日本銀行金融研究所「金融研究」第10巻第4号（平成3年12月）

Encompassing：計量モデルの比較方法に関する新しい考え方

吉　田　知　生

1．　はじめに

ある経済現象を説明する計量モデルを作ろうとするとき、モデルをどのように構築するかは各人各様である。理論レベルで研究者が異なる立場をとることがあることはもちろん、説明変数の選択、ラグ構造の仮定等により、無数のモデルを作ることが可能であり、実際にも同一の事象を説明する複数のモデルが存在することは珍しくない。

こうした状況の下では、当然のことながら複数のライバル・モデルの優劣を比べてみたいというニーズが生じる。また、もしも誰もが納得できるような客観的な方法でそうした比較が可能になれば、それは単なる好奇心の充足にとどまらず経済学全体の進歩にも大きく貢献するはずである。

しかし、あるモデルが他のモデルの特殊型である場合（これを「入れ子型」モデル<nested models>という）については、F検定等の古典的な統計理論をそのまま応用することができるものの、それ以外のケース（「非入れ子型」モデル<non-nested models>）については、1970年代までCox［1961,1962］をはじめとする英国ないし豪州の計量経済学者による散発的な研究が行われてきたにすぎなかった。

ところが、1980年代になると英国ロンドン・スクール・オブ・エコノミックス（LSE）出身の計量経済学者を中心に、encompassingと呼ばれる新しい概念を用いて、従来の断片的な研究成果を統合しつつ計量モデルの優劣比較の問題に正面から取組もうとする動きが活発化してきた。Encompassという言葉は本来「取り囲む」ないし「包容する」という意味であるが、計量経済学では「あるモデル

本論文の4.および補論2.の計測は西島裕子が担当した。本論文の作成にあたっては、本多祐三（神戸大学）、馬場善久（創価大学）、川崎能典（東京大学大学院）の各氏から有益なコメントを頂いた。なお、実証分析のソフトウェアにはPC-GIVEとMICRO-FITを使用した。

（M_1）が別のモデル（M_2）と同等以上の説明力をもつ」といった意味で用いられる。本論文では、この encompassing について、その概念および具体的検定方法につき簡単な解説を行う。

なお、encompassing は、単なる統計的テストにとどまらず、LSE の哲学者 Imre Lakatos の「洗練された反証主義」(sophisticated falsificationism) の実践といった側面を合わせ持ち、David Hendry に代表される英国計量経済学の方法論の中心概念の1つとなっているが、この点については吉田［1991］を参照されたい。

本論文の構成は以下のとおりである。2.では、encompassing の解説に先立ち、それ以前の「入れ子型」モデルならびに「非入れ子型」モデルの統計的比較方法について簡単にサーベイする。3.では encompassing の考え方を「データ生成プロセス」の概念とともに説明した後、具体的な統計的テストの方法について述べる。4.では実際に2つの異なる通貨需要関数を推定し、encompassing test を用いてそれらの優劣比較を行った結果を紹介する。5.では結論に代え今後の研究課題について述べる。

2. 従来のモデル比較方法に関する簡単なサーベイ

(1) 「入れ子型モデル」(nested models)と「非入れ子型モデル」(non-nested models)

いま、被説明変数 y の動きを説明するために、2つの説明変数セット X、Z が用意されているとしよう。モデルが線型であれば、X、Z を説明変数にもつモデル M_1、M_2 はそれぞれ以下のように表される。

$$M_1: y = X\beta + u \qquad u \sim N(0,\ \sigma^2 I) \quad (1)$$

$$M_2: y = Z\gamma + v \qquad v \sim N(0,\ \tau^2 I) \quad (2)$$

（但し、y、X、Z はそれぞれ $T\times 1$、$T\times k$、$T\times \ell$ の行列で、T はサンプル数、k、ℓ はそれぞれ説明変数の数を表す）

このとき、モデル M_1 を規定するパラメータ・ベクトルは $\alpha=(\beta',\ \sigma^2)$、モデル M_2 を規定するパラメータ・ベクトルは $\delta=(\gamma',\ \tau^2)$ と書くことができる。

さて、このように2つの異なるモデルの比較に関する従来の計量経済学の考え方は、両者の関係が「入れ子型」(nested) であるケースと、「非入れ子型」(non-nested) であるケースとを区別し、それぞれ別個に統計的テストを開発すればよいというものであった。2つのモデルの関係が「入れ子型」(nested) であるとは、片方のモデルが他方のモデルの特殊ケースとして得られる場合——すなわち片方のモデルのパラメータに何らかの制約を課すと他のモデルが得られる場合——を指す。次の(3)(4)式で表される2つのモデルは、M_1 が M_2 の説明変数をすべて含む（ないし M_1 パラメータの一部に $\beta_2=0$ とのゼロ制約を課すと M_2 が得られる）形となっており、「入れ子型」モデルの最も典型的なケースである。

$$M_1: y = X_1\beta_1 + X_2\beta_2 + u_1 \qquad u_1 \sim N(0,\ \sigma^2 I) \quad (3)$$

$$M_2: y = X_1\beta_1 + u_2 \qquad u_2 \sim N(0,\ \sigma^2 I) \quad (4)$$

（但し、X_1、X_2 はそれぞれ $T \times k_1$、$T \times k_2$ の行列、M_1 の説明変数の数は $k_1 + k_2$、M_2 の説明変数の数は k_1）

これ以外にも、片方のモデルに様々な線型ないし非線型の制約を課すと他のモデルが得られる場合、両者は「入れ子型」の関係にあるという。[1] これに対し、片方のモデルが他方のモデルの特殊形として表すことが出来ない場合、両者の関係は「非入れ子型」(non-nested) であると呼ばれる。研究者が自分のモデルと他人のモデルを比較するといった実践的なケースにおいては、モデルの片方が他方の特殊形であることはまれであり、「非入れ子型」モデルに関する統計的テストが必要になるケースが多いが、統計理論上まず発達を見たのは「入れ子型」モデルの比較に関する分野であった。一方、「非入れ子型」モデルの比較に関しては、Cox [1961,1962] の先駆的な業績はあったものの、研究が本格化をみたのは1970年代以降のことである。

(2) 「入れ子型モデル」(nested models) の比較

2つのモデルが「入れ子」の関係にある場合にそのどちらを選ぶかは、パラメータに対する制約が妥当するか否かに基本的に依存する。すなわち、制約付きのモデル（上記のモデルでは M_2）は、制約のないモデル（同 M_1）に比べて、制約のある分推定するパラメータの数が少なくて済む利点があるから、利用可能なデータ数に限りのある計量経済モデルにおいては、制約が妥当する限り、自由度確保の観点から小さなモデル (parsimonious model) の方が望ましい。逆に、制約が妥当しなければ、小さなモデルは過度に単純化されていることになり、同モデルは y を説明するものとして適当とはいえない。[2]

モデルのパラメータに対する制約の妥当性を調べるための統計的テストにはいくつかの方法があるが、その中でも計算が簡便なF検定が実際によく利用されている。

F検定は、上記(3)(4)式のようにモデルに対する制約が線型であり、かつ両モデルのパラメータの最良線型不偏推定量 (Best Linear Unbiased Estimator) がともに最小自乗法により計算できるケースにおいて、両モデルの残差平方和を利用して制約の妥当性を調べるものである。[3] 今、M_1 の β に対する線型制約を一般化して

$$R\beta = r \tag{5}$$

1) モデルのパラメータ β に関する制約の一般形は、β の関数 h について

$$h(\beta) = 0$$

との形で表されるが、制約が線型の場合には、$h(\beta)$ はさらに

$$h(\beta) = R\beta - r = 0 \qquad (R、r \text{ はそれぞれ } j \times k、j \times 1 \text{ の行列、} j \text{ は制約の個数})$$

の形で表すことができる。

2) この考え方は、後述する parsimonious encompassing の考え方である。

3) パラメータに制約がある場合の最小自乗法は「制約付き最小自乗法」と呼ばれる。制約がない場合の β の最小自乗推定量 $\hat{\beta} = (X'X)^{-1}X'y$ を用いると、$R\beta = r$ の線型制約の下での β の推定量 $\hat{\beta}_R$ は

$$\hat{\beta}_R = \hat{\beta} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - r)$$

により求めることができる (Grasa [1989] p.74)。

（但し、R は $q \times k$ の行列、β は $k \times 1$ の係数ベクトル、r は $q \times 1$ の定数ベクトル）

の形で表せば、「この制約が正しい」との帰無仮説の下では、両モデルの残差平方和の差を制約の個数 q で除したものを分子に、制約のないモデルの残差平方和を同モデルの自由度（$T-k$）で除したものを分母にした比率

$$F = \frac{(\mathrm{RRSS}-\mathrm{URSS})/q}{\mathrm{URSS}/(T-k)} \tag{6}$$

（但し、URSS (Unrestricted Residual Sum of Squares) は制約がないモデルの残差平方和、RRSS（Restricted Residual Sum of Squares）は制約付きモデルの残差平方和をそれぞれ表す）

は自由度 q、$T-k$ のF分布に従う。したがって、この値が棄却水準を超える場合には帰無仮説が棄却され、制約は妥当しないと結論される。

F検定は、計算が容易なうえ、サンプルサイズに応じた棄却水準を分布表から得ることができる（すなわち「小標本」検定が可能）点が便利であるが、最小自乗法によるモデルの推定が適当でない（①モデルの誤差項が互いに独立でない、②モデルに対する制約が非線型である、等）場合には利用できないという限界がある。このような場合には、モデルの推定を最尤法（Maximum Likelihood Estimation）で行ない、尤度比（Likelihood Ratio, LR）検定（Neyman and Pearson [1928]）により、制約の妥当性を調べることが考えられる。

今、制約がない場合に尤度関数 $L = L(X, \theta)$ を最大化するパラメータを $\hat{\theta}$、線型ないし非線型の制約 $h(\theta) = 0$ のもとで尤度関数を最大化するパラメータを $\bar{\theta}$ とすると、尤度比検定の検定量は、

$$LR = 2\left[logL(\hat{\theta}) - logL(\bar{\theta})\right] \tag{7}$$

で与えられ、この統計量は「制約が正しい」との帰無仮説の下で自由度 q の χ^2 分布に漸近的に従う。なお、参考までに最小自乗法による推定量が最尤推定量と一致するケース（すなわち上述のF検定が可能なケース）では、LR は

$$LR = Tlog\,(\mathrm{RRSS}/\mathrm{URSS})$$
$$= Tlog\left(1 + \frac{q}{T-k}F\right) \tag{8}$$

