計量経済モデルの選択基準について : AIC利用上の問題点とその一つの解決法

林田, 実

https://doi.org/10.15017/2920687

出版情報：経済論究. 66, pp.19-32, 1986-12-05. 九州大学大学院経済学会
バージョン：
権利関係：

計量経済モデルの選択基準について

——AIC利用上の問題点とその一つの解決法——

林　田　　　実

目　次

序　　問題の所在

一般に，計量経済学の課題は，理論経済学によって導かれた理論仮説を確率モデル化し，(1)このモデルに，数理統計学の推定論を適用して，理論的仮説に数量的内容を付与すること，(2)仮説検定論をもちいて，理論的仮説の現実妥当性を検証すること，(3)仮説検定により，その妥当性が確認されたモデルを利用して，予測その他の現実的要請に答えることである注1。

計量経済学にかんする理論的研究の多くは，これまで(1)の「理論仮説に数量的内容を与える」ために，様々な推定法を開発し，さらに，それらの推定量の数理的性質を探求することに集中してきた。しかし，モデルの推定が，推定だけに終わるのであれば，そのこと自体は，さほど重要な意味を持つとは言えない。というのは，一つの経済事象について，複数の経済理論が存在し，それぞれに異なる計量経済モデルを設定するのが普通であるから，対象反映性が吟味されていないモデルは，他のモデルに比較して，現実妥当性が保証されているとは言えないからである。言うまでもなく，計量経済学において，モデルの妥当性を検証する方法は，仮説検定論である。したがって，モデルの推定が有効

であるためには，仮説検定が十分に機能していることが前提とされる。すなわち，同一の経済事象に関する複数の計量経済モデルの優劣を，データに照らして決定する基準として，仮説検定が有効でなければならない。しかし，先に述べたように，計量経済学の理論的研究の力点が推定論におかれていたために，現在の仮説検定論が，この要請に答えているとは言いがたい。たとえば，サローは，次のように述べている注2。

「方程式を正確に測るために用いられる標準的な統計尺度（標準誤差やt統計量）は，計量経済学の方程式への信頼性に，誤った評価を与えがちである。どのような二義的変数がその関係に含まれるべきかを正確に知らず，また用いられるべき正確な関数形がわからないため，方程式はふつう何度も推定される。単純でランダムな探査によって，計量経済学の研究者は“ベスト”な方程式を与えてくれる二義的変数と関数形との組合わせを求めるのである。

ここでいう“ベスト”な方程式は，分析者のあらかじめ抱いている信念に強く左右される。もし分析者が，利子率は貨幣の流通速度に影響しないと信じていれば，彼は自分の事前の信念を確証するような“ベスト”な方程式を見つけるのである。もし分析者が利子率は貨幣の流通速度に確かに影響を与える，と信じていれば，彼は，自分のこの事前の信念を確証する“ベスト”な方程式を見つけるだろう。両方の立場から“ベスト”の方程式が見つかる可能性があるとすれば，どちらの“ベスト”の方程式も相手方の“ベスト”の方程式が誤りであると説得することはできない」

冒頭に掲げた計量経済学の課題にそくしていえば，(2)の「仮説検定による理論的仮説の検証」が，成功したとは言いがたいのである。

次に，計量経済学の第3の課題である「予測，その他の現実的要請に答えること」に論を進めよう。「現実的要請」に答えるためには，予測の精度が確保されることが必要条件であるから，(3)にもりこまれている計量経済学の役割は，簡潔に「正確な予測を行なうこと」と言いかえてよかろう。さて，近年，時系列モデルの研究が進展するとともに，「予測力」において，計量経済モデルが時系列モデルに劣っていることが主張されるようになった。このような論

調のなかで，予測力によってモデルの比較を行なうこと自体が，無意味であると主張する識者もあらわれている注3。しかし，計量経済学の歴史が物語っているように，予測は計量経済学者の基本的関心事であったのだから，そのように予測を軽視する態度は，計量経済学の大きな後退と言わねばならない。ところで，計量経済モデルは，なぜ正確な予測をなしえないのであろうか。この問題に関して，赤池弘次は，次のように指摘している注4。

