

Title	動学的特定化におけるエラー修正モデルの意義：モンテカルロ実験による比較
Sub Title	The significance of error correction model in dynamic specification
Author	大津, 泰介
Publisher	慶應義塾経済学会
Publication year	1999
Jtitle	三田学会雑誌 (Keio journal of economics). Vol.92, No.2 (1999. 7) ,p.415(175)- 434(194)
JaLC DOI	10.14991/001.19990701-0175
Abstract	
Notes	論説
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-19990701-0175

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

動学的特定化におけるエラー修正モデルの意義*

—— モンテカルロ実験による比較 ——

大 津 泰 介

1. 序

本稿は、Sargan (1964) を嚆矢とし、Davison *et al.* (1978) 以降マクロ経済時系列のモデリングの手法として実証分析において広範に応用されるに至ったエラー修正モデル (Error Correction Model: 以下 ECM) の意義を、特に実証分析における動学的特定化のツールとしての有用性に注目して、1) 分布ラグモデルとしての一般性、2) 同等の一般性を持つモデルと比較した場合の多重共線性の低さ、という2つの視点から分析することを目的としている。

ECM の意義は以下のようにまとめることができる。

1. 経済理論から得られる経済関係式を長期関係式として表現し、同時に、データの一時的構造を反映した短期の動学的調整過程もモデル化している。したがって、長期と短期の乗数が同時にかつ容易に計測できる (養谷 (1996))。
2. ECM は均衡への調整方程式として解釈することができ、様々な調整費用を考慮した損失関数から最適調整ルールとして導出できる (Nickell (1985))。
3. 動学的特定化に関して、最も一般的な分布ラグモデルである一般自己回帰分布ラグ (Autoregressive Distributed Lag: 以下 ADL) モデルと同型 (Isomorphic) であり、関心あるパラメータについて等しい推定量をもたらす。そのため、「一般から特殊へ (General to Specific)」の特定化を重視する LSE (London School of Economics) アプローチの方法論と整合的なモデルである (Banerjee *et al.* (1993))。

* 本稿は、筆者の修士論文「実証分析におけるエラー修正モデルの意義」(平成9年度)をもとに作成したものである。論文作成に際しては、養谷千風彦教授、神谷傳造教授、および伊藤幹夫助教授に貴重なコメントをいただいた。また、本誌のレフリーからも有益なコメントをいただいた。ここに記して感謝したい。

4. 均衡の統計学的表現である共和分と深い関係を持ち、共和分関係を有するシステムであることと、(ベクトル) ECM で表現可能であることは同値であることが示されている (表現定理) (Engle and Granger (1987))。
5. 共和分関係が存在し、ECM に現れる変数が全て定常過程になるならば、見せかけの回帰 (Spurious Regression) (Granger and Newbold (1974)) の問題を回避して従来の仮説検定が可能になる。
6. 説明変数に差分を多く用いているので、ADL モデルと比べても (同型であるにもかかわらず) 多重共線性の影響が小さい。したがって、動学的特定化の便法として望ましい方法である (本稿)。

これらの意義は、モデル化する変数の非定常性を考慮に入れない LSE アプローチ本来の動学的特定化の方法としての意義 (1, 2, 3, 6) と、非定常過程の時系列における共和分ベクトルを含むモデルの推定方法としての意義 (4, 5)、という 2 つの範疇に分けられ、前者は構造方程式アプローチ、後者は時系列アプローチの視点から見た ECM の意義として整理できる。

本稿では、これらの意義の中でも、特に 3 の分布ラグモデルとしての一般性 (2 節) と、6 の多重共線性の影響の低さ (3 節) について扱う。どちらもモデルに含む変数の定常性の有無に依存しない問題であるが、本稿では定常過程について分析を行う。

この研究に関する先行研究の系譜は以下のようにまとめることができる。まず、分布ラグモデルとして長期乗数と短期乗数を別個に一つのモデルから得ることができるよう十分なラグ構造を持っている、という意味から ECM の重要性を分析した研究として、Hendry, Pagan and Sargan (1984) がある。Hendry, Pagan and Sargan (1984) は、最も一般的な分布ラグモデルである ADL モデル⁽¹⁾と、それにパラメータ制約を加えて得られるモデルについて包括的に議論し、ADL と同程度の一般性を持つモデルとして ECM の一般性を強調した。また、ADL モデル、ECM の他にも、長期乗数を計測する上での特定化の方法として、Bewley (1979) や Bårdsen (1989) によって様々なモデルが提案された。

これらのモデルを比較した研究として、ADL モデルと Bewley モデルの、関心あるパラメータ推定値に対する同等性を示し、ECM と比べても Bewley モデルの方が実証分析の便宜上望ましいと論じた、Wickens and Breusch (1988) や、ECM の便宜上の利点を強調しつつも、モンテカル

(1) 一般的であるとはいえ、ADL モデルは、VARMA モデルに外生変数ベクトルを加えた VAR-MAX モデルに比べれば、MA 部分を捨象しているので制約的である。しかし、LSE アプローチにおいては、移動平均項のような誤差項のダイナミックスに頼らずに、モデル化する変数による系統的な部分でのモデリングを重視する。したがって、その場合には ADL モデルが最も一般的なモデルとなる。

ロ実験の結果から、共和分ベクトルを推定する際のバイアスの問題があることを論じた、Banerjee, Galbraith and Dolado (1990) がある。また、ECM, Bårdsen 型の ECM, Bewley, 3 つのモデルの統計的性質について整理したものとしては、Banerjee *et al.* (1993) がある。しかし、これらのモデルについて、小標本における性質を比較した研究例としては、わずかに、ADL モデルと ECM の最小 2 乗 (OLS) 推定量における小標本バイアスについて比較した、Kiviet and Phillips (1994) を挙げることができるのみであり、Bårdsen 型の ECM, Bewley モデルを含めた比較はなされていない。本稿では、これらのモデルを多重共線性の観点からモンテカルロ実験によって比較する。

本稿は以下のように構成される。まず、2 節では、ECM の分布ラグモデルとしての一般性について、ECM が部分調整モデルや有限分布ラグモデル等をはじめとする様々なモデルを特殊ケースとして含む一般的なモデルであることを示し、データ発生機構 (Data Generation Process: 以下 DGP) が ECM と同型の ADL モデルの場合についてモンテカルロ実験を行い、特定化を誤った特殊ケースとして得られる各モデルのパフォーマンスを検討する。つぎに、3 節では、ECM と同等の一般性を持つ 3 つのモデル (ADL モデル, Bårdsen 型の ECM, Bewley モデル) を導入し、これらのモデルの特性と、パラメータ推定量および関心あるパラメータについての分散共分散行列の同等性について述べた後に、モデル特定化の上での便宜上の問題として、多重共線性に着目し、モンテカルロ実験によって比較する。実験の結果から、真のラグ構造が分からない状況においては多重共線性の影響が小さい Bårdsen 型の ECM が特定化のツールとして望ましいことが示される。4 節では、結論を述べるとともに、「一般から特殊へ」のアプローチの有効性について言及する。

2. 分布ラグモデルとしての一般性

本節では、Hendry, Pagan and Sargan (1984) および Hendry (1995) の分類をもとに、ADL モデルと ADL モデルにパラメータ制約を与えたモデルとして導出される、DHSY 型の ECM を含む様々な分布ラグモデルを整理し、モンテカルロ実験によって各モデルのパフォーマンスを評価する。

本節では最も単純な ADL (1,1) モデルについて分析する。このモデルは以下のように示される。

$$\text{ADL: } y_t = \beta_0 + \beta_1 x_t + \beta_2 y_{t-1} + \beta_3 x_{t-1} + \varepsilon_t \quad (1)$$

