランダム効果モデルを用いて推定を行う。ただし、その場合には、c_iとγ_tの両方の効果を考慮して推定を行う必要があるので、それぞれ、Two-way固定効果モデル、Two-wayランダム効果モデルと呼ばれる手法が用いられる。これと同様の発想で、クラスター頑健手法に関しても、企業と年度の両方に関して誤差項がクラスタリングしている場合に用いられる、Two-wayクラスター頑健手法が存在する。

最初に、**図表3**(ii)では、誤差項が企業でクラスタリングしている場合の誤差項の共分散行列$\boldsymbol{\Sigma}_{\mathrm{firm}}$の構造を示したが、**図表4**(i)は、誤差項が年度でクラスタリングしている場合の誤差項の共分散行列$\boldsymbol{\Sigma}_{\mathrm{year}}$の構造を表している。年度のクラスタリングでは、同一年度の誤差項間に相関があると仮定されている。例えば、年度1に関しては、その誤差項a_1, b_1は互いに相関していて分散も異なるという仮定がなされており、年度2に関してはa_2, b_2、年度3に関してはa_3, b_3に同じ仮定が置かれている。つまり、同一年度内の誤差項の構造はどのような形でも良いというのが、年度に関する誤

図表4　誤差項の共分散構造（年度のクラスタリングとTwo-wayクラスタリング）

（i）年度のクラスタリング$\mathbf{\Sigma}_{\text{year}}$

		firm a			firm b		
		t=1	t=2	t=3	t=1	t=2	t=3
firm a	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3
	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3
firm b	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3
	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2

（ii）Two-wayクラスタリング$\mathbf{\Sigma}_{\text{Two-way}}$

		firm a			firm b		
		t=1	t=2	t=3	t=1	t=2	t=3
firm a	t=1	a_1^2	a_1a_2	a_1a_3	a_1b_1	a_1b_2	a_1b_3
	t=2	a_2a_1	a_2^2	a_2a_3	a_2b_1	a_2b_2	a_2b_3
	t=3	a_3a_1	a_3a_2	a_3^2	a_3b_1	a_3b_2	a_3b_3
firm b	t=1	b_1a_1	b_1a_2	b_1a_3	b_1^2	b_1b_2	b_1b_3
	t=2	b_2a_1	b_2a_2	b_2a_3	b_2b_1	b_2^2	b_2b_3
	t=3	b_3a_1	b_3a_2	b_3a_3	b_3b_1	b_3b_2	b_3^2

（図表注）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、$\mathbf{\Sigma}_{\text{Two-way}}=\mathbf{\Sigma}_{\text{firm}}+\mathbf{\Sigma}_{\text{year}}-\mathbf{\Sigma}_{\text{hetro}}$という関係が存在している。

差項のクラスタリングである。ただし、年度のクラスタリングでは、異なる年度間の誤差項、すなわち年度を表す下添字１、２、３が異なる誤差項間には相関がないという仮定がなされていることには注意が必要である。

ここで、灰色のエリアの誤差項a_1，a_2，a_3，b_1，b_2，b_3を、その一致推定量である残差$\hat{a}_1$，$\hat{a}_2$，$\hat{a}_3$，$\hat{b}_1$，$\hat{b}_2$，$\hat{b}_3$で置き換えた、$\hat{\mathbf{\Sigma}}_{\text{CR(year)}}$を用いると、年度のクラスタリングに対して頑健な漸近分散推定量が得られる。

$$\text{Est.Asy.Var}[\mathbf{b}]=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{\Sigma}}_{\text{CR(year)}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{5}$$

次に、**図表4**(ii)は、企業と年度の両方に関して誤差項がクラスタリングしている、Two-wayクラスタリングの誤差項の共分散行列$\mathbf{\Sigma}_{\text{Two-way}}$の構造を示したものである。図表からもわかるように、同一企業および同一年度内の誤差項の構造はどのような形でも良いというのが、Two-wayクラスタリングの特徴である。

Two-wayクラスタリングの共分散行列$\mathbf{\Sigma}_{\text{Two-way}}$は、企業と年度の両方で誤差項がクラスタリングしているので、基本的には、**図表3**(ii)の企業のクラスタリングの共分散行列$\mathbf{\Sigma}_{\text{firm}}$と、**図表4**(i)の年度のクラスタリングの共分散行列$\mathbf{\Sigma}_{\text{year}}$を足し合わせれば形成できる。ただし、この場合には対角要素（a_1^2，a_2^2，a_3^2，b_1^2，b_2^2，b_3^2）が重複してしまうので、**図表3**(i)の不均一分散の共分散行列$\mathbf{\Sigma}_{\text{hetro}}$を差し引いてやる必要がある。つまり、$\mathbf{\Sigma}_{\text{Two-way}}=\mathbf{\Sigma}_{\text{firm}}+\mathbf{\Sigma}_{\text{year}}-\mathbf{\Sigma}_{\text{hetro}}$という関係が存在するのである。