「コールズ研究所（初期の計量経済学の発展に多大の貢献をはたした研究所）の研究者等の研究には，このような視点（予測の観点からモデルを評価する視点）は明示的には，取入れられていなかった。このため係数推計の手法の高度の発達にもかかわらず，最終的なモデルの選択決定は極めて経験的なものとなっている」

氏の指摘を私なりに言いかえれば，次のようになる。計量経済モデル選択の基準としては，仮説検定論が存在しているわけであるが，その研究は推定論に比べて著しく立ち遅れており，しかも，そのなかには予測を基準としてモデルを選択するという問題意識は内包されていなかった，と。

こうして，計量経済学の大きな役割と考えられていた計量経済モデルによる理論仮説の検定，および予測は，十分に達成されたとは言えない。そして，その原因が，計量経済学における仮説検定論の立ち遅れにあったことは，すでに述べたとおりである。

本稿は，このような計量経済学の現状に一石を投ずるため，計量経済モデルの一つの選択基準を提示することを目的とする。第1節では，クラインモデルＩを用いて，仮説検定論の実際的不備がどのような形であらわれてくるのかを考察する。さらに，第2節では，計量経済モデルの選択基準として AIC を導入する。その際，若干の問題点を指摘し，これを解決する一つの方法を提示する。

第１節　計量経済学における仮説検定の問題点

今，対抗する理論１と理論２とがあった場合，これらの理論をデータに照ら

して，よりよい理論を取捨選択していくことがもっとも広い意味での仮説検定論と言えるであろう。本節では，クラインモデルＩ注5を素材として検定論の問題点を指摘する関係上，このような一般の場合でなく，次のような状況を考える。すなわち，計量経済モデルは，経済理論から導出される固定した構造とそれ以外の部分とからなっているとする。このような場合，データとの適合性の如何にかかわらず，経済理論によって導かれた構造は存続し，残りの部分はこの理論を前提したところのファクトファインディングに向けられることになろう。

さて，クラインモデルＩで，この固定した構造とみなされたものは，次のようなものであった。

モデルＩ

$$C = \alpha_0 + \alpha_1(W_1 + W_2) + \alpha_2 P + \mu_1 \qquad \text{消費関数}$$

$$I = \beta_0 + \beta_1 P + \beta_2 P_{-1} + \beta_3 K_{-1} + \mu_2 \qquad \text{投資関数}$$

$$W_1 = \gamma_0 + \gamma_1(Y + T - W_2) + \gamma_2(Y + T - W_2)_{-1} + \gamma_3(t - 1931) + \mu_3 \qquad \text{民間賃金関数}$$

$$Y + T = C + I + G \qquad \text{定義式}$$

$$Y = W_1 + W_2 + P \qquad \text{定義式}$$

$$\Delta K = I \qquad \text{定義式}$$

ここで，

C＝消費　I＝投資　W_1＝民間賃金　P＝利潤　Y＝国民所得　K＝期末資本ストック　W_2＝政府，公企業の賃金所得　T＝間接税　G＝政府支出　t＝歴年　である。

クラインはこの構造にさまざまなラグを導入したモデルを考え，これらのモデルの中から最良のモデルを選ぼうとした。その結果が次のモデルⅡである。

モデルⅡ

$$C = 16.78 + 0.80(W_1 + W_2) + 0.02P + 0.03P_{-1} + \mu_1'$$

$$I = 17.79 + 0.23P + 0.55P_{-1} - 0.15K_{-1} + \mu_2'$$

$$W_1 = 1.60 + 0.42(Y+T-W_2) + 0.16(Y+T-W_2)_{-1}$$
$$+ 0.13(t-1931) \quad + \mu_3'$$
$$Y+T = C+I+G$$
$$Y = P+W_1+W_2$$
$$\Delta K = I$$