となり、F と LR の2つの検定量は1対1に対応する（このような場合2つの検定は「同等」(equivalent)であるという）。

さらに、何らかの理由で制約付きのモデルの推定が困難な場合には、制約のないモデルのみを推定し、そのパラメータ（係数およびその分散）の推定値から制約の妥当性を吟味するワルド（W）検定（Wald [1943]）や、逆に制約のないモデルの推定が困難な場合に制約付きのモデルのみを推定し、制約なかりせば尤度がどの程度増すかを吟味するラグランジュ乗数（Lagrange Multiplier, LM）検定（Silvey [1959]）といった方法も可能である。これらの2つの検定量も、LR検定と同様漸近的に自由度 q の χ^2 分布に従う。[4]

4) 尤度関数を $L(X, \theta)$ とし、パラメータ θ に対する線型ないし非線型の制約を $h(\theta) = 0$、制約がな

(3) 「非入れ子型モデル」(non-nested) の比較

前項では、「入れ子型」モデルのどちらを選ぶかの問題については、$h(\theta)=0$ との制約が妥当するか否かの検定問題とみなすことが可能であり、統計学の古典的な検定方法をそのまま適用すればよいことをみた。しかしながら、比較するモデルが互いに他の特殊ケースでない「非入れ子型」モデルの比較については、——少なくともごく最近まで——古典的なテストはそのままの形では妥当せず、別個の統計的テストが必要になると考えられてきた。

もちろん、2つのモデル M_1、M_2 がともに他方の特殊ケースでない場合であっても、M_1、M_2 の説明変数をすべて持つような合成モデル (M_C) を作れば、各々のモデルと合成モデルとは「入れ子」の関係となるから、M_1 と M_C ないし M_2 と M_C との比較に関しては、F検定等を適用することができる。そこで M_C を介して2つのモデルの間接的な比較を行うことは不可能ではない。こうしたアプローチは artificial nesting と呼ばれる。例えば2つのモデルが(1)(2)式のような形で与えられ、両者に重複する説明変数がないとき、M_C は

$$M_C : y = X\beta + Z\gamma + u_C$$
$$u_C \sim N(0,\ \omega^2) \qquad (9)$$

で与えられ、ここで M_C に関して $\beta=0$ との帰無仮説が棄却される一方、$\gamma=0$ との帰無仮説が棄却されなければ、テストの結果は M_1 に有利とされる。

しかしながら、このやり方に対しては、①M_1 ないし M_2 のような理論に裏付けられたモデルと、M_C のような理論的意味づけを持たない機械的モデルを比較することに問題はないか、②自由度の低下やマルチコの発生により M_C の推定精度が落ちるため検出力が低くなるのではないか、などといった懐疑的な見方も少なくなかった。[5)]

このうち、①については、実は後述のように encompassing の観点から新しい解釈を与えることが可能であることが後に判明することになるのであるが、1960～70年代には artificial nesting に頼らず、2つの「非入れ子型」モデルの優劣をダイレクトに比較する新たな統計的検定方法を模索する動きが活発化するところとなった。この分野の研究で大きな役割を果たしたのは、Cox [1961,1962]

い場合の θ の最尤推定量を $\hat{\theta}$、制約付きの最尤推定量を $\bar{\theta}$ で表せば、ワルド (W)、ラグランジュ乗数 (LM) の各検定量は次のように与えられる (Amemiya [1985] pp.140-41)。

$$W = -h(\hat{\theta})' \left\{ \frac{\partial h}{\partial \theta'}\bigg|_{\hat{\theta}} \left(\frac{\partial^2 \log L}{\partial \theta \partial \theta'}\bigg|_{\hat{\theta}} \right)^{-1} \frac{\partial h}{\partial \theta'}\bigg|_{\hat{\theta}} \right\}^{-1} h(\hat{\theta})$$

$$LM = -\frac{\partial \log L}{\partial \theta'}\bigg|_{\bar{\theta}} \left(\frac{\partial^2 \log L}{\partial \theta \partial \theta'}\bigg|_{\bar{\theta}} \right)^{-1} \frac{\partial \log L}{\partial \theta'}\bigg|_{\bar{\theta}}$$

ワルド検定の応用例としては、パラメータに対する非線型制約の検定 (例えば Sargan [1980]) が、ラグランジュ乗数検定の応用例としては、残差項の系列相関、分散不均一性等の検定 (例えば Godfrey [1978]、Breusch and Pagan [1979]) がある。このほか大標本における仮説検定一般の問題に関しては、Engle [1982b,1984] 等を参照。

5) このほか、モデルが線型でない場合は、合成モデルを作るのは必ずしも容易でないという問題もある。

の先駆的業績である。Coxが示したアイデアは「入れ子型」モデルにおける尤度比検定を一般化し、比較するモデルが「非入れ子型」である場合にも適用可能にしようとするものである。

いま具体的に比較するモデルが(1)(2)式で与えられるものとしよう。2つのモデルが「入れ子型」の関係にあれば、「制約が妥当である」との帰無仮説が正しいとき、2つのモデルの対数尤度は等しくなるため、帰無仮説が正しいときの検定量（対数尤度の差）の期待値はゼロになる。ところが、「非入れ子型」のケースでは、「M_1が正しいモデルである」との帰無仮説の下では、M_1の対数尤度推定値がM_2の対数尤度推定値を通常上回るものと予想されるため、検定量の期待値はゼロにはならない。したがって、検定量の期待値をゼロとするためには、何らかの工夫が必要となる。そこでCoxはM_1が正しいとの帰無仮説の下での対数尤度差の期待値

$$E_{\hat{\alpha}}\ \{L_1(\hat{\alpha})-L_2(\hat{\delta})\} \tag{10}$$

を2つのモデルの対数尤度差から差し引いた

$$\begin{aligned} T_\alpha = &\{L_1(\hat{\alpha})-L_2(\hat{\delta})\} \\ &-E_{\hat{\alpha}}\{L_1(\hat{\alpha})-L_2(\hat{\delta})\} \end{aligned} \tag{11}$$

を通常の対数尤度差の代わりに用いることを提案した（ただし、前述のとおり、$\alpha=(\beta', \sigma^2)$、$\delta=(\gamma', \tau^2)$である)。さらにCoxは、この検定量の分布に関して、T_αをその標準偏差で$\{\hat{V}_\alpha\}^{1/2}$で除した検定量Nが漸近的に標準正規分布N（0，1）に従うことをも示した。

以上の議論は、帰無仮説を「M_1が正しいモデルである」とした場合についての検定であるが、同様のテストを「M_2が正しいモデルである」との帰無仮説についても行う必要がある。これは、「非入れ子型」モデルの場合には、2つのモデルがそれぞれ異なる条件付き分布関数をもつため、帰無仮説をとるモデルに変えれば分布関数も変わり、結論も変わってくるためである。つまり、「非入れ子型」モデルの検定においては、2つのモデルを交互に帰無仮説として2回検定を行う必要があり、その結果、①双方の検定においてともに帰無仮説が棄却される（すなわち、どちらのモデルも真でないとされる）ケースや、②帰無仮説がともに受容される（すなわち、どちらのモデルも真とされる）ケースがありうる。

「非入れ子型」モデルの優劣比較に関する上記のCoxのアイデアについて、当時の学界では、対数尤度差の期待値や分散の計算が複雑との理由で、「統計理論的に優れていても実際の利用は困難」（例えばDhrymes *et al*. [1972]）との見方が強く、1960年代中はこの分野の研究はほとんど進まなかった。

ところが、このような先入観は、Pesaran [1974] により、最尤推定量がOLSにより得られる場合にはCoxの検定量が計算可能なことが示されたことにより払拭され（このテストはPesaranのNテストと呼ばれる)、[6] 1970年代後半から「非入れ子型」モデルの優

---

6) Pesaran [1974] が求めた検定量は次のとおりである。

$$N=\frac{(T/2)\,ln\,[\hat{\tau}^2/\{(T\hat{\sigma}^2+\hat{\beta}'X'M_Z X\hat{\beta})/T\}]}{(\hat{\sigma}^2/\hat{\tau}_\alpha{}^4)\hat{\beta}'X'M_Z M_X M_Z X\hat{\beta}}$$

但し、$\hat{\tau}_\alpha{}^2=(T\hat{\sigma}^2+\hat{\beta}'X'M_Z X\hat{\beta})/T$

劣比較に関する統計的検定の研究が活発化した。Davidon and MacKinnon [1981] は、計算が極めて簡単な検定方法（Ｊテスト）を考案した。このテストは、２つのモデルを加重平均してそのウェイトをみるという Cox [1961]、Atkinson [1970] のアイデアを拡張したものである。上記のM_1、M_2を例にとると、２つのモデルを$1-\lambda$とλとで加重平均したモデルは

$$y=(1-\lambda)\ X\beta+\lambda Z\gamma+\varepsilon \tag{12}$$

で与えられる。ここで「M_1が正しいモデルである」ことを帰無仮説とすれば、帰無仮説上では$\lambda=0$となるから、これを検定すればよいことになる。もっとも、⑿式から直接λの値を推定することは不可能であるので、Davidson and MacKinnon [1981] は、⑿式の$Z\gamma$をM_2から得られる$Z\hat{\gamma}=\hat{y}_2$で代用し、

$$y=(1-\lambda)\ X\beta+\lambda\hat{y}_2+\varepsilon \tag{13}$$

としたうえで、λのt値の有意性を検定すればよいとした。この方法は、一種の artificial nesting であるが、自由度の低下やマルチコの発生によりM_Cの推定精度が落ちる、という問題は回避されている。また、Pesaran のＮテストと比べて、通常の回帰分析パッケージさえあれば誰でも手軽に行える点がメリットである。

ＮテストとＪテストの検定量はいずれも検定量の漸近分布に基づいた大標本検定であるのに対し、小標本での検定力を高めたＮＡテスト（Godfrey and Pesaran [1983]）、ＪＡテスト（Fisher and McAleer [1981]）等も考案されている。7) なお、これらのテストは、一見それぞれ別個の検定のように見えるが、実は3.で説明する encompassing の枠組みの中では、encompassing に関する一般的なテストの特殊形と位置づけることができる。

3. Encompassingの考え方と検定方法

(1) Encompassingの考え方

1980年代になると英国の計量経済学者を中心に、モデルの比較問題を encompassing と呼ばれる新しい視点からとらえようとする動きが台頭した。Encompassing の考え方の大きな特徴は、①「データ生成プロセス」(Data Generation Process、以下 DGP と略）という概念を導入した上で、従来「入れ子型」