ここで、Two-wayクラスタリングの灰色のエリアの誤差項a_1，a_2，a_3，b_1，b_2，b_3を、その一致推定量である残差$\hat{a}_1$，$\hat{a}_2$，$\hat{a}_3$，$\hat{b}_1$，$\hat{b}_2$，$\hat{b}_3$で置き換えた、$\hat{\mathbf{\Sigma}}_{\text{CR(Two-way)}}$を用いると、企業と年度の両方のクラスタリングに対して頑健な漸近分散推定量が得られる。

$$\text{Est.Asy.Var}[\mathbf{b}]=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{\Sigma}}_{\text{CR(Two-way)}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{6}$$

ただし、$\hat{\mathbf{\Sigma}}_{\text{CR(Two-way)}}=\hat{\mathbf{\Sigma}}_{\text{CR(firm)}}+\hat{\mathbf{\Sigma}}_{\text{CR(year)}}-\hat{\mathbf{\Sigma}}_{\text{White}}$

Two-wayクラスター頑健手法では、この分散推定量から得られる標準誤差を用いて、係数推定値の有意性検定を行うのである。

5．パネル分析におけるクラスター頑健手法使用に関する注意

本章では、パネル分析にクラスター頑健手法を用いる際の幾つかの注意点について概説する。第１に、クラスター数が少ない場合のクラスター頑

健手法の標準誤差には、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ことが知られている（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クラスター頑健手法を用いると、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、研究者間で意見にバラつきがみられるが、クラスター数がおおよそ20～30個あれば、一般に問題なくクラスター頑健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り、そのための手法が各種提案されている（Cameron *et al.*［2008］，Angrist and Pischke［2009］，Cameron and Miller［2010］，Cameron *et al.*［2011］）。

第２に、(6)のTwo-wayクラスター頑健手法の漸近分散推定量には、正定値性が保証されていないという問題である。つまり、Two-wayクラスター頑健手法を用いて求められた分散は負になってしまう可能性があるのである（注４）。

その対処法としては、(6)をスペクトラル分解して負の固有値はゼロと置き換えるというような方法が提案されてはいるが（Cameron *et al.*［2011］）、非常にアドホックであり、根本的な解決策はまだ見つかっていない。

第３に、会計・ファイナンス分野における企業と年度のパネル・データは、通常、企業数（$i=1, 2, \ldots, N$）に比して年度数（$t=1, 2, \ldots, T$）が非常に少ない、ショート・パネル・データであることが多いということに関する注意点である。

今、(3)～(6)で表される分散推定量を、それぞれ$\mathrm{Var}[\mathbf{b}_{\mathrm{White}}]$，$\mathrm{Var}[\mathbf{b}_{\mathrm{firm}}]$，$\mathrm{Var}[\mathbf{b}_{\mathrm{year}}]$，$\mathrm{Var}[\mathbf{b}_{\text{Two-way}}]$と表すとすると、これらの分散推定量の間には、次のような関係が存在している（Thompson［2011］）。

$$\mathrm{Var}[\mathbf{b}_{\text{Two-way}}] = \mathrm{Var}[\mathbf{b}_{\mathrm{firm}}] + \mathrm{Var}[\mathbf{b}_{\mathrm{year}}] - \mathrm{Var}[\mathbf{b}_{\mathrm{White}}] \qquad (7)$$

このとき、Tを固定して、Nを大きくすると、

$$\lim_{\substack{T\ fixed \\ N\to\infty}} \mathrm{Var}[\mathbf{b}_{\text{Two-way}}] = \mathrm{Var}[\mathbf{b}_{\mathrm{year}}] \qquad (8)$$

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クラスター頑健手法の分散は、年度のクラスター頑健手法の分散とほとんど変わらなくなるのである（注５）。

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クラスター頑健手法を用いることが大切であるということを意味しており、企業数に比して年度数の少ないショート・パネル・データを用いた分析では、年度に関してクラスター頑健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(8)が成り立つので、先に述べた、Two-wayクラスター頑健手法から得られる

（注４）(4)(5)の企業および年度に関するクラスター頑健手法の漸近分散推定量には正定値性が保証されているので、(4)および(5)から得られた分散が負になることはない。