両辺から y_{t-1} を引き、パラメータ再編を行うと、以下の ECM がえられる。

$$\text{ECM: } \Delta y_t = \beta_0 + \beta_1 \Delta x_t + (\beta_2 - 1)(y_{t-1} - Kx_{t-1}) + \varepsilon_t \quad \text{ただし, } K = \frac{\beta_1 + \beta_3}{1 - \beta_2} \quad (2)$$

したがって、ADL と ECM は線形変換によって導出可能な同型 (Isomorphic) な関係にある。また、 y_t, x_t がそれぞれ均衡状態にあり、一定値 (y, x) をとる時の静学的長期関係式は、

$$y = \frac{\beta_0}{1 - \beta_2} + Kx$$

となり、 K は長期乗数となる。また、(2)式より $(\beta_2 - 1)$ は前期の長期関係式からの乖離として生じるエラーに対する調整係数として、 β_1 は x_t の変動に対する短期乗数として解釈することができる。

このように、(2)式は、モデルの直観的解釈は明確であるが、パラメータ K のために直接推定することはできない。そのために、このモデルの推定方法は大きく2つに分けられる。一つは、共和分ベクトルとして K を推定してそれを(2)式に代入してECMを推定するという2段階推定の方法(Engle and Granger (1987))であるが、この方法については本稿では扱わない。もう一つは、(2)式の右辺を直接推定できるようにパラメータ再編を行う方法であり、再編の方法に応じてECMの他にも様々なモデルが導出される。これらのモデルの比較は次節で行う。本節では、(2)式にパラメータ制約を課して導出されるモデルについて分析する。

例えば、(2)式に、 $\beta_1 + \beta_2 + \beta_3 = 1$ という制約を課せば、DHSY型のECMがえられる。

$$\text{DHSY型ECM: } \Delta y_t = \beta_0 + (\beta_2 - 1)(y_{t-1} - x_{t-1}) + \beta_1 \Delta x_t + \varepsilon_t \quad (3)$$

このモデルの静学的長期関係式は、

$$y = \frac{\beta_0}{1 - \beta_2} + x$$

となり、 y, x が対数値であれば、長期弾力性が1であるという制約を加えたモデルであると解釈できる。

この他にもADLモデルやECMにパラメータ制約を課した特殊ケースとして様々なモデルが導出できるが、これらのモデルは表1のように整理できる。ただし、定数項 β_0 は省略する。

以下では、表1に挙げた制約的なモデルのパフォーマンスに注目してモンテカルロ実験を行う。実験に用いるDGPはADL(1,1)モデルをもとに定式化し、DGPがADL(1,1)モデルであるにもかかわらず、特定化を誤って制約的なモデルを用いてモデルを推定した場合に、いかなる問題が生じるかを明らかにする。この分析では、主に、1) 長期乗数(実験のDGPでは $\frac{\beta_1 + \beta_3}{1 - \beta_2} = 0.4$)の推定値のバイアス、2) モデルの標準誤差推定値(実験のDGPでは1)のバイアス、3) 誤差項の1階の自己相関の検定統計量であるDurbin-Watson(DW)比(説明変数が被説明変数のラグ変数を含む時は、Durbin (1970) のh統計量)、4) 特定化の誤りを検定するRESET(2)テスト(Ramsey (1969))、等に注目してモデルの妥当性を考察する。

(2)
実験に用いたDGPは以下のとおりである。

$$y_t = 0.25 + 0.5x_t + 0.75y_{t-1} - 0.4x_{t-1} + \varepsilon_t \quad (4)$$

$$x_t = 0.25 + 0.75x_{t-1} + v_t \quad \text{ただし、} \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \sim \text{IN} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \quad (5)$$

表 1 : 分布ラグモデルの分類

	名前	式	制約
M_0	ADL	$y_t = \beta_1 x_t + \beta_2 y_{t-1} + \beta_3 x_{t-1} + \varepsilon_t$	制約なし
M_1	静学回帰	$y_t = \beta_1 x_t + \varepsilon_t$	$\beta_2 = \beta_3 = 0$
M_2	時系列	$y_t = \beta_2 y_{t-1} + \varepsilon_t$	$\beta_1 = \beta_3 = 0$
M_3	差分	$\Delta y_t = \beta_1 \Delta x_t + \varepsilon_t$	$\beta_2 = 1, \beta_1 = -\beta_3$
M_4	先行指標	$y_t = \beta_3 x_{t-1} + \varepsilon_t$	$\beta_1 = \beta_2 = 0$
M_5	部分調整	$y_t = \beta_1 x_t + \beta_2 y_{t-1} + \varepsilon_t$	$\beta_3 = 0$
M_6	COMFAC	$y_t = \beta_1 x_t + u_t, u_t = \beta_2 u_{t-1} + \varepsilon_t$	$\beta_3 = -\frac{\beta_1}{\beta_2}$
M_7	有限分布ラグ	$y_t = \beta_1 x_t + \beta_3 x_{t-1} + \varepsilon_t$	$\beta_2 = 0$
M_8	Dead Start	$y_t = \beta_2 y_{t-1} + \beta_3 x_{t-1} + \varepsilon_t$	$\beta_1 = 0$
M_9	DHSY	$\Delta y_t = (\beta_2 - 1)(y_{t-1} - x_{t-1}) + \beta_1 \Delta x_t + \varepsilon_t$	$\sum_i \beta_i = 1$
M_{10}	ECM	$\Delta y_t = (\beta_2 - 1)(y_{t-1} - K_1 x_{t-1}) + \beta_1 \Delta x_t + \varepsilon_t$	制約なし

なお、データは初期値の影響を避けるために、100個の y_t, x_t のデータを発生させた後に最初の20個を捨てて残りの80個を用いて推定し、これを10000回繰り返した。表2はその結果の要約である。

表 2 : 特定化を誤った制約的モデルのモンテカルロ実験の結果

M_1 : 静学回帰

	パラメータ	t 値	標準誤差	DW 比	RESET (2)
最小値	-0.6003	-3.4973	0.8068	0.1784	-5.4963
最大値	1.2832	13.3507	2.4589	1.4513	6.5693
中央値	0.4683	4.2239	1.4034	0.5972	-0.0062
平均	0.4668	4.2893	1.4216	0.6147	0.0017
標準偏差	0.2089	2.0701	0.2088	0.1697	1.4138
バイアス (β_1)	-0.0332		0.4216		
長期バイアス	0.0668				

M_2 : 時系列

	パラメータ	t 値	標準誤差	DW 比	RESET (2)	Durbin's h
最小値	0.2628	2.3621	0.8020	1.3723	-4.0419	-3.7066
最大値	0.9014	17.9056	1.5154	2.5070	3.9001	4.6501
中央値	0.7078	8.8418	1.1132	1.9399	0.0112	0.2242
平均	0.6996	8.9829	1.1134	1.9398	0.0061	0.2276
標準偏差	0.0823	2.0527	0.0898	0.1590	0.9014	0.9865
バイアス (β_2)	-0.0504		0.1134			

- (2) このモデルは、Hendry (1995) をもとに、DHSY モデルの制約 ($\beta_1 + \beta_2 + \beta_3 = 1$) を満たさないように修正したモデルである。ここでは、特定化を誤った時の問題点を指摘することを目的としているので、1つのケースについてのみ実験結果を示すが、様々な DGP に対して実験を行い、その結果について回帰分析を用いて結論を一般化する (Response Surface 分析と呼ばれる) ことも可能である。

