

多くの情報変数を用いた 予測方法の有用性について

きたむらとみゆき こいけりょうじ
北村富行 / 小池良司

要 旨

本稿では、さまざまな金融・経済指標を情報変数として用い、CPIおよび実質GDP変化率の予測系列を作成したうえで、個別の変数による予測と多変数の情報を集約した予測のパフォーマンスを比較した。CPI・実質GDPいずれの予測でも、単独で予測を常に改善させる変数は見出せなかったが、個別予測系列の情報を集約することにより予測パフォーマンスは安定化することが示された。この背景を探るため、単純平均による予測統合について、予測誤差の要因分解を行ったところ、主として予測誤差の攪乱的な動きが互いに打ち消し合うことによって誤差縮小が生じていることがわかった。もっとも、予測誤差の縮小効果は変数の追加に伴って急速に減衰する一方で、予測パフォーマンスの劣る系列が追加されるため、予測パフォーマンスの改善効果は高々2~4変数で出尽くしてしまうことも確認された。このため、最適な予測統合対象を得るためには、統合を構成する原予測系列のパフォーマンスと、予測誤差系列間の相関が小さくなるような変数の組み合わせの両方を考えていく必要がある。

キーワード：情報変数、多変数予測、サンプル期間外予測、予測統合

.....
本稿の作成に当たっては、大垣昌夫先生（オハイオ州立大学）、北村行伸先生（一橋大学）ならびに日本銀行企画室、調査統計局、金融研究所のスタッフから有益なコメントを頂いた。ただし、本稿の内容・意見は筆者たち個人に属し、日本銀行ないし金融研究所の公式見解を示すものではない。

北村富行 金融研究所研究第1課（E-mail: tomiyuki.kitamura@boj.or.jp）
小池良司 金融研究所研究第1課（E-mail: ryouji.koike@boj.or.jp）

1. はじめに

本稿では、近年、国内外で注目されている多くの経済変数を利用したインフレ率および実質GDP成長率の予測手法について、そのわが国における有用性を評価するとともに、それらによって予測パフォーマンスが安定的に高まるメカニズムとその限界を検討する。

金融政策の決定において、経済の現状を把握し、将来の見通しを設定することは重要である。そのためには、半年後や1年後、2年後のインフレ率や産出量成長率を予測することが必要不可欠であり、経済予測を巡ってこれまでもさまざまな研究が行われてきた¹。

この中で代表的な手法の1つに、経済理論等から得られた知見によって、予測対象変数の将来の値と関係を有すると推測される個別の変数を、予測における情報変数として用いるという手法がある²。例えば、名目価格の硬直性等を前提とすれば右上がりのフィリップス曲線が導出されるから、失業率等の実体経済変数は将来のインフレの情報を含んでいると考えられよう。あるいは、株価等の資産価格は、理論的には資産から得られる将来収益の割引現在価値に等しいことから、先行きの経済環境の情報を含んでいると考えられよう³。このため、貨幣残高、長短金利差、株価、商品価格指標、失業率などの個別の変数が予測に有用な情報を含んでいるかという視点から、多くの研究が行われている⁴。

しかし、Stock and Watson [2001] が包括的なサーベイによって明らかにしているとおり、すべての時期、国にわたって安定的に予測に有用であるような変数は、これまでのところみつかっていない。すなわち、理論的に想定される各変数間の関係が予測に十分役立つほど安定的であるケースは、ほとんどみられない⁵。

こうした結果を踏まえて、理論的に予測に有用な情報を持っていると考えられるある特定の情報変数のみを用いるのではなく、さまざまな情報変数が持っている情報を集約することでより精度の高い予測を行おうとする試みが、近年多くみ

1 本稿は、インフレ率および実質GDP成長率が将来何パーセントになるかという予測（ピンポイント予測）に焦点を当てている。一方、経済予測においては、物価変動や景気の「転換点」をいかに予測するかということも、重要なテーマの1つである。転換点予測に重点をおいた最近の研究としては、本多・松岡 [2001]、粕谷・真木 [2001] が挙げられる。

2 本稿では、加藤 [1990] に従い、情報変数を「最終目標に対する相関関係と先行性を有する金融経済指標」と考える。

3 翁・白塚 [2002] では、この点を、1980年代のバブル期の経験を踏まえながら、金融政策の運営に引きつけて議論している。

4 これまで予測力の有無が検討されてきた変数としては、ここに挙げたもの以外にも、CP金利と国債金利間のスプレッドや、為替レートなどがある。Stock and Watson [2001] は、これらを含むさまざまな変数についての海外の先行研究を包括的にサーベイしている。個別の変数に着目した、日本における最近の研究としては、長短金利差の景気に対する予測力を検証したHirata and Ueda [1998]、フィリップス曲線の予測力を検証した三尾 [2000] および福田・慶田 [2001] などが挙げられる。

5 こうした結果は、Stock and Watson [1996]、Cecchetti, Chu and Steindel [2000]、Stock and Watson [2001] などの分析によっても得られている。

られるようになっている。

予測に有用な情報を集約する方法としての多変数予測には、種々のものが考えられるが、近年特に注目を集めているものは、大きく2つのアプローチに分けられる。両者の違いは、情報の集約と予測系列作成のどちらを先に行うかという点にある。すなわち、1つは、個々の変数に含まれている共通情報を抽出した少数の指標をまず作成し、これを用いて予測を行うという、「指数アプローチ」である。もう1つは、個々の変数によってまず予測を行い、得られた多数の予測系列を何らかの方法で統合し、最終的な予測系列を作成するという、「予測統合アプローチ」である⁶。