ここで問題となるのは，モデルⅠに適当なラグを付け加え，最良のモデルを得る過程において，検定をくりかえし適用すれば，直線的にモデルⅡを導出することができるかということである。以下，いろいろな検定の手法の中でも，一般に使われている t 検定でこの問題をみてみよう。

まず，考察の対象となるモデルを次のように書く。なお定義式は省略する。

model(i,j,k,ℓ,m)

$$C = \alpha + \alpha_{10}(W_1+W_2)_0 + \cdots + \alpha_{1i}(W_1+W_2)_{-i}$$
$$+ \alpha_{20}P_0 + \cdots + \alpha_{2j}P_{-j} + \mu_1$$
$$I = \beta_0 + \beta_{10}P_0 + \cdots + \beta_{1k}P_{-k} + \beta_{21}K_{-1} + \cdots + \beta_{2\ell}K_{-\ell} + \mu_2$$
$$W_1 = \gamma_0 + \gamma_{10}(Y+T-W_2)_0 + \cdots + \gamma_{1m}(Y+T-W_2)_{-m}$$
$$\gamma_2(t-1931) + \mu_3$$

$i,j = 0,1,2 \quad k,\ell,m = 1,2$

この記述法によればモデルⅠは model (00111) モデルⅡは model (01111) ということになる。さて，後述の AIC によるモデル選択との関係上，制限情報最尤法によってモデルの推定を行なう。各係数について得られた t 比が 2 をこえると該当のパラメータを含んだモデルが選択されることになり，2 より小さいとそのパラメータを含まないモデルが取られることになる。表 1 は，その主な結果を示す。

表 1 の結果からすぐにわかることは 選択されたモデルが一つではなく，model (00111) と model (01111) との 2 つであることだ。これは仮説検定そのものが，帰無仮説下での分布を基礎にしているため，前提されたモデルの型によって，採択されるモデルが変わってくるためである。もっとも良いモデルが 1 個に定まらない，あるいは，一つのオーダーで，多数のモデルを評価できな

いという仮説検定の弱点は，モデル選択にあたって最大のネックとなる。この例のように，仮説検定を通過するモデルが複数ある場合には，結局，他の評価基準が必要となってしまう。

また，クラインモデルＩでは，model (00111) をデータとの適合性にかかわらず固定しているため，検定をクリアーしたモデルは２個であったが，この固定部分が小さい場合，検定を通過するモデルが多くなることは避けられない。たとえば model (01112) の検定結果は次のとおりである。ただし（　）内はｔ比である。

表１

推定されたモデル	選択されたモデル
0 1 1 1 1	0 1 1 1 1
0 2 1 1 1	0 0 1 1 1
0 1 1 1 2	0 1 1 1 1
0 1 2 1 1	0 1 1 1 1
0 2 2 1 1	0 0 1 1 1
0 2 1 1 2	0 0 1 1 1
0 1 2 1 2	0 1 1 1 1
0 1 1 2 1	0 1 1 1 1
0 2 2 1 2	0 0 1 1 1
0 1 1 2 2	0 1 1 1 1
0 2 1 2 1	0 0 1 1 1
0 1 2 2 1	0 1 1 1 1
0 2 1 2 2	0 0 1 1 1
0 2 2 2 1	0 0 1 1 1
0 1 2 2 2	0 1 1 1 1
0 2 2 2 2	0 0 1 1 1
1 1 1 1 1	0 0 1 1 1
0 0 1 1 1	0 0 1 1 1
0 0 1 1 2	0 0 1 1 1
0 0 2 1 1	0 0 1 1 1