---

7) Godfrey and Pesaran [1983] のＮＡテストの検定量は次のとおりである。

$$NT=\frac{(T/2)\,ln[\{\hat{\tau}^2/T-k\}/\{(T\hat{\sigma}^2+\hat{\beta}'X'M_ZX\hat{\beta})/(T-\ell)\}]}{(\hat{\sigma}^2/\hat{\tau}_\alpha^{\,4})\hat{\beta}'X'M_ZM_XM_ZX\hat{\beta}+\frac{1}{2}\hat{\sigma}^2\,Tr\,(B^2)}$$

但し、$Tr\,(B^2)=\ell-Tr\,(A_1A_2)^2-\dfrac{\ell-\{Tr\,(A_1A_2)\}^2}{T-k}$

$$A_1=X\,(X'X)^{-1}X' \quad A_2=Z\,(Z'Z)^{-1}Z'$$

また、ＪＡテストでは⒀式に代えて次の回帰式のμに関するt値が用いられる。

$$y=X\beta\,(1-\mu)+\mu(\hat{A}_1y_2)+\varepsilon$$

このほか「非入れ子」型モデルに関する仮説検定の問題については、Pesaran and Deaton [1978]、Davidon and MacKinnon [1982]、McAleer and Fisher [1982]、MacKinnon [1983]、Ericsson [1983]、Gourieroux *et al*. [1983]、Godfrey and Pesaran [1983]、Godfrey [1984]、McAleer and Pesaran [1986]、McAleer [1987]、Grasa [1989] 等を参照。

および「非入れ子型」の2分法で考えられてきたモデルの比較問題を、ひとつの体系の下で統一的に整理しようとすること、②モデル同士の優劣比較を、「あるモデルが他のモデルと同等以上の説明力を有する（この時「当該モデル」は「他のモデル」をencompassするという）かどうか」という点を基準に行うこと、の2点にある。

この考え方では、従来のモデルが説明し得た事象をすべて説明し、かつ説明し得なかった事象までをも説明する新しいモデルが現れた時（すなわち、新しいモデルが従来のモデルをencompassする一方、従来のモデルにはencompassされない時)、従来のモデルはその役割を終え「用済み」(redundant) となる。David Hendryを中心とする英国の計量経済学者は、このencompassingを基準に新旧モデルの比較を繰り返していけば、説明力の劣るモデルは次々に淘汰され、よりDGPに近いモデルを得ることができると主張する。このように実証研究を「漸進的」(progressive) に進めようとする方法論は、ロンドン・スクール・オブ・エコノミックス(LSE) の哲学者Imre Lakatosの唱えた「洗練された反証主義」(sophisticated falsificationism) ないし「科学的研究プログラム」(Method of Scientific Research Programme) の考え方 (Lakatos [1970,1978]) の影響を強く受けたものである。

第1図は、encompassingの考え方を図解したものである。8) Encompassingを考える上で重要な役割を果たすのが、DGPの概念である。DGPとは、経済変数が相互に影響し合いながら次々と発生していくプロセスを指すものであり、あるひとつの経済変数yのDGPは、yに影響を及ぼしている諸要因をW、影響の及ぼし方を表すパラメータをθで表せば、

$$y \sim D\ (y \mid W、\theta) \tag{14}$$

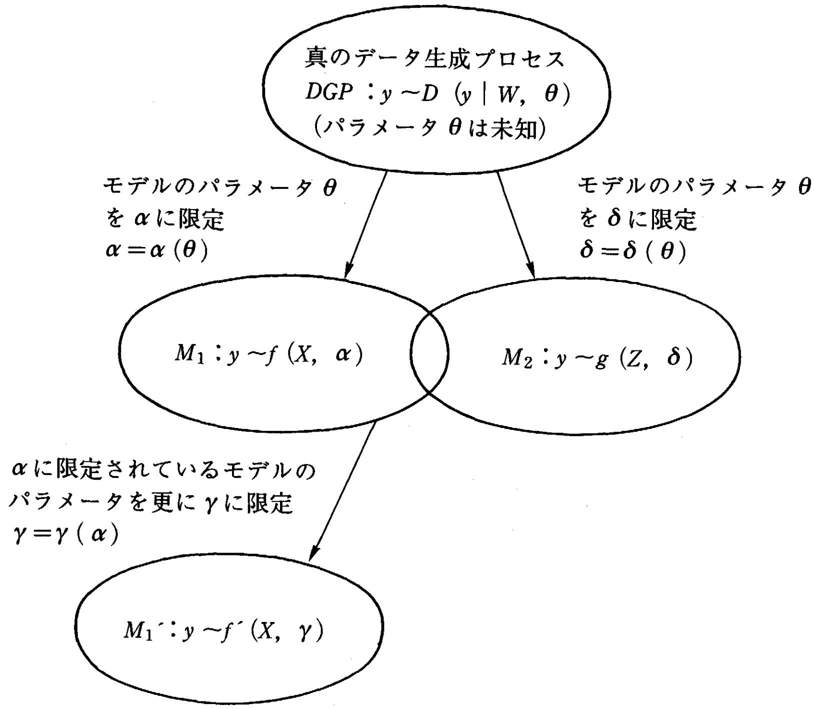
という条件付き確率分布の形で表すことができる。もちろん、W、θとともにDGPの全体の姿は未知であるから、この概念は多分に便宜的なものであるが、このDGPの存在を仮定しておくことは、計量分析の目的、モデルの位置づけといったものを明確にする上で有益である。

モデルを作る目的は、経済理論の妥当性の検証、景気予測、政策効果のシミュレーション等様々であるが、それらに共通することは、研究者がデータのアベイラビリティやモデル全体の自由度等の制約の下で、本来無数の要因が複雑に絡み合ったDGPを模倣 (mimic) する単純かつ推定可能なモデルを構築し、それを手がかりに未知のDGPの姿・特徴を探ろうとしている点である。つまり、研究者が作るモデルは、すべてDGPを各人各様の形で単純化したものとみなすことができる。

例えば、第1図におけるモデルM_1は、DGPの未知のパラメータθに対して$\alpha = \alpha\ (\theta)$との制約を置くことにより、説明変数を$X$に限定したモデルとみなすことができる。同様に、$\delta = \delta\ (\theta)$との制約を置けば、$Z$を説明変数とするモデル$M_2$が得られる。一般に$M_1$、$M_2$は互いに他方の特殊型ではなく、両者の関係は「非入れ子型」である。

8) 本節のencompassingの説明は、Hendry and Richard [1989] に多くの点で依拠している。

第1図　DGP とモデルの関係



(注)　y　：被説明変数
W：DGP に含まれる全ての変数（未知）
X　：M_1 の説明変数
Z　：M_2　　〃
θ　：DGP のパラメータ（未知）
α　：M_1 のパラメータ
δ　：M_2　　〃
γ　：M_1'　　〃
DGP と M_1、M_2、M_1'、etc. との関係：「入れ子型」(nested)
M_1 と M_2 の関係：「非入れ子型」(non-nested)
M_1 と M_1' の関係：「入れ子型」(nested)

他方、M_1に更に$\gamma=\gamma$ $(\alpha(\theta))$ との制約を上乗せしたものがM_1'であり、M_1とその特殊型であるM_1'との関係は「入れ子型」である。

前節でみたように、従来のモデルの比較方法は、①モデルが「入れ子型」の場合には、制約の妥当性の検定、②「非入れ子型」の場合に双方のモデルを交互に「真」とする仮説検定、という形の二分法的アプローチがとられており、両者の関連は不明確であった。しかしながら、上記のようにDGPを頂点とする枠組みの中では、すべてのモデルはその傘下に収まる（ないしDGPに対しては常に「入れ子」の関係にある）ことになる。

こうした枠組みの下では、「あるモデルが他のモデルと少なくとも同等にDGPの特徴を捕まえているか」という点をモデルの優劣の判断基準とすれば、それは比較するモデルが「入れ子型」か「非入れ子型」かを問わず適用することができる。これがencompassingの考え方のポイントである。

まず、比較するモデルの関係が「非入れ子型」の場合を考えてみよう。第1図において、M_2の捕まえているDGPの諸特徴をM_1がM_2に劣らず捕まえている時――ないしM_2の説明できる事象をM_1がすべて説明し得る時――「M_1はM_2をencompassする」と言い、これを記号で$M_1 \mathcal{E} M_2$と表す。他方、そうでない場合――M_1がM_2の説明できる事象の高々一部しか説明できない場合――は、「M_1はM_2をencompassしない」と言い、$M_1 \sim\mathcal{E} M_2$の記号で表す。同様のことはM_2についてもいえるから、これらを総合すると、「非入れ子型」の2つのモデルを比較した結果として、以下の4通りのケースが考えられる。

① $M_1 \mathcal{E} M_2$ and $M_2 \sim\mathcal{E} M_1$
② $M_1 \sim\mathcal{E} M_2$ and $M_2 \mathcal{E} M_1$
③ $M_1 \sim\mathcal{E} M_2$ and $M_2 \sim\mathcal{E} M_1$
④ $M_1 \mathcal{E} M_2$ and $M_2 \mathcal{E} M_1$

①と②のケースでは、片方のモデルが他方のモデルをencompassする一方、他方のモデルにはencompassされず、結果としてencompassされた方のモデルが、モデルの一段の改善に貢献する要素がないという意味で「用済み」(redundant) となる。③は2つのモデルが互いに他をencompassしないケースであり、2つのモデルがともに他のモデルが見逃しているDGPの特徴を捕まえていることになる。この場合は、互いの長所を活かしつつ新しいモデルが構築されねばならない。④の互いに他をencompassするケースは、2つのモデルがほとんど同じで区別できない場合に生じる。

Hendry and Richard [1989] は、母集団上におけるencompassingについて次のように定義している。今、仮に$M_1 \mathcal{E} M_2$であるとすれば、M_1はM_2が捕まえているDGPの特徴をすべて捕まえているから、仮にDGPがM_1に置き換えられたとしても、M_2はDGPから生成されるyとM_1から生成されるy（以下y_1と呼ぶ）とを区別することができない。これはM_1を仮にDGPとしたときのパラメータδ (α) とδ (θ) が一致していることを意味しており、

$$M_1 \mathcal{E} M_2 \iff \delta(\theta) = \delta(\alpha) \tag{15}$$

が成立する。逆に$M_1 \sim\mathcal{E} M_2$であれば、M_2はM_1が見逃しているDGPの特徴を捕まえているから、M_2はこれを用いて、yとy_1とが異なるプロセスから生成されていることを