まず、指数アプローチを用いた研究をみると、Stock and Watson [1998] が米国の170種類の系列に対するダイナミック・ファクター・モデルを推定している⁷。彼らは、得られたファクターのインフレ率および実質鉱工業生産に対する予測力が、ARモデルや失業率等の個別変数によるモデルの予測力に比べて高いことを確認している。同様の研究は、ユーロ地域についてMarcellino, Stock, and Watson [2000] およびForni *et al.* [2002]、英国についてArtis, Banerjee, and Marcellino [2002]、においてなされており、いずれも、ダイナミック・ファクター・モデルの予測パフォーマンスが高いことを報告している。

一方、予測統合アプローチをとっている研究としては、Stock and Watson [2001] が挙げられる⁸。Stock and Watson [2001] は、まず、先進7カ国を対象に、38種類の経済変数をそれぞれ個別に情報変数として用い、産出量成長率およびインフレ率の予測系列群を作成した。そのうえで、これら予測系列群の中央値、平均および刈込平均といった予測統合系列を作成し、そのパフォーマンスを検討した。その結果、彼らは、予測統合系列が個別変数を単独で用いた予測を大きく上回るパフォーマンスを示し、また、予測力が不安定な系列群についても、それらの刈込平均をとって統合すると、予測力が安定的に改善する場合があると報告している⁹。

6 このほかにも、マクロ計量モデルを用いたアプローチ、VAR (Vector Autoregression: 多変量自己回帰モデル) や状態空間モデル等の予測モデルに直接多くの説明変数を取り込むというアプローチも存在する。これらのアプローチをとっている最近の日本の研究としては、前者については伴・齊藤 [2001]、後者については北川・川崎 [2001] が挙げられる。また、より多くの情報を予測に取り込むという考え方は、経済予測の実務でよく用いられている段階的接近法にも通ずると思われる。なお、段階的接近法は、経済を生産・消費など各部門に分割したうえで、各部門の担当者が見通しを設定し、それらが全体として整合性を持つように調整を繰り返す、いわば統計学的手法の対極にある方法である。

7 ダイナミック・ファクター・モデルとは、複数の個別系列の背後に共通のファクター (因子) が存在し、それらファクターと各個別系列との間に、動学的な構造があると想定したモデルである。詳しくは、Stock and Watson [1998] を参照されたい。

8 予測理論の分野においては、こうしたアプローチの歴史は古く、少なくともBates and Granger [1969] にまで遡ることができる。

9 予測系列を統合するアプローチをとっている研究としては、このほかにも例えばMarcellino [2002] が挙げられる。また、わが国では、大山 [2001] が、積み上げ方式による予測と全産業活動指数など4系列を単独で情報変数として用いた予測の合計5つの実質GDP予測系列を算出したうえで、これらをもとに予測統合系列を作成し、それらが高い予測パフォーマンスを示したと報告している。

また、これら2つのアプローチを同時に試している研究として、米国におけるインフレ予測を検討したStock and Watson [1999] がある。彼らは、指数アプローチによる予測方法として主成分分析¹⁰を採用し、また予測統合アプローチによる予測方法として、単純平均、中央値に加え、リッジ回帰によって算出した可変ウエイトの加重平均を試している¹¹。彼らの結果によると、これら2つのアプローチによる予測は、個別変数による予測モデルのパフォーマンスを上回り、なかでも特に、主成分分析により抽出した第1主成分を使ったモデルのパフォーマンスが優れていた。

もっとも、こうした海外における先行研究においては、上述のとおり、情報変数のもつ情報を集約するアプローチの有用性が認められているものの、これらの予測方法がどのようなメカニズムによって予測パフォーマンスを改善させるのか、また、こうした予測パフォーマンスの改善メカニズムが、一体どの程度の数の変数を予測に取り入れるまで働くのかについては、十分解明されていないように思われる。実際、予測パフォーマンスは、予測に取り入れる変数の数が多ければ多いほど向上するわけではなく、変数の数が多すぎるとかえって悪化してしまうとの結果も得られている (Stock and Watson [2001])¹²。したがって、これらのアプローチによる予測パフォーマンスの改善メカニズムとその限界を明らかにし、どの程度の数の変数を予測に取り入れるのが望ましいかを検討しておく必要がある。

そこで、本稿では、わが国における多くの変数を利用した経済予測方法の有用性を、Stock and Watson [1999] の枠組みを踏襲して評価するとともに、予測パフォーマンスの改善が生じるメカニズムとその限界を探る。その分析結果からは、わが国においてもこうした予測方法が有用であることが確認される。また、こうした予測手法によるパフォーマンスの改善および安定化は、主として予測誤差の攪乱的な動きが互いに打ち消し合うことによって生じていることが明らかにされる。さらに、予測誤差の縮小効果は変数の追加に伴って急速に減衰する一方で、予測パフォーマンスの劣る系列が追加されるため、予測パフォーマンスの改善効果は高々2~4変数で出尽くしてしまうことも確認される。

次節以降の構成は以下のとおりである。2節では、Stock and Watson [2001] の方法に基づき、わが国のCPIおよび実質GDP変化率について、自己回帰 (AR: autoregressive) 予測と、さまざまな金融・経済変数を情報変数として用いた予測系列を作成し、それらのパフォーマンスを概観する。3節では多変数による予測系列を作成し、それらのパフォーマンスをみる。4節では、予測パフォーマンス改善の要因とその限界について検討する。5節では、本稿の議論を総括する。

10 Stock and Watson [1998] は、変数の数が十分多い場合、いくつかの技術的な仮定のもとでは、主成分分析によって抽出される主成分が、ダイナミック・ファクター・モデルにおけるファクターについての一一致推定量となることを示した。Stock and Watson [1999] において彼らが主成分分析を用いているのはこのような背景がある。なお、本稿3節の分析でも、こうした背景のもと、主成分分析を採用している。