注，表中の数字は model (i,j,k,ℓ,m) の i,j,k,ℓ,m を示す。

$$C = \underset{(7.6)}{17.6} + \underset{(12.2)}{0.8}(W_1+W_2) - \underset{(-1.0)}{0.2P} + \underset{(2.0)}{0.4P_{-1}} + \mu_1$$

$$I = \underset{(2.1)}{32.1} - \underset{(-0.2)}{0.1P} + \underset{(2.8)}{0.8P_{-1}} - \underset{(-3.0)}{0.2K_{-1}} + \mu_2$$

$$W_1 = \underset{(0.5)}{0.9} + \underset{(6.0)}{0.5}(Y+T-W_2) + \underset{(0.01)}{0.01}(Y+T-W_2)_{-1} + \underset{(1.0)}{0.07}(Y+T-W_2)_{-2} + \underset{(2.2)}{0.09}(t-1931) + \mu_3$$

ｔが２をこえるか否かでパラメータの採否を決めると消費関数，投資関数におけるＰ，賃金決定関数における定数項，$(Y+T-W_2)_{-1}$，$(Y+T-W_2)_{-2}$ を落としたモデルが選択されることになる。このような事態を避け，モデルの選択幅を縮小するためには，この固定した部分を比較的大きく考える必要が出てくる。これが，モデル作成上の恣意性の問題となって表面化してくる可能性を

否定することはできない。

さて，一般に実証的モデル作成の作業は次のような段階をへて進められる注6。

(1) 各構造方程式を想定し係数を推定する。係数推定値の経済的意味，有意性，部分テストを経て各式の「モデル選択」を進める。

(2) システムに含まれる各式の推定結果が得られれば，全体テストによりシステム全体としての性能を調べる。

(3) 全体テストが終われば，最終テストにより，観測期間中でのシステムによる予測の精度を調べる。最終テストが満足すべきものであれば，乗数分析，政策シミュレーション，ショックテスト，そして予測などの経済的な分析を行なう。

ところが先にみたように，クラインモデルⅠの場合すでに(1)の段階で（厳密には(1)の検定の段階で）2つのモデルが候補としてのぼっており，これらをさらに⑥⑦の手続へまわして最終的なモデルを得なければならない。モデルが大規模になれば，すべてのモデルに(1)〜(3)の手続をとることは不可能に近い。その主原因は先にも述べたとおり，検定の段階で十分なモデル選択がおこなわれなかったからである。しかし，より本質的な問題としては，検定論の不備の故に，モデル選択の過程として，(1)〜(3)の手続が定着しているとみなせることである。さらに，全体テスト，最終テストにしても，その結果がよければ，モデルも良いと判断することは極めて経験的なものである（その直観的意味づけを疑うものではないけれども）。赤池が指摘するように「最終的なモデル選択」は実に経験的なものになってしまっている。

第2節　AIC による構造の決定

構造方程式を，一般に次のように書く。

$$\mathbf{B}\mathbf{y}_t + + \Gamma\mathbf{x}_t = \boldsymbol{\mu}_t \quad —— \quad ①$$

ただし，

$$\mathbf{y}_t=\begin{bmatrix} y_{1t} \\ \vdots \\ y_{Gt} \end{bmatrix}$$ $y_{1t}\sim y_{Gt}$ はt期の内生変数

$$\mathbf{x}_t=\begin{bmatrix} x_{1t} \\ \vdots \\ x_{Kt} \end{bmatrix}$$ $x_{1t}\sim x_{Kt}$ はt期の先決変数

$$\boldsymbol{\mu}_t=\begin{bmatrix} \mu_{1t} \\ \vdots \\ \mu_{Gt} \end{bmatrix}$$ $\mu_{1t}\sim\mu_{Gt}$ はt期の撹乱項

$\boldsymbol{\mu}_t\sim N(\theta,\Sigma)$　　$t=1, 2, \cdots, N$　　　である。

また①式の第一方程式は次のようにも書かれるとする。

$$y_{1t}=\sum_{i=1}^{G_1}\beta_i y_{i+1t}+\sum_{j=1}^{K_1}\gamma_j x_{jt}+\mu_{1t} \quad —— \quad ②$$