識別できるはずである。これを同様に表せば、

$$M_1 \sim \mathscr{E} M_2 \iff \delta(\theta) \neq \delta(\alpha) \qquad (16)$$

となる。平易にいえば、M_2の側から「M_1はDGPである」というウソを見破ることができるかどうかが、M_1がM_2をencompassするかどうかの鍵となる（その具体的な検定方法については後述）。

次に、第1図のM_1とM_1'のように2つのモデルが「入れ子型」の関係にある場合をみてみよう。ここで、M_1'はM_1の制約付きモデルであるから、M_1'に対してより一般的なモデルであるM_1がM_1'の説明可能な事象をすべて説明できるのは当然であり、$M_1 \mathscr{E} M_1'$は自動的に成立する。さらに、$M_1 \rightarrow M_1'$の単純化のための制約（第1図における$\gamma=\gamma(\alpha(\theta))$）が妥当する時には、小さいモデルは大きいモデルと比べて遜色がなく、$M_1 \mathscr{E} M_1'$であると同時に$M_1' \mathscr{E} M_1$である。このように「入れ子型」モデルの単純な方がより一般的な方をencompassすることをparsimonious encompassingと呼び、$M_1' \mathscr{E}_p M_1$と表す。[9] Parsimonious encompassingの例としては、より一般的なモデルの係数にゼロ制約を課した場合のほか、COMFACと呼ばれる非線形制約付きモデルのケースがある（詳しくは補論1.参照）。

(2) Encompassingの検定方法

先に述べたとおり、比較するモデルが「入れ子型」の場合（上記の例ではM_1とM_1'には、$M_1 \mathscr{E} M_1'$は自動的に成立するから、検定が必要なのはparsimonious encompassingすなわち$M_1' \mathscr{E}_p M_1$が成立するかどうかだけであり、2.でみたＦ検定（ないしＬＲ、Ｗ、ＬＭ検定）がそのまま適用できる。そこで以下では比較するモデルが「非入れ子型」の場合（上記の例ではM_1とM_2）の検定方法について説明する。

今、２つのモデルM_1、M_2が改めて次の形で与えられるとしよう。[10]

$$M_1 : y = X\beta + u \qquad u \sim N(0,\ \sigma^2 I) \qquad (17)$$

$$M_2 : y = Z\gamma + v \qquad v \sim N(0,\ \tau^2 I) \qquad (18)$$

ここでサンプル数をTとすると、y、X、Zはそれぞれ$T\times 1$、$T\times k$、$T\times \ell$の行列である。さて、このT個のサンプルから２つのモデルを推定すれば、それぞれのモデルの未知のパラメータ$\alpha=(\beta',\ \sigma_2)$、および$\delta=(\gamma',\ \tau_2)$の最尤推定量は通常の最小自乗法により得られるから

$$\hat{\beta} = (X'X)^{-1}X'y \qquad (19)$$

$$\hat{\sigma}^2 = y'M_x y/T$$

$$\text{但し、} M_x = I - X(X'X)^{-1}X' \qquad (20)$$

9) なお、parsimonious encompassingについては「推移律」が成立し、M^mをM_1とM_2の中間のモデルとしたとき、$M_1 \mathscr{E}_p M^m$かつ$M^m \mathscr{E}_p M_2$であれば、$M_1 \mathscr{E}_p M_2$であるが、モデルが「非入れ子型」の場合にはencompassingには「推移律」（transitivity）は成立しないことがある。いま、DGPに$\zeta=\zeta(\theta)$との制約を課したモデルがM_3であるとしよう。$M_1 \mathscr{E} M_2$かつ$M_2 \mathscr{E} M_3$のとき、$\delta(\theta)=\delta(\alpha)$と$\zeta(\theta)=\zeta(\delta)$から、$\zeta(\theta)=\zeta[\delta(\alpha)]$が導かれるが、これは$M_1 \mathscr{E} M_2$の必要十分条件$\zeta(\theta)=\zeta(\alpha)$とは必ずしも一致しない（Hendry［1988］参照）。

10) 以下の議論では、簡略化のためX、Zは重複する変数を持たず、ともにyに対して「強外生」（strongly exogenous）であるものとする。

$$\hat{\gamma} = (Z'Z)^{-1} Z'y \tag{21}$$

$$\hat{\tau}^2 = y'M_Z y / T$$

$$\text{但し、} M_Z = I - Z\,(Z'Z)^{-1} Z' \tag{22}$$

となる。

M_1 が M_2 を encompass しているかどうかは、「M_1 が DGP である」とのウソを M_2 が見破ることができるかどうかで決まるから、これを統計的に検定するためには M_1 を DGP とみなして、そこからyに関するデータ

$$\hat{y}_1 = X\hat{\beta}$$

を生成させ、これを M_2 で回帰した場合のパラメータの推定値 $\delta_{\hat{\alpha}} = (\gamma_{\hat{\alpha}}',\ \tau^2_{\hat{\alpha}})$ が $\hat{\delta} = (\hat{\gamma}',\ \hat{\tau}^2)$ と一致するかどうかを調べればよい。そこで encompassing difference と呼ばれる統計量 $\hat{\phi}$ を

$$\hat{\phi} = \hat{\delta} - \delta_{\hat{\alpha}} \tag{23}$$

と定義すれば、$\hat{\phi} = 0$ が統計的に棄却されるときには、M_2 は M_1 が DGP でないことを見破っていることになり、$M_1\,\mathcal{E}\,M_2$ である。一方、$\hat{\phi} = 0$ が統計的に棄却されなければ、M_2 には M_1 と DGP の区別がつかないことになり、$M_1 \sim \mathcal{E} M_2$ となる。

さて、この検定を行うためには、統計量 $\hat{\phi}$ の分布を知る必要がある。近年の「定式化に誤りのあるモデルにおける最尤推定量の分布」に関する White [1982]、Gourieroux *et al*. [1984] らの研究は、「M_1 が DGP である」との帰無仮説のもとで、$\sqrt{T}\,\hat{\phi}$ の極限分布が $N(0,\ V_\alpha(\hat{\phi}))$ の正規分布となる（但し、$V_\alpha(\hat{\phi})$ は $\hat{\phi}$ の分散共分散行列）ことを明らかにしており、これを利用すれば、次のようにモデル M_1 の下で、漸近的に χ^2 分布に従う（これを以下では $\overset{as}{\underset{M_1}{\sim}}$ と書く）ワルド検定量を作ることができる。

$$\eta_W(\hat{\delta}) = T\hat{\phi}' V_\alpha(\hat{\phi})^{-1} \hat{\phi} \ \overset{as}{\underset{M_1}{\sim}}\ \chi^2(\ell) \tag{24}$$

この検定は、全ての未知のパラメータ（γ および τ^2）を包括的に検定するという意味で、complete parametric encompassing (CPE) test と呼ばれる。[11] もっとも、この検定量は計算が複雑であることに加え、漸近的には次に述べる γ のみに関する coefficient encompassing test と同等 (equivalent) であるため、実際に利用されることは少ない。

次に、encompassing differnce を γ についてのみをとった場合、$\hat{\phi}$ は

$$\begin{aligned}\hat{\phi} &= \hat{\gamma} - \gamma_{\hat{\alpha}} \\ &= (Z'Z)^{-1} Z'y - (Z'Z)^{-1} \\ &\quad Z'X\,(X'X)^{-1} X'y \\ &= (Z'Z)^{-1} Z'\{I - X\,(X'X)^{-1} X'\} y \end{aligned} \tag{25}$$

となり、その分散 $V_\alpha(\hat{\phi})$ は

$$V_\alpha(\hat{\phi}) = T\sigma^2\,(Z'Z)^{-1} Z' M_X Z\,(Z'Z)^{-1} \tag{26}$$

で与えられる。したがって、γ に関する en-

---

11) X、Z が重複する説明変数を持つ場合の CPE test の検定量は次のように与えられる。

$$\eta_W(\hat{\delta}) = T\hat{\phi}' V_\alpha(\hat{\phi})^{+} \hat{\phi} \ \overset{as}{\underset{M_1}{\sim}}\ \chi^2(\rho)$$

ここで $V_\alpha(\hat{\phi})^{+}$ は $V_\alpha(\hat{\phi})$ の一般化逆行列である。この検定量が漸近的に従う χ^2 分布の自由度 ρ は ℓ から X と Z の間で重複している説明変数の数を引いた値で $V_\alpha(\hat{\phi})$ 行列のランクに一致する。

compassingのワルド検定量$\eta_W(\hat{\gamma})$は、

$$\eta_W(\hat{\gamma})=T(\hat{\gamma}-\gamma_{\hat{\alpha}})'V_\alpha(\hat{\phi})^{-1}(\hat{\gamma}-\gamma_{\hat{\alpha}})$$
$$=\hat{\sigma}^{-2}\cdot y'M_XZ(Z'M_XZ)^{-1}$$
$$Z'M_Xy \quad \frac{as}{M_1} \quad \chi^2(\ell) \qquad (27)$$

となり、漸近的に自由度ℓ（モデルM_2の説明変数の数）のχ^2分布に従う。この検定は、前述のCPE testに対して、coefficient encompassing testとも呼ばれる（Gilbert [1986]）。ところで、このcoefficient encompassing testは、次式により同等のF検定に変換することができるため、これを利用すれば小標本検定も可能になる。

$$\ell F=\frac{T-\ell-k}{T}\eta_W(\hat{\gamma})$$
$$\left[1-\frac{1}{T}\eta_W(\hat{\eta})\right]^{-1} \qquad (28)$$

さらに、このF検定量はM_1、M_2を人工的に合成したモデル

$$M_C: y=X\beta+Z\gamma^*+e$$
$$e\sim N(0,\ \omega^2) \qquad (29)$$

における$\gamma^*=0$のF検定量と一致することがMizon [1984]、Mizon and Richard [1986]によって示された。[12] このことは、従来artificial nestingと呼ばれていた手法が実はcoefficient encompassing testであることを意味している。もっとも、encompassing testは2つのモデルを直接に比較するものであり、M_1とM_CないしM_2とM_Cを比較している訳ではない。確かにF検定量の計算の都合上は合成モデルM_Cを推定した方が便利であるが、その場合でもM_Cの存在は便宜的なものにすぎず、M_Cが理論的に解釈できないと