M_3 : 差分

	パラメータ	t 値	標準誤差	DW 比	RESET (2)
最小値	0.1010	0.9014	0.6604	1.3839	-4.5564
最大値	0.9838	9.4791	1.4298	2.9543	3.8473
中央値	0.5070	4.4986	1.0696	2.2159	-0.0057
平均	0.5073	4.5490	1.0706	2.2140	0.0013
標準偏差	0.1190	1.1890	0.0895	0.2070	1.0061
バイアス (β_1)	0.0073		0.0706		

 M_4 : 先行指標

	パラメータ	t 値	標準誤差	DW 比	RESET (2)
最小値	-0.6019	-4.2104	0.8853	0.2386	-5.8712
最大値	1.2719	11.6364	2.5641	1.8128	5.2040
中央値	0.3081	2.6087	1.4934	0.7325	0.0023
平均	0.3048	2.6463	1.5068	0.7508	0.0103
標準偏差	0.2093	1.8666	0.2054	0.1979	1.3431
バイアス (β_3)	0.7048		0.5068		
長期バイアス	-0.0952				

 M_5 : 部分調整

	長期乗数	標準誤差	DW 比	RESET (2)	VIF	Durbin's h
最小値	-0.3795	0.7517	1.1029	-4.1911	1.0000	-3.0245
最大値	2.5405	1.4533	2.3641	4.7668	2.4439	5.6789
中央値	0.6900	1.0549	1.7906	-0.0032	1.1051	1.1706
平均	0.7032	1.0560	1.7921	0.0016	1.1543	1.1911
標準偏差	0.2838	0.0876	0.1678	0.9219	0.1638	1.0771
バイアス		0.0560				
長期バイアス	0.3032					

 M_6 : COMFAC

	パラメータ (CO 法)	t 値	標準誤差	DW 比	パラメータ (BM 法)
最小値	0.0692	0.5229	0.7072	1.2144	0.0720
最大値	1.0536	10.3229	1.3288	2.6231	1.0526
中央値	0.4983	4.4602	0.9875	1.9375	0.4980
平均	0.4990	4.5027	0.9891	1.9364	0.4985
標準偏差	0.1172	1.1812	0.0815	0.1590	0.1163
バイアス (β_1)	-0.0010		-0.0109		-0.0015
長期バイアス	0.0990				0.0985

 M_7 : 有限分布ラグ

	長期乗数	標準誤差	DW 比	RESET (2)	VIF
最小値	-0.5739	0.7982	0.1824	-6.1670	1.0643
最大値	1.4610	2.4203	1.4326	5.5048	5.9509
中央値	0.4494	1.4063	0.5980	-0.0082	2.0601
平均	0.4496	1.4212	0.6164	0.0052	2.1473
標準偏差	0.2442	0.2105	0.1698	1.3376	0.5286
バイアス		0.4212			
長期バイアス	0.3496				

M_8 : Dead Start

	長期乗数	標準誤差	DW 比	RESET (2)	VIF	Durbin's h
最小値	-23.1436	0.7943	1.3297	-3.5559	1.0000	-3.3379
最大値	1.1351	1.5082	2.5232	3.1086	3.1181	4.2244
中央値	-0.0849	1.1115	1.9429	-0.0117	1.2289	0.2058
平均	-0.1667	1.1133	1.9427	-0.0059	1.2903	0.2044
標準偏差	0.5469	0.0917	0.1577	0.9056	0.2556	0.9853
バイアス		0.1133				
長期バイアス	-0.5667					

M_9 : DHSY

	調整係数	標準誤差	DW 比	RESET (2)	VIF
最小値	-0.6339	0.7336	1.3960	-3.8624	1.0000
最大値	-0.0387	1.3485	2.6867	3.7813	1.3592
中央値	-0.2187	1.0144	1.9972	0.0082	1.0342
平均	-0.2282	1.0155	1.9961	0.0023	1.0428
標準偏差	0.0748	0.0817	0.1733	1.0085	0.0367
バイアス (β_2-1)	0.0218	0.0144			

1. パラメータの列のバイアスは対応する DGP のパラメータの値と平均の差であり、長期バイアスは DGP の長期乗数 (0.4) と平均の差を示す。
2. 標準誤差の列のバイアスは DGP における回帰モデルの標準誤差 (1.0) と平均の差である。

実験の結果は、以下のよう⁽³⁾にまとめることができる。

1. **静学回帰モデル**: 伝統的計量経済学の方法において最も基本的なモデルである。まず、回帰係数について、平均でみると、(4)式のパラメータ ($\beta_1=0.5$) 式と比べて約7%の過小推定であり、長期乗数 (0.4) と比べると約14%の過大推定となっている。したがって、これは定常過程のケースであるが、長期乗数を静学回帰によって推定することはバイアスの点からみて支持できない。また、標準誤差は、平均でみて約40%の過大推定であり、モデルの説明力は低い。最後に、DW 比、RESET (2)テストの結果より、RESET (2)テストでは、特定化の誤りはそれほど多く検出されないが、DW 比で見ると、最大値で見ても約1.45しかなく、1階の正の自己相関が発生する比率が高いと考えられる。したがって、このモデルの特定化の誤りは誤差項の自己相関に反映されており、疑似自己相関の典型例である。
2. **時系列モデル**: 時系列モデルにおける1階の自己回帰過程 (AR (1)) モデルである。形式上、外生変数を含まないために長期乗数は計測できないので、(4)式のパラメータ ($\beta_2=0.75$) と比べると、平均でみて約6%の過小推定である。標準誤差は平均で約11%の過大推定であり、このモデルの説明力の低さが示唆される。なお、DW 比をみると、2の周囲に分布しており自己相関が発生する比率は低いと考えられるので、OLS 推定量でも一致性を持つ推定量

(3) ここではモンテカルロ実験の結果についてのみ言及する。各モデルの一般的な性質については、Hendry, Pagan and Sargan (1984), Hendry (1995) において議論されている。

であると考えられる。あわせて Durbin の h 統計量も計算したところ、その標準偏差でみてもほとんど自己相関は検出されない、という結論が得られた。また、回帰係数は最大でも、0.901程度なので非定常過程になる恐れはない。

3. **差分モデル**：差分変数からなるモデルであり、対数差分であれば成長率モデルと考えることができる。まず、回帰係数について、(4)式のパラメータ ($\beta_1=0.5$) と比べると、平均で約1%の過大推定であり、ほとんどバイアスがなく、標準誤差で見ても、平均で約7%しか過大推定しておらず、静学回帰モデル (M_1) と比べてもはるかにあてはまりが良い。また、DW 比、RESET (2)テストでみても、自己相関や特定化の誤りのような問題が生じる比率が低いことがわかる。したがって、DGP が差分モデルのパラメータ制約に近い値に設定してあるために、このモデルの問題点を隠蔽しがちであるという問題はあるにせよ、静学回帰モデルに比して差分モデルのパフォーマンスの良さは明らかであり、この点からも差分モデルを基礎に、長期的関係も反映するようにエラー修正項を説明変数として加えて定式化される ECM を支持することができる。
4. **先行指標モデル**：説明変数に当期の外生変数 (x_t) を用いる静学回帰モデル (M_1) と異なり、1期前のもの (x_{t-1}) のみを用いるモデルである。回帰係数は、平均でみると(4)式のパラメータ ($\beta_3=-0.4$) とは符号が逆であり、長期乗数と比べても平均で約25%の過小推定である。標準誤差でみても平均で約50%の過大推定であり、静学回帰 (M_1) よりも当てはまりが悪い。また、DW 比でみても自己相関が発生する比率が高く、バイアス、標準誤差、自己相関、全ての点で他のモデルと比べて最悪の結果である。
5. **部分調整モデル**：均衡への調整費用を仮定した2次形式の損失関数から最適調整ルールとして導出されるモデルであり、調整方程式として最も広く用いられてきた定式化である。このモデルからえられる長期乗数 ($\frac{\beta_1}{1-\beta_2}$) は平均でみると約75%の過大推定である。したがって部分調整モデルから長期乗数を計測することは全くナンセンスな結果を与えることがわかる。しかし、標準誤差でみても、平均でわずか5%の過大推定であり、DW 比、RESET (2)でみても、モデルの特定化の誤りはほとんど検出されない。また、多重共線性の尺度として VIF をみると、最大でも2.44であり、多重共線性の問題は発生しないと考えられる。したがって、この場合は特定化の誤りがあり、長期乗数において大きなバイアスが発生しているにも関わらず十分な特定化のテストを行わないために適切なモデルとしてこのモデルが支持されてしまう可能性は高いと考えられる。しかし、説明変数に被説明変数のラグ変数が入っていることを考慮して、誤差項の自己相関を Durbin の h 統計量によって検定すると、Durbin-Watson 比と比べて十分に高い比率で自己相関が検出される。したがって、このケースでは、 h 統計量によって自己相関の有無を判断しなければ DW 比の持つ問題によって自己相関が隠蔽される可能性が高くなると考えられる。