11 なお、リッジ回帰を用いた予測手法については3節を参照されたい。

12 なお、Watson [2000] によれば、決定係数でみた情報変数の追加による限界的な予測力の向上は、情報変数の総数が増加するに従い急減する。

2. 2変数予測とその結果

本節では、多変数予測のパフォーマンス改善メカニズムとその限界について検討する準備作業として、Stock and Watson [2001] に従い、情報変数と被予測変数のラグ値を用いた2変数予測を作成し、各情報変数のパフォーマンスをみる。具体的には、半年（2四半期）、1年（4四半期）、2年（8四半期）先までのCPI変化率および実質GDP変化率の予測を行ったうえで、その結果を整理する。

(1) 予測モデル

このモデルは、予測対象変数 Y および情報変数 X の現在時点までの情報、つまり予測時点の値およびそのラグ値を用い、予測対象変数の h 期先までの変化率を予測する¹³。

$$y_{t+h}^h = \alpha + \beta(L)y_t + \gamma(L)X_t + \varepsilon_{t+h}^h. \quad (1)$$

ただし、ここで、 $y_t = \ln(Y_t) - \ln(Y_{t-1})$ は変数 Y_t の1期前の値との対数階差であり、前期比ベースのCPI変化率もしくは実質GDP変化率となる。また、 $y_{t+h}^h = \ln(Y_{t+h}) - \ln(Y_t)$ は変数 Y_{t+h} の h 期前の値との対数階差で t 時点からみた h 期先までの変化率、 X_t は情報変数の候補、 α は定数項、 $\beta(L)$ 、 $\gamma(L)$ はそれぞれ y_t 、 X_t のラグ多項式である。

上記の予測モデルのパフォーマンスは、(1)式で $\gamma(L)=0$ との制約条件を課したAR予測との対比で評価される。すなわち、ARモデルに情報変数を追加した2変数モデルによる、サンプル期間外予測の平均自乗予測誤差(MSFE: mean squared forecast error)を、対応するARモデルのMSFEで基準化した、相対MSFE(MSFE_{2変数予測} / MSFE_{AR予測})によって、予測パフォーマンスの比較を行う¹⁴。したがって、相対MSFEが1を下回れば、情報変数の追加によって、ARモデルの予測力が改善でき、パフォーマンスが向上したことを意味する。

13 (1)式の推計に当たっては、赤池情報基準(AIC: Akaike's information criterion)によって選択されたラグ次数を使い、過去40期間のサンプルによるローリング推計(rolling estimation)を行っている。なお、Stock and Watson [2001]では、サンプル期間の始期を固定し、それ以降のデータをすべて使用する再帰的推計(recursive estimation)を行っているが、サンプル期間が長くなるにつれ、経済構造や変数の情報量が変化することの影響を受けやすくなると考えられる。実際、以下でベンチマークとして用いるARモデルについて再帰的推計も試してみたところ、予測パフォーマンスはローリング推計の場合に比べて概ね低下した。なお、ラグ次数選択にベイズ情報基準(BIC: Bayesian information criterion)を用いた場合も、ARモデルの予測パフォーマンスはAICの場合に比べて概ね低下した。

14 すなわち、本稿では、Stock and Watson [2001]同様、ARモデルをベンチマークとして用いる。ただし、被説明変数の過去の情報のみに依存するARモデルの限界には留意しておく必要がある。

なお、サンプル期間内予測のパフォーマンスは、情報変数の追加によって確実に改善されるが、本稿で着目するサンプル期間外予測については、情報変数を追加することで予測モデルのパフォーマンスが必ず改善するとは限らない点に留意する必要がある¹⁵。

(2) 使用したデータ

情報変数の候補として、原系列の始期が1970~73年、終期は2001年前半まで、四半期ベースで利用可能な56変数を利用した(図表1)¹⁶。対象となった変数は、経済活動水準(鉱工業指数、失業率、短観D.I.など)、物価・賃金・市況関連(卸売物価、CRB指数など)、マネー関連(マネタリー・ベース、M2+CDなど)、資産価格関連(為替、金利、株価、地価など)の4種類に分類した。

上記変数の多くは、原系列に対して加工処理を行ったものを情報変数の候補としている。まず、季節性を有する系列については、季節調整済み系列が公表されている変数は原則としてそれを使用し、季節調整済み系列が公表されていない系列についてはX12-ARIMAを使って季節調整済み系列を作成した¹⁷。次に、月次もしくは日次データを四半期に変換して利用する場合、原則として、個別変数の性質にあわせて期末値もしくは期中平均値のいずれかを利用した(ただし、どちらか一方に特定できない一部の系列については、期末値と平均値の双方を採用)。また、各系列は、原計数だけでなく、対数値、対数階差値、HPフィルタ($\lambda=1,600$)¹⁸によるトレンドからの乖離といった加工系列もあわせて使用した。

以上のような複数の加工処理を施した結果、情報変数の候補として、CPI変化率については148系列、実質GDP変化率については147系列を使用している(図表2)。

15 予測モデルの推計に用いたサンプル期間を0期からT期とするとき、「サンプル期間内予測」とは、0期からT期までの予測値を指す。これに対し、本稿で検討の対象とする「サンプル期間外予測」は、T+1期以降の予測値である。サンプル期間外予測は、予測時点で利用可能な情報のみで算出された予測値であり、予測モデルの優劣の評価には、サンプル期間外予測を用いる方が適切であると考えられる。なお、本稿で用いているデータはすべて最終改定値であるが、こうしたデータは、一般に、予測時点では利用可能ではない。このため、本稿でのサンプル期間外予測は、厳密には「シミュレートされたサンプル期間外予測(simulated out of sample forecast)」と呼ばれる。