12) $\gamma^*=0$に関するF検定は次の式で与えられる。

$$F=\frac{T-\ell-k}{\ell}\cdot\frac{\hat{\gamma}^{*\prime}Z'M_XZ\hat{\gamma}^*}{y'M_Xy-\hat{\gamma}^{*\prime}Z'M_XZ\hat{\gamma}^*}$$

但し、$\hat{\gamma}^*=(Z'M_XZ)^{-1}Z'M_Xy$

ここで

$$\hat{\gamma}^{*\prime}Z'M_XZ\hat{\gamma}^*=y'M_XZ(Z'M_XZ)^{-1}Z'M_XZ(Z'M_XZ)^{-1}Z'M_Xy$$
$$=y'M_XZ(Z'M_XZ)^{-1}Z'M_Xy$$
$$=\eta_W(\hat{\gamma})\cdot\hat{\sigma}^2$$

であるから、上記のF検定量は

$$F=\frac{T-\ell-k}{\ell}\cdot\frac{\eta_W(\hat{\gamma})\cdot\hat{\sigma}^2}{T\hat{\sigma}^2-\eta_W(\hat{\gamma})\cdot\hat{\sigma}^2}$$
$$=\frac{T-\ell-k}{\ell T}\cdot\frac{T\eta_W(\hat{\gamma})}{T-\eta_W(\hat{\gamma})}$$

と書き直すことが可能であり、これを整理すれば(28)式が得られる。

なお、X、Zが重複する説明変数を持つ場合のcoefficient encompassing testは、重複する説明変数をZから除外した上でM_Cを次のようにおき、$\gamma^*=0$をF検定すればよい。

$$M_C: y=X\beta+\bar{Z}\gamma^*+e \qquad e\sim N(0,\ \omega^2)$$

（但し、$\bar{Z}$はZのうちXと共通でない、ないしXの線形変換で得られない説明変数）

いった artificial nesting に対する過去の批判はこの encompassing test に関しては当たらない。

上記は encompassing difference を係数パラメータ γ についてとったケースであったが、同様のことは分散のパラメータ τ^2 についても行うことができる。この場合 encompassing difference $\hat{\phi}$ は

$$\hat{\phi} = \hat{\tau}^2 - \tau^2{}_{\hat{\alpha}}$$

であり、これを基に τ^2 に関する encompassing のワルド検定量 $\eta_W(\tau^2)$ を作ることができる。[13] この検定は variance encompassing test と呼ばれ、検定量は説明変数の多さにかかわらず自由度1の χ^2 分布に従う。この variance encompassing test は、対象を分散パラメータのみに絞った部分的なテストであるが、coefficient encompassing の F 検定のように、自由度不足やマルチコ問題により検定力が低下する惧れがない点でメリットがあり、実際上は両者の併用が望ましいとされている。

さて、これまでは encompassing difference $\hat{\phi}$ をモデルのパラメータ $\hat{\delta}$ ないしその一部（$\hat{\gamma}$ ないし $\hat{\tau}^2$）について考えたが、$\hat{\phi}$ を $\hat{\delta}$ と T 個のデータの観測値 $Y_T^1 = (Y_1, \ldots, Y_T)$ の関数

$$b = b\ (Y_T^1,\ \hat{\delta})$$

についてとるように一般化することも可能である。この場合の $\hat{\phi}$ は

$$\hat{\phi} = b\ (Y_T^1,\ \hat{\delta}) - E_{\hat{\alpha}}\ \{b\ (Y_T^1,\ \hat{\delta})\} \tag{30}$$

となり、これを基にさらに多くの検定量を作ることができる。例えば、b に2つのモデルの対数尤度差をとり

$$b\ (Y_T^1,\ \hat{\delta}) = (1/T)\ \{L_1(\hat{\alpha}) - L_2(\hat{\delta})\} \tag{31}$$

とすれば、(30)式は

$$\begin{aligned} T\hat{\phi} = &\{L_1(\hat{\alpha}) - L_2(\hat{\delta})\} \\ &- E_{\hat{\alpha}}\{L_1(\hat{\alpha}) - L_2(\hat{\delta})\} \end{aligned} \tag{32}$$

となる。この検定量は、2つのモデルが「入れ子型」の場合には通常の尤度比検定と一致し、「非入れ子型」の場合には前節でみた Cox の「一般化尤度比検定」(generalised likelihood ratio test) と一致する。このほかにも、ϕ のとり方を工夫すれば、N-test、J-test 等これまで散発的に開発されてきた「非入れ子型」モデルの仮説検定はいずれも encompassing test の一種として位置づけることが可能となる。[14]

4.　Encompassing test による2つの通貨需要関数の比較

(1)　2つの異なるモデルの推定

本節では、5. の encompassing test が実際どのようなものであるかを例証するため、2つの全く異なる通貨需要関数に関するモデル

13)　τ^2 に関する検定量 $\eta_W(\hat{\tau}^2)$ は、次のように与えられる (Mizon [1984])。

$$\eta_W\ (\hat{\tau}^2) = \frac{[(T-\ell)(\hat{\tau}^2 - \hat{\sigma}^2) - \hat{\beta}' X' M_Z X \hat{\beta}]^2}{4\,\hat{\sigma}^2 \hat{\beta}' X' M_Z M_X M_Z X \hat{\beta}} \quad \frac{as}{M_1} \quad \chi^2\ (1)$$

14)　N-test の場合の $\hat{\phi}$ は $\hat{\phi} = T^{-1/2}\ (\hat{\gamma}' Z' Z \hat{\gamma} - \gamma_{\hat{\alpha}}' Z' Z \gamma_{\hat{\alpha}})$、J-test の $\hat{\phi}$ は $\hat{\phi} = 2\,T^{-1/2}\ \{\hat{\gamma}' Z' Z\ (\hat{\gamma} - \gamma_{\hat{\alpha}})\}$ である (Mizon and Richard [1986])。

の比較を試みる。[15] 今回比較するモデルとは、①通貨需要関数の分野で伝統的によく用いられてきた部分調整（ＰＡ）モデル（例えば、Goldfeld [1976]）、そして②特定の理論に基づかず、単に名目通貨需要を名目総需要（名目 GNP に輸入を加えたもの）と金利のみの関数で表現したモデル（以下 NTD モデル）、の２つである。推計期間はいずれも1957年第１四半期から1989年第３四半期の約30年間とした。

部分調整（ＰＡ）モデルには、実質通貨需要の動きを当期の実質所得、金利および実質通貨需要の１期ラグにより説明する「実質調整」(real adjustment) モデルと、説明変数

第１表　部分調整（ＰＡ）モデルによる通貨需要関数（M_1）

EQ (1) Modelling　ΔLM2　by OLS
The Sample is 1957 (1) to 1988 (3) less 0 Forecasts

VARIABLE	COEFFICIENT	STD ERROR	H.C.S.E.	t-VALUE
LGNPR	0.0413	0.0176	0.0153	2.34
CALL	-0.0025	0.0004	0.0005	-6.86
L (M2/P) (-1)	-0.0449	0.0134	0.0116	-3.29
ΔLGNPD	-0.2924	0.0594	0.0670	0.492
CONSTANT	-0.0327	0.0858	0.0740	0.38

$R^2=0.604$　S.E=0.00790　F (4,125)=47.62　DW=0.55
RSS=　0.0077936　for 5 Variables and 130 Observations
Normality χ^2 (2) =17.47 [0.00]　AR 1-5 F[5,120] =35.67 [0.00]
ARCH 4 F[4,117] =27.28 [0.00]　Xi∗Xj F[8,116] = 3.50 [0.00]
RESET F[1.124] = 0.89 [0.34]

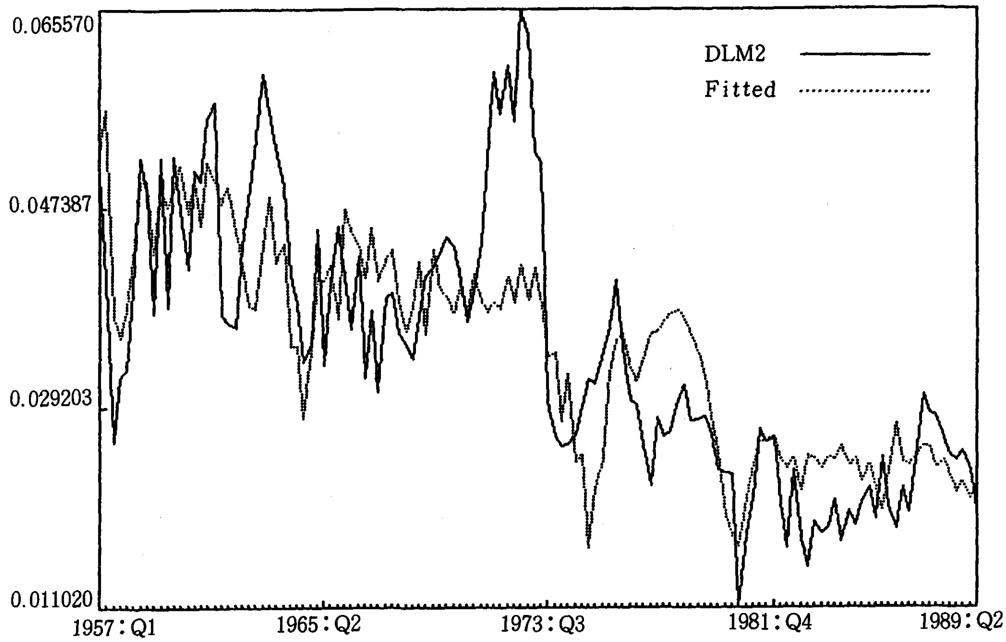
(注) LGNPR　：名目 GNP 対数値
CALL　：コールレート
L(M2/P)　：実質マネーサプライ対数値
ΔLGNPD　：GNP デフレータ前期比変化率
(　)内はラグを表す。
H.C.S.E.　：分散不均一性の影響を受けない標準誤差（White [1980]）
Normality χ^2　：残差項の正規性に関する検定（Jaque and Bera [1980]）
AR 1-5 F　：残差項の系列相関に関する検定（Godfrey [1978]）
ARCH 4 F　：残差項のARCHに関する検定（Engle [1982a]）
Xi∗Xj F　：残差項の分散不均一性に関する検定（White [1980]）
RESET F=　：一般的な「定式化の誤り」に関する Ramsey [1969] のテスト
[　]内は当該検定量が得られる確率を示す。
なお、これらの検定量全般についての解説としては Hendry [1989] を参照。