6. **Common Factor モデル**：初期の LSE アプローチにおいて重点的に研究されたモデルであり、誤差項が自己相関するモデルである。 u_t を消去すると、

$$M_6^*: y_t = \beta_1 x_t + \beta_2 y_{t-1} - \beta_1 \beta_2 x_{t-1} + \varepsilon_t \quad (6)$$

となり、ADL(1,1)モデルと比べると $\beta_3 = -\beta_1 \beta_2$ という非線形制約が課されたモデルとなっている。ここでは、Cochrane and Orcutt (1949) (CO) 法と、Beach and MacKinnon (1978) の最尤法 (BM) による推定結果を示す。まず、回帰係数について、静学回帰 (M_1) モデルと比べて若干パラメータは大きくなり、CO 法、BM 法どちらでみても (4) 式のパラメータ ($\beta_1=0.5$) とのバイアスが改善されている。また、回帰係数でみる限り CO 法と BM 法の間に大差はない。また、標準誤差でみると、平均ではわずか 1% の過小推定であり、DW 比でみると、静学回帰モデルにおいては高い比率で検出された自己相関は、この場合ほとんど検出されない。すなわち、この結果は自己相関が特定化の誤りによってもたらされているにもかかわらず、誤差項本来の自己相関としてその影響を除去する推定法を採用することによりその誤りが隠蔽されてしまう典型例となっている。

7. **有限分布ラグモデル**：ADL(1,1)モデルと比べると被説明変数のラグ変数 (y_{t-1}) が除かれており、外生変数 (x_t) の動学的影響のみがモデル化された定式化である。まず、回帰係数について、長期乗数を比べると平均で約 9% の過大推定である。標準誤差も平均で約 40% の過大推定であるため、モデルのあてはまりは良くないことがわかる。また、DW 比でみると正の自己相関が高い比率で発生しており、特定化の誤りが反映された結果となっている。すなわち、モデルを正しく特定化するためには、説明変数のラグ構造を精緻化するだけでは不十分であり、被説明変数そのものの動学的特性を反映するための特定化、すなわち被説明変数のラグ変数を用いてデータの持つ短期的構造もモデル化すること（データ重視のアプローチ）の必要性を示唆している。

8. **Dead Start モデル**：当期の外生変数 (x_t) を使わずに、説明変数が先決変数 (y_{t-1} , x_{t-1}) のみからなるので、Dead Start と呼ばれる。長期乗数を比べると、符号が逆であり、バイアスが非常に大きい。標準誤差でみても平均で約 10% の過大推定であり、部分調整モデル (M_5) と比べても、さらに悪い結果であり、説明変数として x_t と x_{t-1} を使うことの差異は明確である。しかし、DW 比、RESET (2)、VIF でみてもモデルの問題が検出される比率は低い。自己相関については、Durbin の h 統計量でも同様の結論が得られる。したがって、このモデルにおいても部分調整モデルと同様に特定化の誤りがあるにもかかわらず適切なモデルとして受容されてしまい易いという問題が生じる。

9. **DHSY 型 ECM**：長期関係式の弾力性を 1（ただし、モデルに含まれる変数が全て対数の場合）と仮定して導出される ECM であり、長期乗数を計測する必要がなく、その簡便さから現在でも広く用いられるモデルである。ADL(1,1)と比較すると、 $\sum_i \beta_i = 1$ というパラメータ制

約が課されている。現在では、共和分ベクトルを推定し、このパラメータ制約が受容される場合にこのモデルを用いて推定するケースも多くある。実験の結果をみると、まず、調整係数 ($\beta_2 - 1 = -0.25$) のバイアスについては、平均で約 9 % の過大推定である。また、標準誤差でみると、平均でわずか 1 % の過大推定であり、COMFAC モデル (M_6) に次いであてはまりがよく、DW 比、RESET (2)、VIF をみても、自己相関や特定化の誤り、多重共線性の問題が検出される比率は低いと考えられる。特に多重共線性については、平均や最大値でみると他のモデルと比べて最も小さく、この例からも多重共線性を考慮した場合の ECM の優秀性が示唆される。

3. 多重共線性による比較

本節では、ADL モデルと ECM の動学的特定化の手法としての有効性を比較するとともに、ADL モデルや ECM と同等の一般性を持つ動学的線形回帰モデルである、Bårdsen (1989) 型の ECM と Bewley (1979) モデルについて、その定式化や統計的性質を整理し、特に多重共線性に注目して、モンテカルロ実験による比較を行う。

3.1 モデルの定式化

ADL モデルは、Hendry, Pagan and Sargan (1984) において動学的線形回帰モデルの最も一般的な形式として議論された。ADL モデルと ECM の関係は前節で述べたとおりであり、互いに線形変換によって導出可能な同型 (Isomorphic) なモデルである。ADL モデルの一般形は以下のよう

ADL

$$y_t = \alpha_0 + \sum_{i=1}^m \alpha_i y_{t-i} + \sum_{i=0}^n \beta_i x_{t-i} + \varepsilon_t \quad (7)$$

長期関係式は、 $y_t = y$, $x_t = x$ として、

$$y = k + \theta x, \text{ ただし, } \theta = \frac{\sum_{i=0}^n \beta_i}{1 - \sum_{i=1}^m \alpha_i}, \quad k = \frac{\alpha_0}{1 - \sum_{i=1}^m \alpha_i} \quad (8)$$

となり、 θ が長期乗数である。

前節で述べた ADL (1,1) モデルの場合と同様にして一般形の ECM が導出される。

ECM (a)

$$\Delta y_t = \alpha_0 + \sum_{i=1}^r \eta_i (y_{t-i} - x_{t-i}) + \beta_0 \Delta x_t + \sum_{i=1}^r \zeta_i x_{t-i} + \sum_{i=r+1}^n \beta_i x_{t-i} + \sum_{i=r+1}^m \alpha_i y_{t-i} + \varepsilon_t \quad (9)$$

ただし、

$$\eta_1 = \alpha_1 - 1, \eta_i = \alpha_i \ (i > 1), \ r = \min(m, n),$$

$$\zeta_1 = \alpha_1 + \beta_0 + \beta_1 - 1, \zeta_i = \alpha_i + \beta_i \ (i > 1)$$