以上の前提のもとで，データ（$\mathbf{y}_t$ $\mathbf{x}_t$）が独立に N コ与えられたとする。この時，制限情報最尤法を②式に適用して，その AIC を求めると結局，

$$AIC=N\log\lambda-2(G_1+K_1)$$

となる。ここで λ は最小分散比である。すなわち，λ は，

$$L=\frac{\boldsymbol{\beta}_\Delta' W_{\Delta\Delta}^* \boldsymbol{\beta}_\Delta}{\boldsymbol{\beta}_\Delta' W_{\Delta\Delta} \boldsymbol{\beta}_\Delta} \quad —— \quad ③$$

を $\boldsymbol{\beta}_\Delta$ の関数とみた時の最小値である。ここで，

$$\boldsymbol{\beta}_\Delta=\begin{bmatrix} 1 \\ -\beta_1 \\ \vdots \\ -\beta_{G_1} \end{bmatrix}$$

$$W_{\Delta\Delta}^*=Y_\Delta' Y_\Delta-Y_\Delta' X_1(X_1' X_1)^{-1}X_1' Y_\Delta \quad —— \quad ④$$

$$W_{\Delta\Delta}=Y_\Delta' Y_\Delta-Y_\Delta' X(X'X)^{-1}X' Y_\Delta \quad —— \quad ⑤$$

$$Y_\Delta=\begin{bmatrix} y_{11} & y_{21} \cdots y_{G_1+1\ 1} \\ \vdots & \vdots \\ y_{1N} & y_{2N} \cdots y_{G_1+1\ N} \end{bmatrix}$$ ←第一期のデータ ←第N期のデータ

$$X_1=\begin{bmatrix} x_{11} & x_{21} \cdots x_{K_1\ 1} \\ \vdots & \vdots \\ x_{1N} & x_{2N} \quad x_{K_1\ N} \end{bmatrix}$$ ←第一期のデータ ←第N期のデータ

$$X=\begin{bmatrix} x_{11} & \cdots & x_{K1} \\ \vdots & & \vdots \\ x_{1N} & \cdots & x_{KN} \end{bmatrix} \begin{array}{l} \leftarrow \text{第一期のデータ} \\ \\ \leftarrow \text{第N期のデータ} \end{array}$$

③④⑤から明らかなように，λ は β_Δ，Y_Δ，X_1，X に依存している。このうち β_Δ，Y_Δ，X_1 は第一方程式の特定化に際して同時に決定されるのに対して，X は他の方程式に，どのような先決変数を説明変数としてもってくるかによって変化する。一方，AIC は，

$$AIC=N\log\lambda-2(G_1+K_1)$$

だった。(G_1+K_1) も第一方程式の特定化によって定まるので，AIC は第一方程式の特定化と X，すなわち他の方程式に含まれる先決変数に依存して決まってくる。したがって，ある特定化をなされた第一方程式が AIC 基準でもっとも良いと判断されたとしても，それは，他の方程式の悪さをもたらす場合がある。これをクラインモデル I を使って例証してみよう。表 2 がそれである。

たとえば，model (11211) と model (00212) とを比較してみよう。消費関数についていえば，model (11211) の AIC が 20．1，model (00212) の AIC が 26.1と model (11211) がはるかに良い。しかし，投資関数をみると，前者の AIC が 17.4，後者の AIC が，11.5 と逆に，model (00212) の方が model (11211) をしのいでいる。このような事例は，表 2 の随所にみられる。この問題は，連立方程式モデルに AIC を適用する際の基本的問題とい

表 2

MODEL (i j k ℓ m)	AIC-C	AIC-I	AIC-W1
1 1 2 1 1	20.1	17.4	22.4
0 0 2 1 2	26.1	11.5	22.6
1 2 1 1 1	22.1	15.8	22.4
0 0 1 2 1	26.9	11.6	22.2
0 0 1 2 2	26.2	11.9	23.0
1 2 2 1 1	22.1	17.4	22.4
0 0 2 2 1	26.2	13.7	22.5
1 0 1 1 1	25.9	14.8	22.2
0 0 2 2 2	26.2	13.9	23.4
1 1 1 2 1	20.8	21.2	22.3
1 1 1 1 2	22.3	21.3	23.1
2 1 1 1 1	25.2	18.6	23.2
1 0 2 1 1	27.9	17.4	22.4
2 1 2 1 1	25.2	20.3	23.2
1 1 1 2 2	22.5	23.2	23.2
1 2 1 2 1	24.0	22.4	22.5
1 1 2 2 1	22.2	24.3	22.5
2 2 1 1 1	27.2	18.6	23.2
1 1 2 1 2	22.3	23.5	23.3
1 2 1 1 2	24.0	21.9	23.3