15) Encompassing test の応用例としては、Ahumuda [1985]、Muscatelli [1989] 等がある。

に当期のインフレ率を追加した「名目調整」（nominal adjustment）モデルの２種類があるが、ここでは説明変数の多い後者を採用した。[16] このモデル（M_1）の推定結果は第１表に示されている（NTD モデルとの比較を容易にするため、被説明変数が名目マネーサプライの増加率になるよう parametrization に変更を加えてある）。第１表を見ると、定数項以外の説明変数は一応有意である。ただ、モデルの残差項に対する各種診断テストはいずれもこのモデルに「定式化の誤り」（misspecification）があることを示唆している。また、推定値と現実値のプロット（第２図）を見ても、このモデルは DGP の重要な特徴を見逃しているように思われる。

第２のモデル（NTD）は、経済理論を度外視して、名目マネーサプライの増加率を、名目総需要（名目 GNP に輸入を加えたもの）の増加率と金利の分布ラグ（distributed-lag）モデルで説明しようとするモデルである。具体的には、２つの説明変数のラグを当初４期までとり、有意な説明変数のみを残すことで、最終的に第２表のようなモデル（M_2）を得た。この場合も定数項を含めた４つの説明変数はいずれも有意であるが、モデルの残差項に対する各種の診断テストは M_1 のケースと同様このモデルに「定式化の誤り」があることを示唆しており、推定値と現実値のプロット（第３図）も満足のいくものではない。

第２図　PA モデル（M_1）による推計値と現実値



16)　「実質調整」モデルと「名目調整」モデルとの違いについては、例えば Cuthbertson［1985］を参照。

第２表　名目総需要を説明変数とした通貨需要関数（NTD）の推定結果（M_2）

EQ (2) Modelling　ΔLM2　by OLS

The Sample is 1957 (1) to 1989 (3) less 0 Forecasts

VARIABLE	COEFFICIENT	STD ERROR	H.C.S.E.	t-VALUE
ΔLTD (-1)	0.2593	0.0486	0.0498	5.33
ΔLTD (-2)	0.1453	0.0508	0.0505	2.86
CALL	0.0011	0.0005	0.0004	2.40
Δ4CALL	-0.0022	0.0004	0.0004	-5.26
CONSTANT	0.0151	0.0036	0.0027	4.18

$R^2=0.388$　S.E=0.00982　F (4,125)=19.79　DW=0.54

RSS=　0.0120420　for 5 Variables and 130 Observations

Normality χ^2 (2) =13.92 [0.00]　AR 1-5 F[5,120] =40.66 [0.00]

ARCH 4 F[4,117] =31.82 [0.00]　Xi∗Xj F[8,116] = 0.57 [0.80]

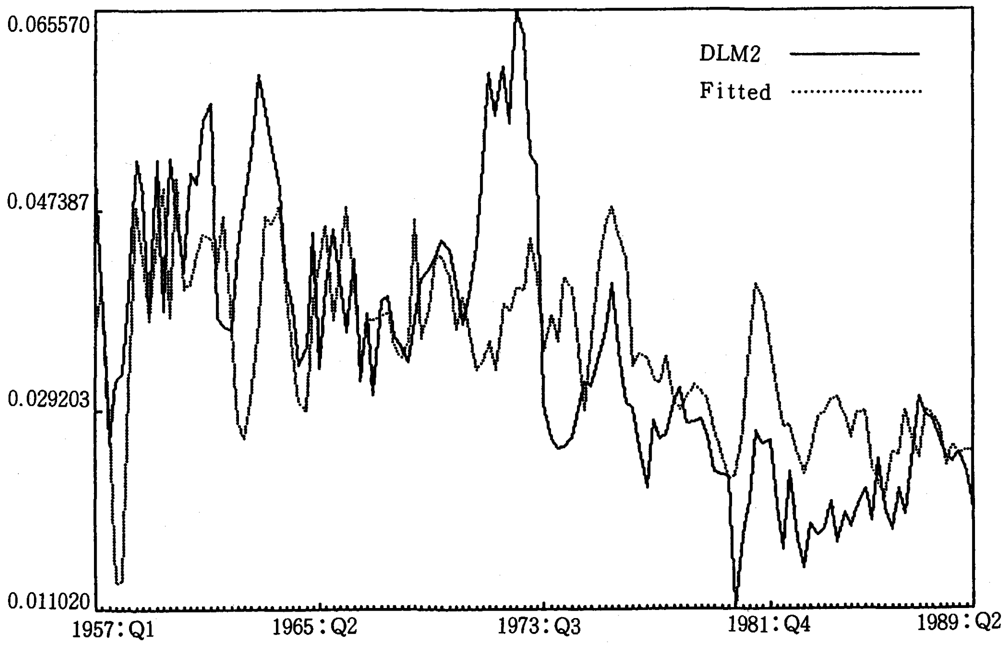
RESET F[1.124] = 0.34 [0.56]

ΔLTD　：名目総需要変化率

CALL　：コールレート

Δ4 CALL　：コールレート前年比変化幅

第３図　NTD モデル（M_2）による推計値と現実値



(2) Encompassing testの結果

上記のM_1・M_2のモデルのいずれがより良いモデルであるかは、推計結果やグラフをみた限りでは必ずしも明らかでない。そこで、各種の encompassing test を用いて２つのモデルを比較してみよう。ここでは、coefficient encompassing のＦ検定、variance encompassing のワルドテスト (W-Test) のほか、encompassing test の一種とみなすことが可能な各種の「非入れ子型」モデルの仮説検定（Pesaran［1974］のＮテスト、Godfrey and Pesaran［1983］のＮＴテスト、Davidson and MacKinnon［1981］のＪテスト、Fisher and McAleer［1981］のＪＡテスト）についても併せて行った。

M_1、M_2の encompassing test の結果を第３表に示した。第３表の左列は「M_1が DGP である」という帰無仮説をM_2に対して検定した結果であり、検定量はいずれも５～１％棄却水準ぎりぎりのところにあり、$M_1 \mathcal{E} M_2$か$M_1 \mathcal{E} M_2$か判定するのは難しい。一方、右列は「M_2が DGP である」という帰無仮説をM_1に対して検定した結果であり、検定量はいずれも帰無仮説を十分に棄却している。したがって$M_2 \sim \mathcal{E} M_1$である。２つの検定結果を総合的にみれば、結果はM_1に有利であるといえよう。

このことは、M_2がM_1に比べて redundant なモデルであることを示唆するが、これはむろんM_1が究極のモデルであることを意味しない。上述のように、M_1に対する各種診断テストの結果は、M_1に「定式化の誤り」があり、推定値と現実値のプロットも必ずしも満足のいくものではないからである。

実際、共和分の考え方を取り入れたエラー修正モデル（ECM、詳しくは吉田［1989］を参照）により各種診断テストや推定値のプロットについて一応満足のいくモデル（M_3）を得ることができ、M_3はM_1、M_2を encompass する（比較の詳細は補論2.参照)。さらに、このM_3も他の何らかのモデルに encompass されれば redundant なものとなる。このように説明力の劣るモデルを次々淘汰し、より DGP に近いモデルを得ようとするのが encompassing の基本的考え方なのである。

5. 今後の課題

Davison *et al*.［1978］が encompass という言葉を計量経済学ではじめて使用した当時、それは定義すらはっきりしない漠然としたアイデアにすぎなかった。以来 Dastoor、Mizon、Richard、Hendry らの努力により en-

第３表　M_1（PA）とM_2（NTD）との Encompassing Test

Test Statistic	M_1 against M_2		M_2 against M_1	
F-Test	F（3,122）	2.44 [0.93]	F（3,122）	25.95 [1.00]
W-Test		-2.63 [0.99]		-9.33 [1.00]
N-Test		-3.16 [1.00]		-13.95 [1.00]
NT-Test		-2.83 [1.00]		-12.69 [1.00]
J-Test		2.67 [0.99]		8.84 [1.00]
JA-Test		2.51 [0.99]		6.03 [1.00]

compassingの考え方は理論的に体系化され、統計的テストの方法も整備されつつある。しかしながら、今なお研究を要する課題も決して少なくなく、以下ではこうした点につき3点を指摘しておくこととする。

まず、第1に、3.で紹介した統計的テストの頑健性（robustness）の問題がある。F、W、N、J等encompassingに関するテストは、すべて比較する両方のモデルの残差項に系列相関や分散不均一性がないことを前提としたものであり、こうした前提が満たされない場合には検定力が低下するおそれがある。

両方のモデルがともに残差項がホワイトノイズに見えるようにデザインされている場合はこうした問題は生じない。しかしながら、単に理論モデルのパラメータを推定しただけの計量モデルは往々にして残差項の性質に問題があるケースが多く、こうしたモデルをも含めてencompassingの考え方を適用していくためには、各種のテストが残差項の系列相関や分散不均一性に対してどの程度頑健であるのかをモンテ・カルロ・シミュレーション等により明らかにしておく必要がある。こうした研究は既に一部で始められているが（McAleer［1987］のサーベイを参照）、未だ十分とは言い難く、今後研究成果の積み重ねが必要である。

第2に、これまでのencompassing testは単一方程式モデル同士を比較するものに限られてきたが、encompassingのアイデア自体は連立方程式モデル同士の比較へも適用可能であるだけに、こうした面での統計的テストの開発が期待される。さらに、大型のマクロ計量経済モデルを比較対象にする場合には、モデルの一部のみしか公表されていなかったり、推定するパラメータ数が多すぎて合成モデルの推定が不可能といった問題が生じることもあろう。その場合には、モデルが生みだした予測値のみを使ってencompassingをテストするといった工夫（Chong and Hendry［1986］）も役に立つかもしれない。[17)]

最後に、これも連立方程式モデルと関連するが、parsimonious encompassingの考え方の応用として、VARをparsimoniouslyにencompassし、かつ理論的に解釈可能な小型の構造型モデルを作ることができれば、VARをredundantにすることも不可能ではない。この面での試みとしてHendry and Mizon［1990］があるが、こうした手法が確立されれば多変量時系列モデルの分野における大きな貢献になるであろう。

補論1．COMFACモデルとコークラン・オーカット法適用の問題点

以下では、parsimonious encompassingの応用例としてCOMFAC（Common Factor）モデルを紹介するとともに、実証分析においてよく用いられるコークラン・オーカット法（Cochrane and Orcutt［1949］）の適用に関する問題点を指摘する。

いま2つの時系列モデルが

$$M_1 : y_t = \zeta x_t + \varepsilon_t$$
$$\varepsilon_t = \rho\varepsilon_{t-1} + u_{1t}$$
$$|\rho| < 1,\ u_{1t} \sim N(0,\ \sigma_1^2) \quad \text{(A-1)}$$