一般形の ECM (a) では、DHSY 型のエラー修正項 $(y_{t-i} - x_{t-i})$ がラグを持つように定式化され、 η_i は各期のエラーに対する調整係数である。また、短期乗数は β_0 ある。なお、このモデルの第 5 項および第 6 項は ADL におけるラグの長さが等しくないとき ($m \neq n$) のみ現れる。説明変数としては、ラグを持つ DHSY 型のエラー修正項 $(y_{t-i} - x_{t-i})$ 、当期の外生変数の差分 (Δx_t)、各変数のレベルでのラグ変数 (y_{t-i}, x_{t-i}) が含まれる。このモデルの長所は各期のエラーに対しての調整係数 (η_i) が容易に計測、検定できる点にある。

なお、このモデルの長期乗数は以下のように求められる。

$$\theta = - \frac{\sum_{i=1}^r (\zeta_i - \eta_i) + \sum_{i=r+1}^n \beta_i}{\sum_{i=1}^r \eta_i + \sum_{i=r+1}^m \alpha_i} \quad (10)$$

さらに、このモデルのエラー修正項を、(2) 式のように $(y_{t-i} - \theta x_{t-i})$ という形式に変換することもできる。

ECM (b)

$$\Delta y_t = \alpha_0 + \sum_{i=1}^r \eta_i (y_{t-i} - \theta_i x_{t-i}) + \beta_0 \Delta x_t + \sum_{i=r+1}^n \beta_i x_{t-i} + \sum_{i=r+1}^m \alpha_i y_{t-i} + \varepsilon_t \quad (11)$$

ただし、

$$\theta_i = 1 - \frac{\zeta_i}{\eta_i}$$

である。ただし、先述したように、このモデルは直接推定できないので、外生的に θ_i を与える必要がある。本稿ではこのモデルは扱わない。

次に、Bårdsen 型の ECM について述べる。このモデルも ECM と同様に ADL からの線形変換として得られ、ECM (a) や ECM (b) と比べて、説明変数に差分変数が多く現れるように定式化されている。実証分析においてはこのモデル (特に Bårdsen (b) モデル) が最も頻繁に用いられ、このモデルが ECM と称して分析されていることも多い。定式化は以下のとおりである。

Bårdsen (a)

$$\Delta y_t = \alpha_0 + \sum_{i=1}^{m-1} \alpha_i^* \Delta y_{t-i} + \sum_{i=0}^{n-1} \beta_i^* \Delta x_{t-i} + \alpha_m^* y_{t-m} + \beta_n^* x_{t-n} + \varepsilon_t \quad (12)$$

ただし、

$$\alpha_i^* = \sum_{h=1}^i \alpha_h - 1 = \sum_{h=1}^i \eta_h, \quad \beta_i^* = \sum_{h=0}^i \beta_h$$

このモデルの特徴は、ある期 (この場合はラグ切断期) のレベル変数以外は全て差分になっている

点にある。なお、ECM (a)と比較すると、パラメータ α_i^* , β_i^* はそれぞれ調整係数 (η_i) と短期乗数 (β_i) の累積的な効果を示すパラメータとなっている。

また、このモデルの長期乗数は、

$$\theta = -\frac{\beta_n^*}{\alpha_m^*} \quad (13)$$

として求められる。この分子と分母はそれぞれ、 x_{t-n} , y_{t-m} の係数であり、ADL モデルや ECM の長期乗数 ((8), (10)) と比べて計算が容易である。

さらに、このモデルのレベル変数の部分を整理して、ECM (a)において定式化されたエラー修正項 ($y_{t-i} - x_{t-i}$) がある任意の期だけ現れるように修正することができる。例えば、1 期前のエラー修正項 ($y_{t-1} - x_{t-1}$) のみが含まれる形式で定式化すると以下ようになる。

Bårdsen (b)

$$\Delta y_t = \alpha_0 + \alpha_m^* (y_{t-1} - x_{t-1}) + (\alpha_m^* + \beta_n^*) x_{t-1} + \beta_0 \Delta x_t - \sum_{i=1}^{m-1} \alpha_i^* \Delta y_{t-i} - \sum_{i=1}^{n-1} \beta_i^* \Delta x_{t-i} + \varepsilon_t \quad (14)$$

ただし、

$$\alpha_i^* = \sum_{h=i+1}^m \alpha_h, \quad \beta_i^* = \sum_{h=i+1}^n \beta_h$$

この定式化は、実証分析において Engle and Granger (1987) の 2 段階法とともに最も多く用いられる定式化であり、これが代表的にパラメータ再編された ECM と呼ばれることが多い。

次に、Bewley モデルについて述べる。定式化は、以下のとおりである。

Bewley (a)

$$y_t = \lambda \alpha_0 - \lambda \sum_{i=1}^m \alpha_i (y_t - y_{t-i}) + \theta x_t - \lambda \left[\sum_{i=1}^n \beta_i (x_t - x_{t-i}) \right] + \lambda \varepsilon_t \quad (15)$$

ただし、

$$\lambda = \frac{1}{1 - \sum_{i=1}^m \alpha_i}, \quad \theta = \frac{\sum_{i=0}^n \beta_i}{1 - \sum_{i=1}^m \alpha_i}$$

Bewley モデルの特徴は、ECM, Bårdsen 型の ECM と異なり、被説明変数が差分変数 (Δy_t) ではなくレベル変数 (y_t) で定式化されていることと、 x_t のパラメータが長期乗数であるため、他のモデルのようにパラメータ推定値を用いて長期乗数を計算する必要がないことである。長期乗数が直接計測できることの利点は、同時にパラメータの分散が直接計測できることである (他のモデルを用いて推定した場合は、長期乗数がパラメータに関して非線形の関数になっているために、分散の計算が面倒である)。ただし、Bewley モデルは、説明変数の中に当期の被説明変数 (y_t) を含むので、OLS で推定すると、誤差項と説明変数の間に相関が生じ推定量が一致性を持たない。したがって、操作変数 (IV) 法で推定する必要がある。

Wickens and Breusch (1988) では, Bewley (a)モデルを変換して, 目的に応じた3つのタイプの Bewley モデルを提案した。

Bewley (b)

$$y_t = \lambda \alpha_0 - \lambda \sum_{i=0}^{m-1} \left(\sum_{j=i+1}^m \alpha_j \right) \Delta y_{t-j} + \theta x_t - \lambda \sum_{i=0}^{n-1} \left(\sum_{j=i+1}^n \beta_j \right) \Delta x_{t-j} + \lambda \varepsilon_t \quad (16)$$

このモデルでは, 説明変数が x_t 以外は全て1期前との差分変数になっており, 短期乗数の解釈も容易なモデルとなっている。

Bewley (c)

$$y_t = \lambda \alpha_0 + \sum_{i=1}^m \phi_i \Delta^i y_t + \theta x_t + \sum_{i=1}^n \gamma_i \Delta^i x_t + \lambda \varepsilon_t \quad (17)$$

ただし,

$$\phi = \lambda P_k \alpha, \quad \gamma = \lambda P_k \beta$$

$$\alpha' = (-1, \alpha_1, \dots, \alpha_m), \quad \beta' = (\beta_1, \dots, \beta_n), \quad \phi' = (-1, \phi_1, \dots, \phi_k), \quad \gamma' = (\theta, \gamma_1, \dots, \gamma_k),$$

$$P_k = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & -1 & -2 & -3 & \cdots & -k \\ . & 0 & 1 & 3 & & . \\ . & & 0 & -1 & & . \\ . & & & 0 & & . \\ 0 & & & & & (-1)^k \end{bmatrix}$$