16 以下の分析では、改定済みのデータを使用した。なお、Bernanke and Boivin [2001]は、Stock and Watson [2001]の手法をリアル・タイム・データと最終改定値のみからなるデータ・セット双方に適用し、両者の予測パフォーマンスに大きな違いはないとの結果を示している。

17 消費税の直接的な影響を受けている統計(名目・実質GDP、GDPデフレーター、CPI、新車登録台数、大型小売店販売、百貨店販売)については、消費税の影響をX12-ARIMAの季節調整オプションで除去した。国内卸売物価については、消費税および夏季電力料金の影響を除いた。

18 HPフィルタは時系列データから「滑らか」なトレンド成分を抽出する変換法の一種であり、 λ は滑らかさを表すパラメータである。詳しくは肥後・中田 [1998]を参照されたい。

図表1 予測力テストに利用した変数

変数名	コード	季節調整	名目 / 実質		平均 / 期末		加工計算			
			名目	実質 r	平均 av	期末 ed	レベル lev	対数 ln	対数階差 ln1d	GAP gap
実質GDP	rgdp	SA								
鉱工業生産指数	ip	SA								
第3次産業活動指数	sanji	SA								
鉱工業稼働率指数	capu	SA								
短観業況判断DI	tnkzen	SA								
同（製造業）	tnksej	SA								
同（非製造業）	tnkhi	SA								
完全失業率	unemp	SA								
有効求人倍率	kyujin	SA								
機械受注額（民需）	kijmi	SA								
機械受注額（製造業）	kijse	SA								
機械受注額（非製造業）	kijhi	SA								
建設受注高	kenjal	SA								
建設受注高（民間）	kenjmi	SA								
建設受注高（非製造業）	kenjhi	SA								
公共工事請負額	ukeall	SA								
公共工事請負額（国）	ukekun	SA								
公共工事請負額（地方）	ukechi	SA								
住宅着工戸数	juckko	SA								
住宅着工床面積	juckme	SA								
建築着工床面積	ckhime	SA								
新車登録台数	car	SA								
大型小売店販売額	kouri	SA								
百貨店販売額	hyaka	SA								
通関輸出総額	expt	SA								
通関輸入総額	impt	SA								
名目GDP	ngdp	SA								
GDPデフレーター	pgdp	SA								
CPI	cpi	SA								
国内WPI・総平均	wpi	NSA								
国内WPI・中間財	wpiin	NSA								
輸入物価・総平均	ipiav	NSA								
輸入物価・素原材料	ipiso	NSA								
賃金指数	earn	SA								
原油（\$, ¥）	oil	NSA								
国内商品指数（¥）	commed	NSA								
ロイター指数（\$, ¥）	reu	NSA								
CRB指数（\$, ¥）	crb	NSA								
金（\$, ¥）	gld	NSA								
マネタリー・ベース	mon0	SA								
M1	mon1	SA								
M2+CD	mon2	SA								
最広義信用量	mon4	SA								
銀行貸出	lended	SA								
信用乗数	mlp	SA								
M2+CD流通速度	velo	SA								
銀行券発行残高	note	SA								
長期国債	jgb	NSA								
長短金利差	sprd	NSA								
貸出総合平均金利	alnd	NSA								
円/ドル・レート	rate	NSA								
実効為替レート	efrat	NSA								
株価・日経平均	nik	NSA								
株価・TOPIX	tpx	NSA								
東証出来高	tosho	NSA								
地価・6大都市商業地	land	NSA								