注，AIC-C，AIC-I，AIC-W_1 はそれぞれ消費関数，投資関数，民間賃金関数の AIC を示す。

えよう。この問題を解決するためには，単一方程式の良さを示すために AIC を用いるのではなく，計量経済モデル全体の良さの基準として AIC を導入することが望ましい。このような考え方は，連立方程式モデル自体が全体として一つの意味を持つように構築されているのだから，きわめて自然な発想と言ってよかろう。具体的には，全ての構造方程式の AIC の算術平均でもって，諸計量経済モデル間の比較基準とする。クラインモデルでは，

$$\mathrm{AIC}=\frac{1}{3}(\mathrm{AIC\_C}+\mathrm{AIC\_I}+\mathrm{AIC\_W_1})$$

となる。表3は，この基準でもっとも良いモデルを上から20コ選んだものである。周知のように，AIC はその絶対値ではなく，モデル間の差に意味があるので，便宜のため，

$$\mathrm{Criterion}=\frac{1}{3}[(\mathrm{AIC\_C}-\min(\mathrm{AIC\_C}))+(\mathrm{AIC\_}-\min(\mathrm{AIC\_I}))+(\mathrm{AIC\_W_1}-\min(\mathrm{AIC\_W_1}))]$$

を表中に入れた。

クラインが最終的に選んだ model (01111) が，第1位を占めている。おそらく，クラインは model (01111) に到達するまでに，様々な検定をくり返さなければならなかったであろう。しかし，AIC を利用すれば，このように，実に簡単に，model (01111) を

表3

MODEL (i j k ℓ m)	Criterion	AIC_C	AIC_I	AIC_W1
1 1 1 1	0.00000	15.8	8.7	22.2
2 1 1 1	0.43333	16.8	8.8	22.4
1 1 1 2	0.53333	16.6	9.1	22.6
1 2 1 1	0.70000	16.0	10.4	22.4
2 2 1 1	0.96667	16.8	10.4	22.4
2 1 1 2	1.13333	17.6	9.9	22.6
1 2 1 2	1.36667	16.7	11.5	22.6
1 1 2 1	1.40000	17.1	11.6	22.2
2 2 1 2	1.66667	17.6	11.5	22.6
1 1 2 2	1.80000	17.2	11.9	23.0
2 1 2 1	1.96667	18.3	11.8	22.5
1 2 2 1	2.23333	17.2	13.7	22.5
2 1 2 2	2.30000	18.3	11.9	23.4
2 2 2 1	2.60000	18.3	13.7	22.5
1 2 2 2	2.60000	17.2	13.9	23.4
2 2 2 2	2.96667	18.3	13.9	23.4
1 1 1 1 1	2.96667	18.6	14.8	22.2
1 1 1	3.16667	25.3	8.7	22.2
1 1 2	3.70000	26.1	9.1	22.6
2 1 1	3.90000	25.6	10.4	22.4

最良のモデルとして選択できるのである。これは，第2節で明らかにしたように，従来の検定論では，望むべくもなかったことである。また，第2節の t 検定では，model (01111) と model (00111) とが，採択すべきモデルとして残されたものであるが，表3をみると，この二つのモデルの間に model (02111) から model (11111) までの15コのモデルが含まれている。すなわち，固定した構造とされた，model (00111) より良いと判断されたモデルが model (01111) を除いて15コも存在している。このように，きめの細い統計的情報が与えられれば，これらのモデルを実質科学としての経済学にひきもどして考える際にこの情報が非常に有効なものとなろう。統計理論的意味の明白な一つの基準 (AIC) で全ての計量経済モデルを一度に評価できることは，経済学的に構造方程式を比較・評価する際，あいまいな統計情報によって無用な混乱に落ち入らないことを意味している。