このモデルの特徴は, 説明変数が x_t 以外は全て1階以上の高階の差分変数からなっている点であり, その利点は各変数が定常成長化にあるときの動学的長期関係式の計測が容易である点にある。すなわち, 定常成長化では, $\Delta y_t = g_y$, $\Delta x_t = g_x$ という一定値であるとする, $\Delta^i y_t = \Delta^i x_t = 0$, $i > 1$ であるから, 動学的長期関係式は,

$$y = (\phi_1 g_y + \gamma_1 g_x) + \theta x$$

となり, Bewley (c)モデルのパラメータから容易に計測できることがわかる。

Bewley (d)

$$y_t = \lambda \alpha_0 - \lambda \sum_{i=1}^m i \alpha_i \Delta y_t + \lambda \sum_{i=2}^m \alpha_i [i \Delta y_t - (y_t - y_{t-i})] + \theta x_t$$

$$- \lambda \sum_{i=1}^n i \beta_i \Delta x_t + \lambda \sum_{i=2}^n \beta_i [i \Delta x_t - (x_t - x_{t-i})] + \lambda \varepsilon_t \quad (18)$$

このモデルの特徴は平均ラグが容易に計算できる点にある。このモデルでは, Δy_t と Δx_t の係数をそれぞれ p, ν とすると, 平均ラグ μ は, $\mu = -(p + \frac{\nu}{\theta})$ として得られる。

3.2 モデルの統計的性質

ADL モデルと同型な諸モデルの統計的性質を比較した研究として、Wickens and Breusch (1988) と Banerjee *et al.* (1993) がある。ここでは、これらの研究を整理する。

まず第一に、同型な関係にあるならば、各モデルの OLS 推定量（ただし、Bewley モデルについては、同型の ADL モデルの説明変数を操作変数とする IV 推定量）が、関心あるパラメータについて同一のパラメータ推定量をもたらすことが証明されている。⁽⁴⁾したがって、どのモデルを用いて推定を行ったとしても、例えば長期乗数のような関心あるパラメータについて、モデルから得られた推定値をもとに計算すれば、等しい数値が得られる。

第二に、関心あるパラメータの分散の推定量について、一般的に、あるパラメータ ψ が、 $\psi = f(p_1, \dots, p_n)$ のようにパラメータ $(p_1, \dots, p_n)$ の関数として示されるとき、 ψ の推定量は、 $\hat{\psi} = f(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_n)$ で得られ、 $\hat{\psi}$ の分散の推定量は、以下のように近似される。

$$\widehat{\text{var}}(\hat{\psi}) \cong \sum_{i=1}^L \left(\frac{\partial \psi}{\partial p_i} \Big|_{p=\hat{p}} \right)^2 \widehat{\text{var}}(\hat{p}_i) + 2 \sum_{h < i} \left(\frac{\partial \psi}{\partial p_h} \Big|_{p=\hat{p}} \right) \left(\frac{\partial \psi}{\partial p_i} \Big|_{p=\hat{p}} \right) \widehat{\text{cov}}(\hat{p}_h, \hat{p}_i) \quad (19)$$

また、ベクトルであれば、 $\psi = f(p)$, $J_r = \frac{\partial \psi}{\partial p'}$ として、以下のように近似できる。

$$\widehat{\text{var}}(\hat{\psi}) = \widehat{J_r'} \widehat{\text{var}}(\hat{p}) \widehat{J_r} \quad (20)$$

ここで関心の対象が長期乗数の分散であるならば、分散の推定量の計算の容易さという点からみると、直接に分散が推定できる Bewley モデルが最も便利である。ADL, ECM, Bårdsen 等のモデルは、上述した近似計算をしなければならない。ADL モデルは (8) より、ECM は (10) より、Bårdsen モデルは (13) よりそれぞれ長期関係式が得られるが、ADL と ECM の長期関係式はパラメータが多く含まれるために計算が面倒であるのに対し、Bårdsen モデルの場合は、 α_m^* , β_n^* という 2 つのパラメータしか含まないために計算が容易であり、近似式は以下になる。

$$\widehat{\text{var}}(\hat{\theta}) \cong \left(\frac{1}{-\widehat{\alpha_m^*}} \right)^2 \widehat{\text{var}}(\hat{\beta_n^*}) + \left(\frac{\widehat{\beta_n^*}}{(\widehat{\alpha_m^*})^2} \right)^2 \widehat{\text{var}}(\widehat{\alpha_m^*}) + 2 \left(\frac{1}{-\widehat{\alpha_m^*}} \right) \left(\frac{\widehat{\beta_n^*}}{(\widehat{\alpha_m^*})^2} \right) \widehat{\text{cov}}(\hat{\beta_n^*}, \widehat{\alpha_m^*}) \quad (21)$$

3.3 モンテカルロ実験による比較

前項までの議論から、ADL, ECM, Bårdsen, Bewley 各モデルは、パラメータ再編によって相互に導出可能な同型の関係にあるならば、関心あるパラメータについて等しい推定値と漸近的に等しい分散の推定値を得られる。したがって、DGP（具体的には動学的線形回帰モデルのラグ構造）

(4) 証明は全て ADL(2,2) モデルについてなされている。

が既知の状況下では、これらのモデルを比較する必要は、計算の手間の差を除いては全く存在しない。

しかし、実証分析は未知のDGPを特定化する作業であり、動学的線形回帰モデルは、経済理論では完全に特定化することができない短期動学的なラグ構造を、データに基づいて特定化するためのツールである。したがって、これらのモデルの比較は、DGPが未知の状況下でのラグ構造特定化のツールとしての有効性、という視点からなされねばならない。

もし、ラグ構造の特定化に、実証分析において通常なされるように、「一般から特殊へ」の方法論的要請に従って、最長のラグを設定し、 t 値でみて有意でない変数を削除して、より単純なモデルに⁽⁵⁾特定化して再推定し、全ての変数が有意になるまで続ける、という方法を採用するならば、モデルに現れる説明変数が規定するパラメータの推定精度、すなわち多重共線性の高さは、これらのモデルの比較基準として決定的に重要である。なぜならば、高い多重共線性が発生している場合には、OLS推定量の分散が大きくなるために、ラグ特定化の判断材料である t 値が小さくなり、説明変数が必要以上に削除されて、DGPよりも単純なモデルが得られてしまう可能性が高くなるからである。

一般的に、回帰モデルの特定化において、真のモデル(DGP)よりも説明変数が少ないモデルでは、OLS推定量が不偏性を持たないのに対し、余分な説明変数を含むモデルでは、OLS推定量が不偏性を持ち、このことが「一般から特殊へ」の特定化が指示される理由の一つとなっている。また、前節のモンテカルロ実験においても、DGPよりも単純なモデルは長期乗数に関して大きなバイアスをもたらすことが明らかにされている。したがって、ラグ構造の特定化には、DGPよりも単純なモデルを選択することがないよう、出来る限り多重共線性の低いモデルを用いるべきである。

そこで、本項では、各モデルの多重共線性の程度をモンテカルロ実験によって比較する。実験のDGPは、定常過程を想定して、ADL(2,2)モデルを用いて以下のように設定する。

$$y_t = 0.25 + 0.75y_{t-1} + 0.15y_{t-2} + 0.5x_t - 0.4x_{t-1} + 0.15x_{t-2} + \varepsilon_t \quad \text{ただし, } \varepsilon_t \sim \text{IN}[0, 1] \quad (22)$$

$$x_t = 0.25 + 0.9x_{t-1} + v_t \quad \text{ただし, } v_t \sim \text{IN}[0, 1] \quad (23)$$

なお、推定するデータ数は、 $T=100, 300$ であり、10000回繰り返した結果である。 $T=300$ について分析するのは大標本下での性質を明らかにするためである。