備考：1. 「SA」は季節調整あり、「NSA」は季節調整なし。

2. 加工計算の「レベル」、「対数」、「階差」、「GAP」はそれぞれ、原計数、対数値、対数階差値、HPフィルタによる加工値を意味する。

図表2 使用した各系列のコード、変数および変換方法

系列コード	変数名 (変換方法)	系列コード	変数名 (変換方法)
rgdp	In1d 実質GDP (対数階差)	ipisoav	In1d 輸入物価・素原材料・期中平均(対数階差)
rgdp	gap 実質GDP (gap)	ipisoed	In1d 輸入物価・素原材料・期末(対数階差)
ip	In1d 鉱工業生産指数(対数階差)	earn	In1d 賃金指数(対数階差)
ip	gap 鉱工業生産指数(gap)	rearn	In1d 賃金指数・実質(対数階差)
sanji	In1d 第3次産業活動指数(対数階差)	oil	In1d 原油・北海プレント・ドル建(対数階差)
sanji	gap 第3次産業活動指数(gap)	roil	In1d 原油・北海プレント・実質・ドル建(対数階差)
capu	In1d 鉱工業稼働率指数(対数階差)	lcoil	In1d 原油・北海プレント・円建(対数階差)
capu	gap 鉱工業稼働率指数(gap)	rlcoil	In1d 原油・北海プレント・実質・円建(対数階差)
tnksei	lev 短観DI・製造業(レベル)	commed	In1d 国内商品指数(対数階差)
tnkhi	lev 短観・非製造業(レベル)	rcommed	In1d 国内商品指数・実質(対数階差)
tnkzen	lev 短観業況判断DI・全体(レベル)	reued	In1d ROIター指数(対数階差)
unemp	lev 完全失業率(レベル)	rreued	In1d ROIター指数・実質(対数階差)
unemp	gap 完全失業率(gap)	lcreued	In1d ROIター指数・円建(対数階差)
kyujin	lev 有効求人倍率(レベル)	rlcreued	In1d ROIター指数・実質・円建(対数階差)
kyujin	gap 有効求人倍率(gap)	crbed	In1d CRB指数(対数階差)
kijmi	In1d 機械受注額 民需・除船舶、電力 (対数階差)	rcrbed	In1d CRB指数・実質(対数階差)
kijmi	gap 機械受注額 民需・除船舶、電力 (gap)	lccrbed	In1d CRB指数・円建(対数階差)
kijse	In1d 機械受注額 製造業 (対数階差)	rlccrbed	In1d CRB指数・実質・円建(対数階差)
kijse	gap 機械受注額 製造業 (gap)	glded	In1d 金価格(対数階差)
kijhi	In1d 機械受注額 非製造業・除船舶、電力 (対数階差)	rglded	In1d 金価格・実質(対数階差)
kijhi	gap 機械受注額 非製造業・除船舶、電力 (gap)	lclgded	In1d 金価格・円建(対数階差)
kenjal	In1d 建設受注高・総計(対数階差)	rlclgded	In1d 金価格・実質・円建(対数階差)
kenjal	gap 建設受注高・総計(gap)	mon0	In1d マネタリー・ベース(対数階差)
kenjmi	In1d 建設受注高・民間等計(対数階差)	mon0	gap マネタリー・ベース(gap)
kenjmi	gap 建設受注高・民間等計(gap)	rmon0	In1d マネタリー・ベース・実質(対数階差)
kenjhi	In1d 建設受注高・非製造業計(対数階差)	rmon0	gap マネタリー・ベース・実質(gap)
kenjhi	gap 建設受注高・非製造業計(gap)	mon1	In1d M1(対数階差)
ukeall	In1d 公共工事請負額・合計(対数階差)	mon1	gap M1(gap)
ukeall	gap 公共工事請負額・合計(gap)	rmon1	In1d M1・実質(対数階差)
ukekun	In1d 公共工事請負額・国(対数階差)	rmon1	gap M1・実質(gap)
ukekun	gap 公共工事請負額・国(gap)	mon2	In1d M2+CD(対数階差)
ukechi	In1d 公共工事請負額・地方(対数階差)	mon2	gap M2+CD(gap)
ukechi	gap 公共工事請負額・地方(gap)	rmon2	In1d M2+CD・実質(対数階差)
juckko	In1d 住宅着工戸数・総計(対数階差)	rmon2	gap M2+CD・実質(gap)
juckko	gap 住宅着工戸数・総計(gap)	mon4	In1d 最広義信用量(対数階差)
juckme	In1d 住宅着工床面積・総計(対数階差)	mon4	gap 最広義信用量(gap)
juckme	gap 住宅着工床面積・総計(gap)	rmon4	In1d 最広義信用量・実質(対数階差)
ckhime	In1d 建築着工床面積・民間非居住用(対数階差)	rmon4	gap 最広義信用量・実質(gap)
ckhime	gap 建築着工床面積・民間非居住用(gap)	lended	In1d 銀行貸出(対数階差)
car	In1d 新車登録台数・乗用車計(対数階差)	lended	gap 銀行貸出(gap)
car	gap 新車登録台数・乗用車計(gap)	rlended	In1d 銀行貸出・実質(対数階差)
kouri	In1d 大型小売店販売額計(対数階差)	rlended	gap 銀行貸出・実質(gap)
kouri	gap 大型小売店販売額計(gap)	mlp	lev 信用乗数(レベル)
hyaka	In1d 百貨店販売・全国(対数階差)	mlp	gap 信用乗数(gap)
hyaka	gap 百貨店販売・全国(gap)	velo	In1d 流通速度(対数階差)
lcexpt	In1d 通関輸出・円建(対数階差)	velo	gap 流通速度(gap)
lcexpt	gap 通関輸出・円建(gap)	note	In1d 銀行発行高(対数階差)
lcimpt	In1d 通関輸入・円建(対数階差)	note	gap 銀行発行高(gap)
lcimpt	gap 通関輸入・円建(gap)	rnote	In1d 銀行発行高・実質(対数階差)
expt	In1d 通関輸出・ドル建(対数階差)	rnote	gap 銀行発行高・実質(gap)
expt	gap 通関輸出・ドル建(gap)	igbed	lev 長期国債(レベル)
impt	In1d 通関輸入・ドル建(対数階差)	rigbed	lev 長期国債・実質(レベル)
impt	gap 通関輸入・ドル建(gap)	sprded	lev 長短金利差(レベル)
ngdp	In1d 名目GDP(対数階差)	alndav	lev 貸出総合平均金利(レベル)
pgdp	In1d GDPデフレータ(対数階差)	raindav	lev 新規貸出約定金利(レベル)
cpi	In1d CPI(対数階差)	rateav	In1d 円/ドル・レート・期中平均(対数階差)
wpi	In1d 国内WPI・総平均(対数階差)	rateav	gap 円/ドル・レート・期中平均(gap)
wpiinv	In1d WPI・中間財・期中平均(対数階差)	rateed	In1d 円/ドル・レート・期末(対数階差)
wpiined	In1d WPI・中間財・期末(対数階差)	rateed	gap 円/ドル・レート・期末(gap)
ipiv	In1d 輸入物価・総平均(対数階差)	ratest	lev 円/ドル・レート標準偏差(レベル)