さて，構造方程式全体を評価する基準として，従来用いられてきたものは，全体テストと最終テストである。今，構造方程式を次のように書く。

$$\mathbf{B}\mathbf{y}+\mathbf{A}\mathbf{y}_{-p}+\Gamma\mathbf{x}=\boldsymbol{\mu}$$

①式と異なるのは，$\mathbf{y}_{-p}$ がラグ付内生変数 (最大ラグ p)，$\mathbf{x}$ が外生変数であることだ。これを変形して，

$$\mathbf{y}=-\mathbf{B}^{-1}\mathbf{A}\mathbf{y}_{-p}-\mathbf{B}^{-1}\Gamma\mathbf{x}+\mathbf{B}^{-1}\boldsymbol{\mu}$$

ここで，$\mathbf{u}$ を無視し，$\mathbf{y}_{-p}$ に現実値を入れて $\mathbf{y}$ の予測値を出し，この予測値を現実値と比較するのが全体テストである。これにたいして，最終テストは $\mathbf{y}_{-p}$ に $\mathbf{y}$ の予測値を入れて，$\mathbf{y}$ を計算し，これと現実値とを比較するものである。その名のとおり，最終テストの結果が，モデル選択の最終的判断基準とされる。さて，クラインモデルⅠで考察の対象となっているモデルは，内生変数のラグ構造の異なるものであるから，この最終テスト，全体テストの結果によるモデルの優劣と AIC 基準のモデルの良し悪しとを比較することは興味深い。そこで，最終テスト，全体テストによる評価基準として，表2に含まれる全モデルに関して，予測値と現実値との差の絶対値を各々の内生変数について求

表4

MODEL (i j k ℓ m)	Criterion	FINALTEST	TOTAL TEST	AIC_N	FINAL_N	TOTAL_N
1 1 1 1	0.00000	3.25799	1.62598	1	2	17
2 1 1 1	0.43333	3.42724	1.58267	2	6	13
1 1 1 2	0.53333	3.24534	1.58672	3	1	14
1 2 1 1	0.70000	3.40090	1.56466	4	5	8
2 2 1 1	0.96667	3.56760	1.59780	5	13	15
2 1 1 2	1.13333	3.38671	1.57858	6	4	11
1 2 1 2	1.36667	3.36525	1.57765	7	3	10
1 1 2 1	1.40000	3.52229	1.53496	8	12	2
2 2 1 2	1.66667	3.51724	1.58187	9	10	12
1 1 2 2	1.80000	3.49197	1.54893	10	8	3
2 1 2 1	1.96667	3.67800	1.57153	11	18	9
1 2 2 1	2.23333	3.50363	1.53432	12	9	1
2 1 2 2	2.30000	3.63287	1.55622	13	15	6
2 2 2 1	2.60000	3.65099	1.56340	14	17	7
1 2 2 2	2.60000	3.46642	1.55405	15	7	4
2 2 2 2	2.96667	3.60635	1.55477	16	14	5
1 1 1 1 1	2.96667	3.52102	1.59993	17	11	16
1 1 1	3.16667	3.64351	1.70228	18	16	19
1 1 2	3.70000	3.68933	1.71418	19	19	20
2 1 1	3.90000	3.82643	1.67271	20	20	18

め，これらの総和を出す。この総和を（データ数×内生変数の個数）で割ったものを両テストの評価基準とする。表4の finaltest, totaltest の欄が各々，最終テスト，全体テストによるこれらの値である。final-n，total-n は，それぞれの基準によるモデルの順位を表わす。