分析対象としては、まずラグの特定化を誤った場合について、DGP(ADL(2,2))に比べてラグを短く誤った場合(ADL(1,1))とラグを長く誤った場合(ADL(3,3))の長期乗数とその分散の推定値を比較する。次に、ADL(2,2)、ECM(a)、Bårdsen(a)、Bårdsen(b)、Bewley(a)、ADL(1,1)、ADL(3,3)それぞれのモデルについて多重共線性の指標として分散拡大要因(VIF)の最

(5) この方法の他にも、赤池情報量基準(AIC)やSchwartzのBayes情報量基準(SBIC)のような情報量基準を用いる方法や、包括テストを併用する方法もある。

大値と状態数 (Condition Number) (蓑谷 (1992)) を計測し比較した。ただし, Bewley モデルは IV 推定を行うので, 説明変数の操作変数による推定値 (VIF 1, 状態数 1) と説明変数自体 (VIF 2, 状態数 2) とを用いて 2 通りの方法で多重共線性の指標を計測した。

結果は表 3 に要約される。

表 3 : 動学的線形回帰モデルのモンテカルロ実験の結果 ($T=100$)

$T=100$ のケース

	ADL (2,2)		ADL (1,1)		ADL (3,3)	
	長期乗数	長期分散	長期乗数	長期分散	長期乗数	長期分散
最小値	0.2645	0.0204	0.2901	0.0306	0.3670	0.0298
最大値	9.8601	117.9884	7.7097	33.3849	9.6816	240.8430
中央値	2.2550	0.2965	2.0050	0.3307	2.2624	0.2927
平均	2.3441	0.8178	2.0445	0.6246	2.3788	1.1453
標準偏差	0.8133	4.0073	0.7136	1.4261	0.8703	8.6953
長期バイアス	-0.1559		-0.4555		-0.1212	

	ADL (2,2)		ECM (a)	
	VIF	状態数	VIF	状態数
最小値	2.6014	3.6559	2.0473	3.1714
最大値	67.2693	23.2594	40.1138	15.9726
中央値	12.0794	9.5712	8.2159	6.7600
平均	14.6413	10.0622	9.3095	7.0139
標準偏差	8.7196	3.2433	4.5739	1.7485

	Bardsen (a)		Bardsen (b)	
	VIF	状態数	VIF	状態数
最小値	1.2133	1.6453	1.1604	1.5458
最大値	6.5328	4.9530	3.2611	3.4995
中央値	1.9090	2.5621	1.3996	1.9461
平均	2.0994	2.6388	1.4404	1.9889
標準偏差	0.7293	0.5148	0.1890	0.2285

	Bewley (a)			
	VIF 1	状態数 1	VIF 2	状態数 2
最小値	1.9788	2.7468	1.6517	2.4377
最大値	185.7024	32.3398	3.7616	5.1326
中央値	7.5489	6.9940	2.4942	3.7030
平均	11.7671	7.9315	2.5314	3.7205
標準偏差	15.3919	3.7303	0.3172	0.4144

	ADL (1,1)		ADL (3,3)	
	VIF	状態数	VIF	状態数
最小値	1.2855	1.6896	2.9714	4.4270
最大値	15.8324	8.5737	111.8946	32.1990
中央値	4.8401	4.3845	18.7115	12.6743
平均	5.3268	4.5233	22.9350	13.6185
標準偏差	2.2061	1.0852	14.9471	4.5832

$T=300$ のケース

	ADL (2,2)		ADL (1,1)		ADL (3,3)	
	長期乗数	長期分散	長期乗数	長期分散	長期乗数	長期分散
最小値	1.3778	0.0222	1.0747	0.0339	1.3548	0.0204
最大値	4.4036	0.9691	4.2615	1.1968	4.6106	1.7785
中央値	2.4192	0.1149	2.2010	0.1337	2.4313	0.1179
平均	2.4439	0.1388	2.2103	0.1561	2.4490	0.1458
標準偏差	0.3888	0.0994	0.3856	0.0981	0.3978	0.1197
長期バイアス	-0.0561		-0.2897		-0.0510	

	ADL(2,2)		ECM(a)	
	VIF	状態数	VIF	状態数
最小値	5.8608	6.3162	4.0275	4.6045
最大値	56.1828	21.3983	31.0024	14.1794
中央値	15.8740	11.1842	9.7317	7.2539
平均	16.9949	11.3731	10.2334	7.4174
標準偏差	6.6554	2.3662	3.2755	1.3355

	Bardsen(a)		Bardsen(b)	
	VIF	状態数	VIF	状態数
最小値	1.2815	1.8557	1.1760	1.5741
最大値	4.5612	4.1226	1.7669	2.2793
中央値	1.9893	2.5684	1.3216	1.8149
平均	2.0541	2.5955	1.3316	1.8259
標準偏差	0.4200	0.3117	0.0733	0.1065

	Bewley(a)			
	VIF 1	状態数 1	VIF 2	状態数 2
最小値	2.2721	3.2614	2.0191	3.0407
最大値	29.5181	13.6371	2.9684	4.6546
中央値	6.1499	6.2629	2.3634	3.6209
平均	6.8850	6.4857	2.3796	3.6368
標準偏差	3.0854	1.4112	0.1637	0.2310

	ADL(1,1)		ADL(3,3)	
	VIF	状態数	VIF	状態数
最小値	2.3142	2.7835	8.5239	8.2404
最大値	11.5706	7.3080	94.8927	29.3202
中央値	5.6134	4.8302	25.2205	15.0154
平均	5.7847	4.8681	27.2654	15.3156
標準偏差	1.4205	0.7070	11.1580	3.2849

1. 長期分散は、長期乗数の OLS 推定量の分散の推定値を示す。
2. VIF は、それぞれの回帰モデルにおける VIF の最大値を示す。
3. Bewley(a)の結果について、VIF 1 及び状態数 1 は説明変数の操作変数への回帰における説明変数の推定値を、VIF 2 及び状態数 2 は説明変数自体を用いた結果である。

表より以下のことがわかる。

1. ラグ選択を誤ったケースについて、長期乗数の真の値 (2.5) と比べると、 $T=100$ の場合には、平均でみるとどのモデルも過小推定であり、特に推定バイアスが大きいのは ADL(1,1) モデルで約18%の過小推定である。また、推定バイアスが最も小さいのは ADL(3,3) (約5%の過小推定) であり、正しく特定化されたモデルである ADL(2,2) (約6%の過小推定) よりも若干バイアスが小さい。したがって、先述したとおり、バイアスからみるとラグ特定化においては真のモデルよりも短く特定化をすることの方が長く特定化することよりも深刻なバイアスをもたらし、真のモデルよりも長いラグを用いても必ずしもバイアスが増大するわけではないことがわかる。
2. 長期乗数の分散を比較すると、どのモデルにおいても平均が中央値よりも明らかに大きく、右に歪んだ分布であることがわかる。中央値で比較すると、どのモデルもほぼ同じ大きさであるが、ADL(1,1)は若干大きな分散をもたらす。
3. $T=300$ の場合について、長期乗数を平均で比較すると、どのモデルにおいてもバイアスは改善されているが、ADL(1,1)モデルは、なお12%の過小推定であるのに対し、ADL(2,2) および ADL(3,3) モデルは約2%の過小推定であり、バイアスは非常に小さくなっている。したがって、ラグ特定化を短く誤った場合のバイアスは大標本下においても十分に改善されないことがわかる。
4. $T=300$ の場合について、長期乗数の分散を比較すると、 $T=100$ の場合と比べて平均と中央