図表2 続き

系列コード	変数名 (変換方法)	系列コード	変数名 (変換方法)		
ratesk	lev	円/ドル・レート歪度 (レベル)	niked	ln1d	日経平均・期末 (対数階差)
ratekr	lev	円/ドル・レート尖度 (レベル)	nikst	lev	日経平均・標準偏差 (レベル)
rrateav	ln1d	実質為替レート・期中平均 (対数階差)	niksk	lev	日経平均・歪度 (レベル)
rrateav	gap	実質為替レート・期中平均 (gap)	nikkr	lev	日経平均・尖度 (レベル)
rrateed	ln1d	実質為替レート・期末 (対数階差)	rnikav	ln1d	日経平均・実質・期中平均 (対数階差)
rrateed	gap	実質為替レート・期末 (gap)	rniked	ln1d	日経平均・実質・期末 (対数階差)
efratav	ln1d	実効為替レート・期中平均 (対数階差)	tpxav	ln1d	TOPIX・期中平均 (対数階差)
efratav	gap	実効為替レート・期中平均 (gap)	tpxed	ln1d	TOPIX・期末 (対数階差)
efrated	ln1d	実効為替レート・期末 (対数階差)	rtpxav	ln1d	TOPIX・実質・期中平均 (対数階差)
efrated	gap	実効為替レート・期末 (gap)	rtpxed	ln1d	TOPIX・実質・期末 (対数階差)
refratav	ln1d	実質実効為替・期末 (対数階差)	tosho	ln	東証出来高 (対数値)
refratav	gap	実質実効為替・期末 (gap)	tosho	ln1d	東証出来高 (対数階差)
refratav	ln1d	実質実効為替・期中平均 (対数階差)	land	ln1d	地価・6大都市商業地 (対数階差)
refratav	gap	実質実効為替・期中平均 (gap)	rland	ln1d	地価・6大都市商業地・実質 (対数階差)
nikav	ln1d	日経平均・期中平均 (対数階差)			

備考：データの始期は1970年。ただし、第3次産業活動指数は1973年、建設受注 (kenjal, kenjmi, kenjhi) は1971年、公共工事請負額 (ukeall, ukekun, ukechi) は1973年、長期国債 (rjgb) は1971年、円/ドル・レートとその関連指標 (rateav, rateed, ratest, ratesk, ratekr, rrateav, rrateed, efratav, efrated, refrated, refratav) は1973年、一部の日経平均 (nikst, niksk, nikkr) は1972年。

(3) 予測結果

前述の予測モデルに基づき、以下では、CPI変化率、実質GDP成長率について、ベンチマークとなるARモデルに対するパフォーマンスの変化をみる。なお、以下では、サンプル期間外予測とその実績からの評価が可能な1983年第1四半期から1999年第2四半期 (以下99/2Qと略記)までを、プリ・バブル期 (1983～86年) バブル生成期 (1987～90年) バブル崩壊期 (1991～94年) ポスト・バブル期 (1995～99/2Q) の4つの時期に分割して、各期間ごとの予測パフォーマンスの違いをチェックする。