（注５）直観的な解釈としては、(7)の右辺の３つの項は、全て一致推定量である$\mathbf{b}$の分散推定量であるので、第１項の$\mathrm{Var}[\mathbf{b}_{\mathrm{firm}}]$は企業数$N$が大きくなるとゼロに近付き、第２項の$\mathrm{Var}[\mathbf{b}_{\mathrm{year}}]$は年度数$T$が大きくなるとゼロに近付き、第３項の$\mathrm{Var}[\mathbf{b}_{\mathrm{White}}]$は全体の観測値数が大きくなると（すなわち企業数$N$と年度数$T$のどちらが大きくなっても）ゼロに近付く。したがって、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項と第３項が非常に小さくなるので、結果として、$\mathrm{Var}[\mathbf{b}_{\text{Two-way}}]$は第２項の$\mathrm{Var}[\mathbf{b}_{\mathrm{year}}]$とほぼ一致するのである。

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クラスター頑健手法を用いることが１つの解決策であるといえる。

６．おわりに

本稿では、パネル・データを用いた分析におけるクラスター頑健手法の使用について論じている。最初に、企業と年度のパネル・データを用いる回帰モデルでは、観察不可能な各企業固有の効果の取り扱い方によって係数パラメータの推定方法が違ってくる。一般に、説明変数と観察不可能な各企業固有の効果の間に相関があると思われる場合には固定効果モデル、相関がないと思われる場合にはランダム効果モデルと呼ばれる係数パラメータの推定方法が用いられる。しかしながら、固定効果モデルやランダム効果モデルを用いて得られた係数推定値の有意性を統計的に検定する際に、誤差項の標準的仮定が満たされていない場合には、係数推定値の分散が正しく求められず、結果として誤った結論を導いてしまう危険性がある。クラスター頑健手法とは、そのような誤差項の標準的仮定が満たされない場合により正しく分散を求める方法であり、それによって、より適切に係数推定値の統計的検定を行うことが可能になるのである。

次に、パネル分析にクラスター頑健手法を使用するパターンとしては、固定効果モデルとランダム効果モデルに用いる以外に、単純なPooled回帰モデルに用いる場合が考えられる。これは、ランダム効果モデルにクラスター頑健手法を用いるということは、ランダム効果モデルが想定する誤差項の共分散構造の形が間違っていたということを意味しており、そのような場合には、単純なPooled回帰にクラスター頑健手法を用いる方がかえって好ましい場合があるからである。したがって、ランダム効果モデルにクラスター頑健手法を用いる代替的方法として、Pooled回帰モデルにクラスター頑健手法を用いることが考えられるのである。

第３に、パネル分析では、係数パラメータの推定方法として、固定効果モデル、ランダム効果モデル、Pooled回帰モデルの３手法があり、さらに、係数の分散推定方法としてクラスター頑健手法を用いるか否かという選択があるので、計６つのパターンがあるといえる。これらの中で、実用上の観点からはどれを使用するのが良いかといえば、状況が許すのであれば、固定効果モデルにクラスター頑健手法を用いるのが最も好ましいといえる。なぜなら、固定効果モデルは、説明変数と各企業固有の効果の間に相関がある場合のみならず、相関がない場合にも使用可能であるが、ランダム効果モデルは、相関がない場合にしか使用できないからである。

しかしながら、固定効果モデルには、説明変数にダミー変数が使えない、定数項が１つに定まらない、自由度が減少するといったもろもろの短所が存在しており、実際的問題として使用不可能である場合も多い。そのような場合には、ランダム効果モデルにクラスター頑健手法を用いる、あるいはその代替的方法として、Pooled回帰にクラスター頑健手法を用いることによって、パラメータに関する何らかの考察を得ざるをえないのである。

最後に、パネル・データを用いる近年の実証研究で頻繁に用いられているクラスター頑健手法に基づく標準誤差は、最新の統計ソフトを用いれば簡単に計算できる。しかしながら、あまりに簡単に使えるが故に、その理論的背景や問題点については見過ごされがちで、結果として誤った使い方

をしてしまう危険性がある。そのようなことを防ぐためにも、クラスター頑健手法の概要を理解して各研究に応じて適切に使用することが、正しい統計的推定ならびに結果の解釈につながると思われる。

〔参考文献〕

Angrist, J. D. and J. Pischke［2009］*Mostly Harmless Econometrics : An Empiricist's Companion*, Princeton University Press.

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), pp. 414-427.

――― ［2011］"Robust Inference With Multiway Clustering," *Journal of Business and Economic Statistics* 29(2), pp. 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, Chapman & Hall/CRC, pp. 1-28.

Greene, W. H.［2012］*Econometric Analysis, 7th ed.*, Prentice Hall.

Koreisha, S. G. and Y. Fang［2001］"Generalized Least Squares with Misspecified Serial Correlation Structures," *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 63(3), pp. 515-531.

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

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

Wooldridge, J. M.［2010］*Econometric Analysis of Cross Section and Panel Data, 2nd ed.*, MIT Press.