値の差が小さく分布の歪みが小さくなっていると考えられる。また、どのモデルにおいても $T=100$ の場合と比べて分散が小さくなっていることがわかる。

5. 多重共線性について、ADL モデルの中では、説明変数の少ない順にパフォーマンスが良く、ADL(3,3)モデルにおいては、VIF の平均が20を越えており、深刻な多重共線性が発生していることがわかる。また、状態数からも同じ結論がえられる。すなわち、推定バイアスの場合とは逆にラグが短く説明変数の数が少ないほど多重共線性の影響は小さく、ADL モデルにおいては推定バイアスと多重共線性はトレードオフの関係にあると考えられる。
6. 多重共線性について、平均、中央値、最大値などの点からみても最もパフォーマンスが良いのは Bardsen (b)モデルである。このモデルは ADL(1,1)モデルと比べても明らかに多重共線性の影響は小さく、バイアスの点では前項の議論より ADL(2,2)と等しい結果をもたらすことを考えれば、このモデルを用いることによって ADL モデルの特定化において生じた推定バイアスと多重共線性のトレードオフは解消できると考えられる。
7. ECM (a)や Bewley (a)等も同型の ADL(2,2)モデルと比べれば多重共線性の影響は小さいが、Bardsen (b)モデルと比べると、どちらも劣っている。
8. 多重共線性の指標については、 $T=300$ の場合においても、結論は同様であり、数値にも大差はない。したがって、大標本においても多重共線性の影響は解消されないことがわかる。

4. 結 論

動学的線形回帰モデルとしての ECM の意義について、本稿では1) ECM と同型の ADL モデルから特殊ケースとして導出される様々なモデルの性質（2節）と、2) ECM と同型なモデル間での比較（3節）、という2点について議論した。本稿の結論は以下のようにまとめられる。

第一に、1) について、本稿では ADL(1,1)モデルを例に ADL モデルと ECM が同型であることを示し、このモデルにパラメータ制約を課して、従来の実証分析において広く応用されてきた様々なモデルが導出されることを示すとともに、DGP が ADL(1,1)モデルである場合に、特定化を誤ってこれらの制約的モデルを用いることの問題を、モンテカルロ実験によって分析した。この結果として、特定化を誤った制約的なモデルを用いた場合には常に大きなバイアスが生じることが示された。しかし、この分析で用いた Durbin-Watson 比、h 統計量、RESET テストだけではモデルの特定化の誤りを発見できない場合もあり、従来の検定統計量の問題も示唆される。また、Common Factor モデルに対するモンテカルロ実験は、特定化の誤りによる疑似自己相関であるにもかかわらず、自己相関の影響を除去する推定法を用いた場合にその誤りが隠蔽されてしまう典型例となっている。これらの結果から、ADL、ECM をはじめとする一般的なラグ構造を持つモデルの優越性が示された。

第二に、2) について、3 節では、2 節で一般的な定式化であることが示された ADL, ECM をはじめとする一般モデルのクラスの中でのモデル間の比較を行った。具体的には、これらの一般モデルのパラメータ推定量が同等であることを示すとともに、モデル間の優劣を比較するために、特定化の道具として多重共線性の影響が小さいモデルが望ましい、という基準を導入し、ADL と同型なモデルである、ECM, Bardsen 型の ECM, Bewley モデルについてモンテカルロ実験を行い比較した。この結果として、最もパフォーマンスがよいのは Bardsen 型の ECM であり、これは実証分析において通常このモデルがパラメータ再編された ECM の代表的定式化として用いられることに支持を与える根拠の一つとなる。

第三に、モデルの特定化を単純なモデルから始めるか（特殊から一般へ）、一般的なモデルから始めるか（一般から特殊へ）、という問題について本稿の分析の帰結を述べておく。2 節の分析から、DGP と比べて特殊なモデルは回帰係数において大きなバイアスをもたらすにもかかわらず、特定化の誤りが検出されない状況が多く存在することが、モンテカルロ実験によって示された。また、3 節では、DGP の ADL(2,2) モデルに対し、ラグの特定化を長く誤った ADL(3,3) モデルの方が、短く誤った ADL(1,1) モデルよりも長期乗数に関して推定バイアスが小さいことが示された。さらに、Bardsen 型の ECM のようなモデルであれば、十分な長さのラグを持ちながらも多重共線性の影響の低い定式化が可能であると考えられる。したがって、特定化の検定に十分に高い検出力を期待できない状況においては、「一般から特殊へ」の特定化を採用すべきであると考えられる。

最後に、本稿の分析の限界として、ここでの結果は特定の DGP のもとでえられた結果に過ぎず、本稿の帰結を一般化するためには、より具体的な統計分析による理論的裏付けと、Response Surface 分析を用いた実験結果の一般化、という 2 つの方向での分析の拡張が必要であることを指摘しておきたい。

(経済学部研究助手)

参 考 文 献

- [1] 蓑谷千凰彦 (1992) 『計量経済学の新しい展開』, 多賀出版
- [2] 蓑谷千凰彦 (1996) 『計量経済学の理論と応用』, 日本評論社
- [3] Banerjee, A., Dolado, J. J., Galbraith, J. D., and Hendry, D. F. (1993) *Cointegration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press
- [4] Banerjee, A., Galbraith, J. D., and Dolado, J. J. (1990) Dynamic specification and linear transformations of the autoregressive-distributed lag model, *Oxford Bulletin of Economics and Statistics*, 52, 95-104
- [5] Bardsen, G. (1989) Estimation of long run coefficient in error correction models, *Oxford Bulletin of Economics and Statistics*, 51, 345-350
- [6] Beach, C. M., and MacKinnon, J. G. (1978) A maximum likelihood procedure for regression

with autocorrelated errors, *Econometrica*, 46, 51-58

- [7] Bewley, R. A. (1979) The direct estimation of the equilibrium response in a linear dynamic model, *Economics Letters*, 3, 357-361
- [8] Cochrane, D., and Orcutt, G. H. (1949) Application of least squares regression to relationships containing autocorrelated error terms, *Journal of the American Statistical Association*, 44, 32-61
- [9] Davidson, J. E. H., Hendry, D. F., Srba, F., and Yeo, S. (1978) Econometric modelling of the aggregate time series relationship between consumers' expenditure and income in the United Kingdom, *Economic Journal*, 88, 661-692
- [10] Durbin, J. (1970) Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables, *Econometrica*, 38, 410-421
- [11] Engle, R. F., and Granger, C. W. J. (1987) Co-integration and error correction : representation, estimation, and testing, *Econometrica*, 55, 251-276
- [12] Granger, C. W. J., and Newbold, P. (1974) Spurious regressions in econometrics, *Journal of Econometrics*, 2, 111-120
- [13] Hendry, D. F. (1995) *Dynamic Econometrics*, Oxford University Press
- [14] Hendry, D. F., Pagan, A. R., and Sargan, J. D. (1984) Dynamic specification, ch. 18 in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, 2, North-Holland, Amsterdam, 1023-1100
- [15] Kiviet, J. F., and Phillips, G. D. A. (1994) Bias assessment and reduction in linear error correction models, *Journal of Econometrics*, 63, 215-243
- [16] Nickell, S. J. (1985) Error correction, partial adjustment and all that: an expository note, *Oxford Bulletin of Economics and Statistics*, 47, 119-129
- [17] Ramsey, J. B. (1969) Tests for specification errors in classical linear least-squares regression analysis, *Journal of the Royal Statistical Society*, B, 31, 350-371
- [18] Sargan, J. D. (1964) Wages and prices in the United Kingdom: a study in econometric methodology, ch. 10 in Hendry, D. F. and Wallis, K. F. (ed.), *Econometrics and Quantitative Economics*, Basil Blackwell, 1984
- [19] Wickens, M. R., and Breusch, T. S. (1988) Dynamic specification, the long-run and the estimation of transformed regression models, *Economic Journal* (Conference), 98, 189-205