17) モデルの予測値のみが公表されている場合、予測パフォーマンス改善のために定数項修正が行われている可能性もあるので、注意が必要である。

$$\mathrm{M}_2 : y_t = \pi_1 y_{t-1} + \pi_2 x_t + \pi_3 x_{t-1} + u_{2t}$$
$$|\pi_1| < 1,\ u_{2t} \sim N(0,\ \sigma_2^2) \quad \text{(A-2)}$$

で表されるとすると、M_1は、

$$y_t = \rho y_{t-1} + \zeta x_t - \rho\zeta x_{t-1} + u_{1t} \quad \text{(A-3)}$$

と書き直すことができるため、M_1はM_2のパラメータに

$$\pi_3 = -\pi_1 \pi_2 \quad \text{(A-4)}$$

という非線形の制約をつけたモデルであることがわかる。したがって、(A-4) 式の制約が妥当すれば、$M_1 \mathscr{E}_p M_2$、すなわちM_1はM_2を parsimonious に encompass するから、M_2を推定する代わりにM_1を推定し、推定するパラメータの数を3個 (π_1, π_2, π_3) から2個 (ζ, ρ) に減らし、自由度を上げることができる。[18] すなわち、(A-4) 式が成立する場合の、残差項の系列相関は、通常考えられているような“nuisance”ではなく、“convenient simplification”となる (Hendry and Mizon [1978])。このM_1のようなモデルは COMFAC (Common Factor) と呼ばれる。これは、(A-3) 式をラグ・オペレーター(L) を用いて表すと

$$(1-\rho L)y_t = (1-\rho L)\zeta x_t + u_{1t} \quad \text{(A-5)}$$

となり、y_tとx_tの時系列が common factor $(1-\rho L)$ を共有しているとみることができることによる。この common factor の存在、すなわち$M_1 \mathscr{E}_p M_2$の検定については、Sargan [1980] の方法がある。[19]

さて、上記のM_1は、実証分析でDW比が低いときに、残差に1階の系列相関があると仮定して推定するコークラン・オーカット法のモデルとしてよく見かけるものであるが、上記の議論から明かなように、M_2のような一般型をM_1で代用してよいのは、$M_1 \mathscr{E}_p M_2$つまり (A-4) 式が妥当する場合のみに限られる。したがって、DW比が低いからといって直ちにコークラン・オーカット法に頼るのは手続き的にみて誤りであり、まず、(A-4) 式の制約に関する検定が先行しなければならない。DW比は残差項の1階の系列相関に対する検定としてデザインされたものではあるが、より一般的な「定式化の誤り」、例えば omitted variables や分散不均一性についても低いDW比が観察されるので、DW検定で帰無仮説が棄却されたとしても、それは必ずしも「真のモデルの誤差項は1階の系列相関をもつ」ということを意味しない (Harvey [1990] pp.156-7参照)。また、コークラン・オーカット法を適用してM_1のようなモデルを推定する場合には、(A-3) 式から明らかなように、被説明変数の自己ラグ項が回帰式にインプリシットに含まれることになるため、自己ラグ項が無いことを前提としたDW比による検定はもはや役に立たなくなる。したがって、こうしたケースでDW比を表示するのは誤解を避ける意味で慎むべきであろう。

補論2．エラー修正モデル（ECM）のパフォーマンス

この補論では、エラー修正モデル (ECM、

18) 但し、この場合M_1は OLS で推定不可能であるから、最尤法を用いなければらない。

19) この点に関しては、Mizon and Hendry [1980]、Gregory snd Veall [1985,1986] も参照。

以下M_3）と、本文で取り上げたＰＡモデル（M_1）およびNTDモデル（M_2）とのencompassingをテストした結果を紹介する。M_3は、M_2+CDに対する通貨需要を実質GNP、GNPデフレーター、金利から説明するもので、吉田［1989］のモデルと基本的に同一であるが、NTDモデルとの比較を容易にするため、被説明変数を実質マネーサプライ増加率から名目マネーサプライ増加率に変更してある（吉田［1989］のECMは、当期のインフレ率を説明変数に有しているめ、計測式の両辺にインフレ率を加えるだけで被説明変数が実質マネーサプライの変化率から名目マネーサプライの変化率になるようにreparametriseすることができる)。なお、その他の主な変更点としては、①金利に関する変数を利付き金融債応募者利回りから金利スプレッド（コールレートーM_2+CD自己金利）に変更したこと、②株価の変動係数については説明変数として有意であるものの、残差項の分散不均一性を増大させる傾向があるため、今回は除外したこと、の２点がある。

M_3の推定にあたっては、最初に各変数の４期までのラグを含むADL（Auto-regressive Distributed Lag）モデルから出発し、reparametrisationを繰り返しながら徐々にモデルを小さくしていく、General-to-Simple Modelling Strategy（吉田［1989］参照）に従った。エラー修正項（EC）における実質通貨需要の所得弾力性については、Yoshida and Rasche［1990］で求めた値（1.2）を使用した。

M_3の推計結果は第A-1表にまとめられて

第A-１表　エラー修正モデル（ECM）による通貨需要関数（M_3）

EQ (3) Modelling　ΔLM2　by OLS
The Sample is 1957 (1) to 1989 (2)

VARIABLE	COEFFICIENT	STD ERROR	H.C.S.E.	t-VALUE
ΔLM2 (-1)	0.5896	0.0801	0.0948	7.35
ΔLM2 (-2)	0.2267	0.0808	0.0891	2.80
Δ^2 LGNPD	0.1173	0.0377	0.0504	3.11
Δ2SP	-0.0016	0.0004	0.0005	-4.64
EC (-4)	-0.02140	0.0065	0.0075	-3.28
CONSTANT	-0.1017	0.0318	0.0362	-3.20

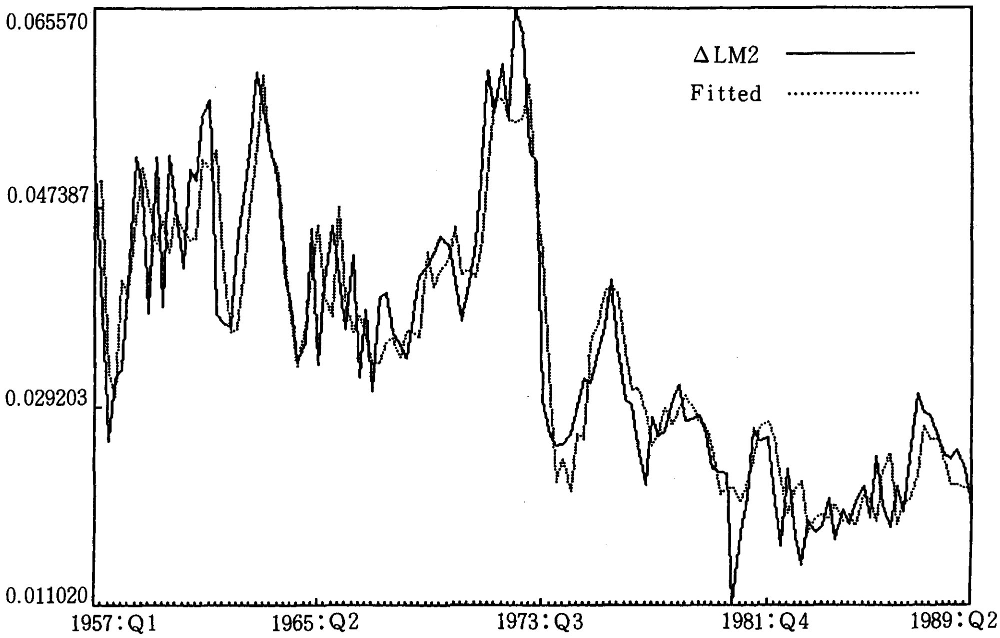
$R^2=0.845$　S.E＝0.00456　F (5,124)＝135.31　DW＝1.99
RSS＝0.0030466　for 6 Variables and 130 Observations
Normality χ^2 (2) ＝4.58 [0.10]　AR 1−5 F[5,119] ＝1.61 [0.16]
ARCH 4 F[4,116] ＝1.21 [0.31]　Xi∗Xj F[10,113] ＝1.90 [0.52]
RESET F[1,120] ＝0.22 [0.64]

（注）ΔLM2　：名目マネーサプライ前期比変化率
Δ^2 LGNPD　：インフレ率の加速度（＝前期比変化率の前期差）
Δ2SP　：（コールレートーM2＋CD自己金利）の２四半期前対比変化幅
EC　：*ln* (M2＋CD)－*ln* (GNPデフレータ) －1.2 *ln* (実質GNP)

いる。第A-1表に記されている各種診断テストの結果をみると、残差項に系列相関と分散不均一性、ARCH（Autoregressive Conditional Heteroskedasticity）は看取されず、正規性も満たされている。第A-1図は、M_3によるマネーサプライの推定値を現実値とともにプロットしたものである。

第A-2表は、M_3とM_1を比較したものであり、その左列は「M_3がDGPである」という帰無仮説をM_1に対して検定した結果、右列は「M_1がDGPである」という帰無仮説をM_3に対して検定した結果を示している。第A-2表左列をみると、検定量はすべて5％の棄却水準に達しておらず、帰無仮説は棄却されない――すなわちM_1はM_3とDGPを区別できないことがわかる。したがって$M_3 \mathcal{E} M_1$

第A-1図　ECM（M_3）による推計値と現実値



第A-2表　M_3（ECM）とM_1（PA）とのEncompassing Test

Test Statistic	M_3 against M_1		M_1 against M_3	
F-Test	F (4,120)	0.79 [0.47]	F (5,120)	39.03 [1.00]
W-Test		-0.74 [0.54]		-10.74 [1.00]
N-Test		-0.89 [0.63]		-20.46 [1.00]
NT-Test		-0.76 [0.55]		-18.82 [1.00]
J-Test		0.92 [0.64]		14.15 [1.00]
JA-Test		0.73 [0.54]		11.87 [1.00]

第A-3表　M_3（ECM）と M_2（NTD）との Encompassing Test

Test Statistic	M_3 against M_2		M_2 against M_3	
F-Test	F (4,120)	0.59 [0.33]	F (5,120)	72.75 [1.00]
W-Test		0.05 [0.04]		-14.29 [1.00]
N-Test		0.05 [0.04]		-32.16 [1.00]
NT-Test		-0.03 [0.02]		-29.17 [1.00]
J-Test		0.08 [0.06]		19.29 [1.00]
JA-Test		0.03 [0.02]		18.19 [1.00]

である。一方、第A-2表右列をみると、検定量はすべて棄却水準を大きく越えており、帰無仮説は棄却される――つまり M_3 は M_1 が DGP でないことを識別できることを示している。したがって $M_1 \mathcal{E} M_3$ であり、２つの検定結果を総合すると、M_1（PA）は M_3（ECM）に対して DGP に関するなんらの追加的情報を持たない redundant なモデルであると結論される。

次に第A-3表は、M_3 と M_2 を比較したものであり、その左列は「M_3 が DGP である」という帰無仮説を M_2 に対して検定した結果、右列は「M_2 が DGP である」という帰無仮説を M_3 に対して検定した結果を示している。この結果は第A-2表の場合と同様 $M_3 \mathcal{E} M_2$ である一方 $M_2 \sim \mathcal{E} M_3$ であり、M_2（NTD）も M_3 に対して redundant なモデルであると結論された。

以　上

［日本銀行金融研究所研究第一課副調査役（現営業局）］

【参考文献】

吉田知生、「通貨需要関数の安定性をめぐって－ＥＣＭ（Error Correction Model）による計測」、『金融研究』第８巻第３号、1989年

――――，「英国計量経済学の考え方」、研究資料(3)研１－９、日本銀行金融研究所、1991年

Ahumuda, H., "An Encompassing Tests of Two Models of the Balance of Trade for Argentina," *Oxford Bulletin of Economics and Statistics*, Vol. 47, No. 1, 1985, pp. 51-70.