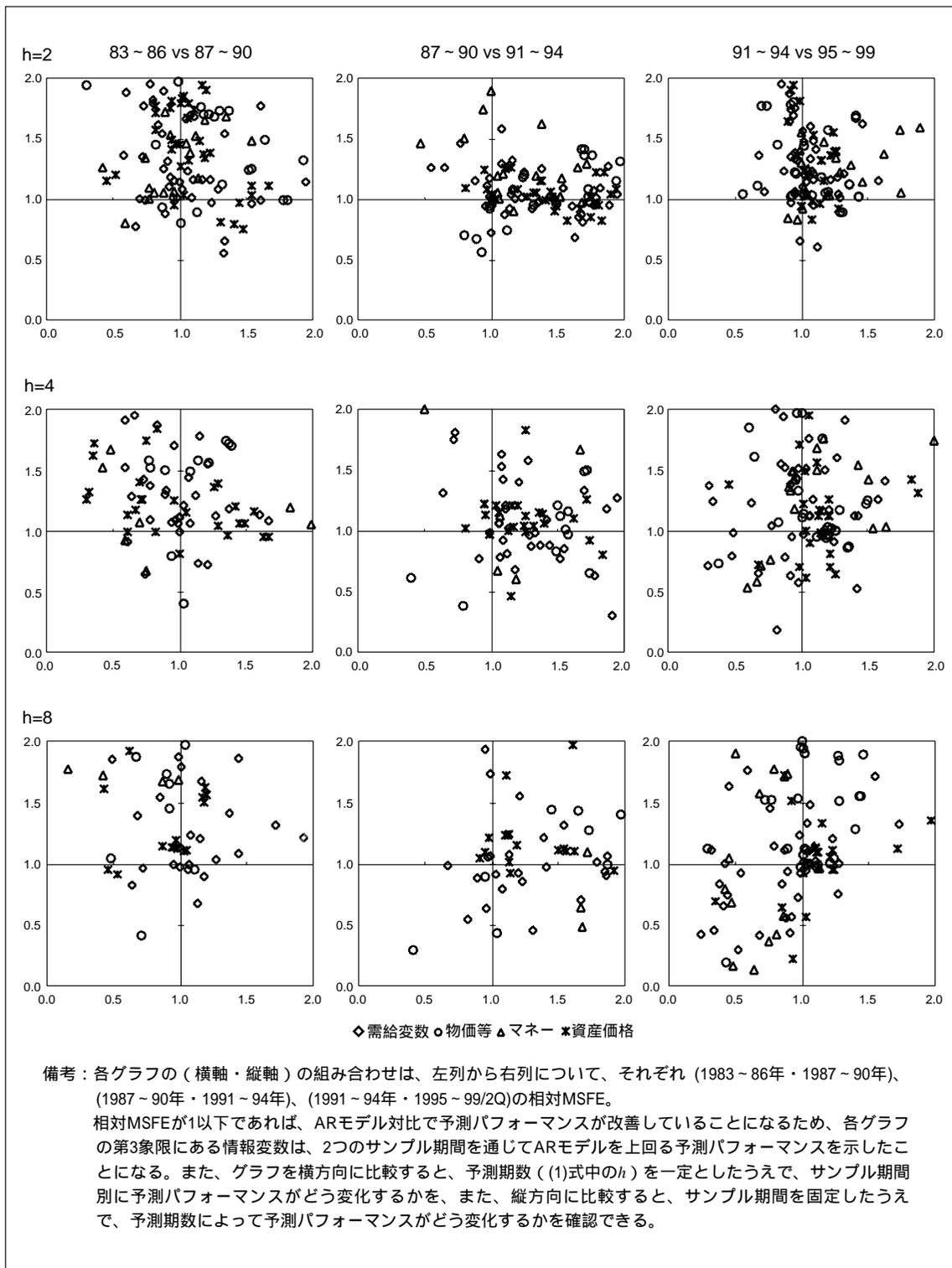
イ．CPI変化率予測

まず、CPI変化率の予測パフォーマンスを、サンプル期間別、予測期数別に比較するために、連続する2つのサンプル期間における2変数予測モデルの相対MSFEを予測期数別に散布図化したグラフ (図表3) をみると、最右列が最も第3象限の密度が高く、バブル崩壊期とポスト・バブル期に予測力が高まった系列が増加していることがわかる。また、情報変数の予測パフォーマンスを個別に確認すると (図表4) 物価・賃金関連指標の予測改善効果が相対的には高いことがわかる¹⁹⁾。しかし、すべての予測期数、サンプル期間を通じて、予測パフォーマンスを改善させるような変数は、残念ながら見当たらない²⁰⁾。

19 個々のサンプル期間をみると、物価・賃金等では、名目GDP (ngdp) や賃金 (earn) を中心に予測を改善している。また、マネー関連では、M2+CD (mon2)、実質銀行貸出 (rlended) がバブル崩壊以降、すべての先行期間で有効となっている。しかし、いずれもサンプル期間や予測期数が変わると、パフォーマンスが悪化している。

20 全予測期数、サンプル期間を通じて平均的に相対MSFEを最小化しているのは名目GDPであるが、それでも2期先予測の1995～99年、8期先予測の1987～90年においては相対MSFEが1を上回っており、ARモデルよりも予測パフォーマンスが悪化している。