全体テストによる順位は，AIC による順位とむしろ逆の様相を示しているが，最終テストによる順位と AIC による順位とは，順位相関係数を出してみると0.77と高い相関を示している。したがって，これまでのように最終テストにおいて，良いモデルを得るという基準に照らしてみても，AIC がその目安となりうることが期待される。

第3節　結びにかえて

第1節で展開した計量経準学における検定論の立ち遅れは，実は，計量経済学に特有のものではない。それは，検定論一般に内在する本質的な問題点なのである。計量経済学の場合この問題が増幅された形ででてくるので，その点がふつうの回帰分析と異なる所といえるかもしれない。それはそれとして，計量経済モデルにおいては，同時方程式体系という形をとるために，通常の回帰モデルで使われている t 検定やF検定などがどの程度まで適用可能であるのか，実は，はっきりしていない。この点にかんして森棟は次のように述べている注7。

「計量モデルの検定法は，従来もっとも開発の遅れていた分野で，望ましい検定法が何なのか，また，どういうわけで望ましい検定法なのか，一般的に明らかにされていなかった。そのためか，線型回帰式で常套的に使われているF比検定法や t 比検定法が，同時方程式モデルの分析でも，モデルの複雑さを無視して，いわば盲目的に利用されてきた」

こうした現状を打開するために，従来の検定論の枠組の中でもモデル選択基準の研究が進められるようになった。今後の研究の進展が期待される。一方，AIC を計量経済モデルに適用する試みはまだ始まったばかりである。この方向で研究を進めていくには，連立方程式モデルにおけるいわゆる modeling の問題，さらに，計量経済学に特有な推定量の AIC との関係等，未解決な問題が多い。これらを考究することが筆者の今後の課題である。

注1.　計量経済学の役割については以下の箇所によっている。経済学辞典　第2版　岩波書店　1979年 p. 323～p. 325「計量経済学」

注2.　サロー『デンジャラスカレンツ』p. 175　東洋経済新報社　1983.

注3.　たとえば，佐和隆光は次のように述べている。

「マクロ計量モデルにしろ，時系列モデルにしろ，現実経済にたいする，ひとつの幻影を，統計的モデルによって記述したにすぎず，“客観的”なデータにもとづいて，両者に甲乙をつけることは不可能なばかりか，“予測力”で勝負をつけるなどというのは的外れもはなはだしい。」

佐和隆光「マクロ計量モデルの有効性」『経済セミナー』1980年，2月号。

注4．赤池弘次 「時系列モデル批判に反論する」『経済セミナー』1980年　5月号。( )内は筆者。

注5．L. R. Klein "Econmic Fluctuations in the United States 1921-1941" Cowles Commission for Research in Economics Monograph No. 11, 1950年。

注6．この部分は次の箇所によっている。
森棟公夫『経済モデルの推定と検定』p. 16，共立出版，1985年。

注7．『経済モデルの推定と検定』p. 175。

参考文献

前掲の著書，論文は省略する。

Akaike, H. "Information theory and an extention of the maximum likelihood principle", 2nd Inter. Symp. on Information Theory (Petrov, B. N. and Csaki, F. eds.), Akademiai Kidao, Budapest 1973年。

赤池弘次「情報量基準 AIC とは何か」『数理科学』No. 153　1976年。

同「統計的検定の新しい考え方」『数理科学』No. 198 1979年。

同「エントロピーとモデルの尤度」『日本物理学会誌』第35巻，第7号 1980年。

同「モデルによってデータを測る」『数理科学』No. 213 1981年。

坂本慶行，石黒真木夫，北川源四郎『情報量統計学』共立出版 1983年。

L. R. Klein 著 佐和隆光訳『経済予測の理論』筑摩書房 1973年。

Kimio Morimune and Takamitsu Sawa.
"Decision Rules For The Choic to Structual Equations" Journal to Econometrics 14 1980年.

パネル・ディスカッション：実証分析の方法
—計量経済分析と時系列分析—　金融研究資料　第6号，1981年，日本銀行特別研究室。