Amemiya, T., *Advanced Econometrics*, Oxford: Basil Blakwell, 1985.

Atkinson, A.C., "A Method for Discriminating between Models," *Journal of the Royal Statistical Society*, Series B, Vol. 32, 1970, pp. 323-53.

Breusch, T.S., and A.R. Pagan, "A Simple Test for Heteroskedasticity and Random Coefficient Variation," *Econometrica*, Vol. 47, 1978, pp. 1287-94.

Chong, Y.Y., and D.F. Hendry, "Ecnometric Evaluation of Linear Macro-Economic Models," *Review of Economic Studies*, Vol. 53, 1986, pp. 671-90.

Cochrane, D., and G.H. Orcutt, "Application of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms," *Journal of the American Statistical Association*, Vol. 44, 1949, pp. 32-61.

Cox, D.R., "Tests of Separate Families of Hypotheses," in *Proceedings of the Forth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 1961, pp. 105-23.

———, "Further Results on Tests of Separate Families of Hypotheses," *Journal of the Royal Statistical Society*, Series B, 1962, pp. 406-24.

Cuthbertson, K., The Supply and Demand for Money, Oxford: Basil Blackwell, 1985.

Dastoor, N.K., "Some Aspects of Testing Non-nested Hypotheses," *Journal of Econometrics*, Vol. 21, 1983, pp. 213-28.

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo, "Economic Modelling of the Aggregate Time Series Relationship between Consumer's Expenditure and Income in the United Kingdom," *Economic Journal*, Vol. 88, 1978, pp. 661-92.

Davidson, R., and J.G. MacKinnon, "Several Tests for Model Specification in the Presence of Alternative Hypothesis," *Econometrica*, Vol. 49, No. 3, 1981, pp. 781-93.

———, and ———, "Some Non-Nested Hypotheses Tests and the Relations among Them", *Review of Economic Studies*, Vol. 49, 1982, pp. 551-65.

Dhrymes, P.J., E.P. Howrey, S.H. Hymans, J. Kmenta, E.E. Leamer, R.E. Quant, J.B. Ramsey, H.T. Shapiro, and V. Zarnowitz, "Criteria for Evaluation of Econometric Models," *Annals of Economic and Social Measurement*, Vol. 1, 1972, pp. 291-324.

Engle, R.F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, Vol. 50, No. 4, 1982a, pp. 987-1007.

———, "A General Approach to Laglange Multiplier Model Diagnostics," *Journal of Econometrics*, Vol. 20, 1982b, pp. 83-104.

———, "Wald, Likelihood Ratio and Laglange Multiplier Tests in Econometrics," in Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics*, Vol. 2, 1984.

Ericsson, N.R., "Asymptotic Properties of Instrumental Variables Statistics for Testing Non-Nested Hypotheses," *Review of Economic Studies*, Vol. 50, 1983, pp. 287-304.

Fisher, G.R., and M. McAleer, "On the Interpretation of the Cox Test in Econometrics," *Economics Letters*, Vol. 4, 1979, pp. 287-304.

———, and ———, "Alternative Procedures and Associated Tests of Significance for Non-Nested Hypotheses," *Journal of Econometrics*, Vol. 16, 1981, pp. 103-19.

Gilbert, C.L., "The Development of British Econometrics 1945-85," *University of Oxford Applied Economics Discussion Paper*, No. 8, 1986.

Godfrey, L.G., "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables," *Econometrica*, Vol. 46, 1978, pp. 1303-10.

———, "On the Uses of Misspecification Checks and Tests of Non-Nested Hypotheses in Empirical Econometrics," *Economic Journal Conference Papers*, Vol. 64, 1984, pp. 69-81.

———, and M.H. Pesaran, "Tests of Non-nested Regression Models: Small Sample Adjustments and Monte Carlo Evidence," *Journal of Econometrics*, Vol. 21, 1983, pp. 133-54.

Gourieroux, C., A. Monfort, and A. Trognon, "Testing Nested or Non-nested Hypotheses," *Journal of Econometrics*, Vol. 21, 1983, pp. 83-115.

———, ———, and ———, "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, Vol. 52, No. 3, 1984, pp. 681-700.

Grasa, A.A., *Econometric Model Selection: A New Approach*, Dordrecht: Kluwer Academic Publishers, 1989.

Gregory, A.W., and M.R. Veall, "Formulating Wald Tests of Nonlinear Restrictions," *Econometrica*, Vol. 53, No. 6, 1985, pp. 1465-68.

———, and ———, "Wald Tests of Common Factor Restrictions," *Economics Leters*, Vol. 22, 1986, pp. 203-208.

Harvey, A.C., *The Econometric Analysis of Time Series*, 2nd ed., London: Philip Allan, 1990.

Hendry, D.F., "Encompassing," *National Institute Economic Review*, August, 1988, pp. 88-92.

———, *PC-GIVE: An Interactive Economtric Modelling System*, Institute of Economics and Statistics, University of Oxford, 1989.

———, and G.E. Mizon, "Evaluating Dynamic Econometric Models by Encompassing the VAR," *University of Oxford Applied Economics Discussion Paper*, No. 102, 1990.

———, and J-F. Richard, "Recent Developments in the Theory of Encompassing," in B. Cornet, and H. Tulkens, eds., *Contributions to Operations Research end Economics: The Twenties Anniversary of CORE*, MIT Press, 1989.

Jarque, C.M. and A. K. Bera, "Efficient Tests for Normality, Homoskedastyicity and Serial Independence of Regression Residuals," *Economics Letters*, 6, 1980, pp. 255-59.

Lakatos, I., "Falsification and the Methodology of Scientific Research Programmes," in I. Lakatos, and A. Musgrave eds., *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press, 1970, pp. 91-196.（森博監訳、『批判と知識の成長』、木鐸社）

———, *The Methodology of Scientific Research Programmes*, Cambridge: Cambridge University Press, 1978.（村上陽一郎ほか訳、『方法の擁護：科学的研究プログラムの方法論』、新曜社）

Mackinnon, J.G., "Model Specification Tests against Non-Nested Alternatives," *Econometric Reviews*, Vol. 2, No. 1, 1983, pp. 85-110.

McAleer, M., "Specification Tests for Separate Models: A Survey," in M.L. King, and E.A. Giles, eds., *Specification Analysis in the Linear Model*, London: Routledge & Kegan Paul, 1987.

———, and G. Fisher, "Testing Separate Regression Models Subject to Specification Error," *Journal of Econometrics*, Vol. 19, 1982, pp. 125-45.

———, and M.H. Pesaran, "Statistical Inference in Non-nested Econometric Models," *Applied Mathematics and Computation*, Vol. 20, 1986, pp. 271-311.

Mizon, G.E., "The Encompassing Approach in Econometrics," in D.F. Hendry and K.F. Wallis, eds., *Econometrics and Quantitative Economics*, Oxford: Basil Blackwell, 1984.

———, and D.F. Hendry, "An Empirical Application and Monte Carlo Analysios of Tests of Dynamic Specification," *Review of Economic Studies*, Vol. 47, 1980, pp. 21-45.

———, and J-F. Richard, "The Encompassing Principle and Its Application to Testing Non-Nested Hypothesis," *Econometrica*, Vol. 54, 1986, pp. 657-78.

Muscatelli, V.A., "A Comparison of the Rational Expectations' and General-to-Specific' Approaches to Modelling the Demand for M1," *Oxford Bulletin of Economics and Statistics*, Vol. 51, No. 4, 1989, pp. 353-75.

Neyman, J. and E.S. Pearson, "On the Use and Interpretation of Certain Test Criteria for the Purposes of Statistical Inference," *Biometrika*, A, Vol. 20, 1928, pp. 175-240 (PartI), pp. 263-94 (PartII).

Pesaran, M.H., "On the General Problem of Model Selection," *Review of Economic Studies*, Vol. 41, 1974, pp. 153-71.

———, and A.S. Deaton, "Testing Non-nested Non-Linear Regression Models," *Econometrica*, Vol. 46, 1978, pp. 677-94.

Ramsey, J.B., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Royal Statistical Society*, Series B, Vol. 31, 1969, pp. 350-71.

Sargan, J.D., "Some Tests of Dynamic Specification for a Single Equation," *Econometrica*, Vol. 48., 1980, pp. 879-97.

Silvey, S., "The Laglangian Multiplier Test," *Annals of Mathematical Statistics*, Vol. 30, 1959, pp. 389-407.

Wald, A., "Tests of Hypotheses Concerning Several Parameters When the Numbers of Observations is Large," *Transactions of the American Mathematical Society*, Vol. 54, 1943, pp. 426-82.

White, H., "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica* 48, 1980, pp. 817-38.

White, H., "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, Vol. 50, 1982, pp. 1-25.

Yoshida, T., and R.H. Rasche, "M2 Demand in Japan: Shifted and Unstable?" *Monetary and Economic Studies*, Vol. 8, No. 2., Institute for Monetary and Economic Studies, Bank of Japan, 1990, pp. 9-30.