図表3 CPI変化率予測のパフォーマンス



図表4 CPI変化率予測のパフォーマンス

Indicator	Trans.	h=2				h=4				h=8			
		83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99
AR	RMSFE	1.38	0.91	0.72	0.55	1.66	0.76	0.69	0.59	2.98	0.62	0.88	0.77
rgdp	ln1d	2.03	1.63	0.68	1.36	2.02	2.56	0.48	0.79	0.55	4.10	0.52	0.29
rgdp	gap	0.91	1.31	0.92	1.74	1.15	1.78	0.63	1.23	0.59	2.57	0.42	1.00
ip	ln1d	0.73	1.77	1.07	1.60	0.58	2.27	0.92	1.38	0.71	2.50	0.89	0.93
ip	gap	0.58	1.36	0.98	2.11	0.59	1.52	0.77	2.38	0.69	2.48	0.84	2.03
sanji	ln1d	1.49	2.51	0.92	0.97	1.12	12.71	0.93	0.95	2.58	97.55	0.99	0.92
sanji	gap	1.61	1.77	0.95	1.04	1.52	4.73	1.07	1.12	2.68	29.89	1.28	1.00
capu	ln1d	0.80	1.79	1.01	1.10	0.71	2.22	1.04	1.51	0.87	2.57	1.04	1.33
capu	gap	0.93	1.18	1.00	2.38	1.12	1.29	0.96	2.42	1.44	1.86	0.90	2.05
tnksei	lev	1.08	1.69	0.81	2.01	1.10	2.93	0.31	1.37	0.94	5.33	0.38	0.83
tnkhi	lev	0.70	1.00	0.72	1.06	0.59	1.91	0.30	0.71	0.82	20.39	0.24	0.42
tnkzen	lev	0.94	2.03	0.85	2.12	0.95	3.52	0.34	1.24	0.89	8.12	0.41	0.65
unemp	lev	4.16	1.38	1.25	3.59	4.49	2.14	1.41	4.70	1.12	3.01	2.40	1.62
unemp	gap	1.11	1.68	0.87	2.71	1.40	2.15	0.45	2.89	1.72	1.31	0.45	1.63
kyujin	lev	0.54	8.09	1.12	0.60	0.44	28.76	0.82	0.18	0.51	92.43	0.68	0.41
kyujin	gap	0.75	3.33	0.99	3.14	0.39	4.97	0.46	3.38	0.36	6.44	0.31	2.50
kijmi	ln1d	1.06	1.23	1.01	1.56	1.07	1.44	0.88	1.52	1.08	1.23	0.85	0.83
kijmi	gap	0.84	1.61	0.94	2.15	0.99	2.05	0.83	2.55	1.16	1.67	0.70	2.01
kijse	ln1d	0.72	1.35	0.97	1.32	0.89	1.30	0.87	1.94	0.95	2.38	0.87	1.11
kijse	gap	0.76	2.03	0.91	1.64	0.78	2.58	0.93	2.38	1.01	2.81	0.66	2.20
kijhi	ln1d	1.95	1.14	0.92	1.35	1.60	1.13	1.02	0.97	1.37	1.41	0.97	0.72
kijhi	gap	1.61	0.99	1.17	2.12	2.60	1.08	1.53	2.35	1.93	1.21	1.55	1.71
kenjal	ln1d	1.09	1.01	1.01	1.04	0.97	1.09	0.92	0.63	1.27	1.03	0.91	0.43
kenjal	gap	0.95	1.01	0.95	1.06	0.94	1.07	0.78	1.04	1.44	1.08	0.79	1.14
kenjmi	ln1d	0.78	1.95	1.04	1.14	0.90	1.33	0.98	0.57	1.18	0.89	0.88	0.55
kenjmi	gap	0.88	1.25	1.22	1.16	0.73	1.42	1.09	1.26	1.07	0.99	1.06	1.48
kenjhi	ln1d	0.87	1.54	1.09	1.16	0.78	1.37	0.88	0.78	1.15	1.20	0.92	0.56
kenjhi	gap	0.80	1.82	1.09	1.21	0.64	1.28	1.58	1.26	0.95	0.99	1.73	1.32
ukeall	ln1d	1.00	1.14	1.28	1.04	1.00	1.11	1.21	1.17	0.99	1.87	1.06	1.00
ukeall	gap	1.23	1.16	1.32	1.21	0.96	1.70	1.33	1.91	2.07	6.75	1.10	1.02
ukekun	ln1d	1.02	1.08	1.29	1.01	1.00	1.11	1.18	1.01	1.00	1.13	1.07	0.99
ukekun	gap	1.29	1.08	1.58	1.15	1.67	1.08	1.63	1.41	3.00	6.16	1.23	1.20
ukechi	ln1d	0.96	1.15	1.04	1.04	0.99	1.06	1.06	1.12	1.01	1.79	1.01	1.01
ukechi	gap	1.10	1.16	1.08	1.23	0.83	1.87	1.18	1.50	1.54	5.94	1.03	0.94
juckko	ln1d	1.24	0.97	1.11	1.41	1.08	1.06	1.06	1.76	0.68	1.39	1.21	2.32
juckko	gap	1.34	0.65	1.26	2.24	1.21	0.72	1.75	2.90	1.06	0.95	1.93	3.53
juckme	ln1d	1.54	0.96	0.94	1.69	1.27	1.12	0.81	2.00	0.49	1.85	0.93	2.60
juckme	gap	1.33	0.55	1.26	2.90	1.14	0.73	1.81	3.63	0.85	1.54	1.31	4.63
ckhime	ln1d	1.17	1.16	1.07	1.33	1.03	1.21	1.40	1.12	0.64	0.82	0.54	0.92
ckhime	gap	0.74	0.99	1.02	1.47	0.61	0.91	0.77	2.43	0.72	0.96	0.63	2.21
car	ln1d	0.98	1.01	0.99	0.65	1.37	1.18	0.68	0.65	0.80	8.89	0.44	0.74
car	gap	0.60	1.88	1.27	1.18	0.41	5.04	1.25	0.91	0.46	13.18	0.76	1.45
kouri	ln1d	0.92	1.10	0.87	1.23	0.78	1.09	1.42	0.52	0.47	4.05	0.34	0.45
kouri	gap	1.67	2.39	0.95	1.33	1.80	6.92	0.49	0.98	1.78	27.46	0.59	1.76
hyaka	ln1d	1.34	1.54	1.09	1.48	1.21	1.55	0.85	1.55	0.53	3.81	0.32	1.11
hyaka	gap	1.05	1.66	0.85	1.95	1.35	3.84	0.57	2.35	1.68	9.76	0.38	2.68
lceopt	ln1d	0.67	0.77	1.46	1.62	0.66	1.95	1.27	1.60	0.92	5.16	1.03	2.54
lceopt	gap	0.64	2.11	1.17	3.01	0.69	5.84	1.28	3.46	1.01	15.65	1.06	4.78
lcimpt	ln1d	0.95	2.47	1.09	0.95	0.88	2.49	1.43	1.12	1.06	2.58	1.28	2.12
lcimpt	gap	0.75	4.09	0.86	2.85	0.86	6.59	1.01	2.86	1.08	15.46	0.93	3.05
expt	ln1d	0.88	1.89	0.95	1.75	1.00	0.99	0.98	1.51	1.00	0.97	1.05	1.10
expt	gap	0.89	0.88	1.15	2.30	0.74	0.64	1.31	2.34	1.13	0.67	0.98	2.08
impt	ln1d	1.10	3.96	0.95	1.38	1.08	7.55	0.97	1.43	0.98	7.90	1.27	0.75
impt	gap	1.26	3.20	0.91	1.87	1.33	5.36	0.75	2.07	1.17	11.11	0.98	1.23
ngdp	ln1d	0.87	0.93	0.56	1.04	0.94	0.79	0.38	0.73	0.48	1.04	0.43	0.19
pgdp	ln1d	1.32	1.12	0.74	1.77	1.35	1.74	0.65	1.61	1.11	0.95	0.89	1.12
wpi	ln1d	0.44	2.28	0.82	1.45	0.42	2.58	1.16	1.76	1.03	4.20	0.97	1.53
wpiinav	ln1d	0.54	2.01	1.20	1.57	0.40	2.41	1.01	2.04	0.95	3.66	0.99	1.95
wpiined	ln1d	0.30	1.94	1.15	1.18	0.26	2.07	1.04	2.02	0.66	2.96	1.02	1.90
ipiaiv	ln1d	0.82	1.45	1.02	1.04	0.89	1.50	1.21	0.97	0.90	1.73	1.27	1.88
ipisoav	ln1d	1.30	1.73	1.02	1.03	1.40	2.07	1.20	1.03	0.97	4.85	1.28	1.51
ipisoed	ln1d	1.37	1.73	1.08	1.03	1.59	4.41	1.19	0.94	1.04	1.97	1.40	1.28
earn	ln1d	1.01	0.80	0.70	1.77	1.03	0.40	0.61	1.85	0.71	0.41	0.29	1.12
rearn	ln1d	1.13	0.89	0.67	1.11	1.08	1.49	0.83	1.07	0.67	1.87	0.99	0.97
oil	ln1d	1.32	2.64	1.43	1.02	0.90	4.31	1.22	1.01	0.86	3.02	1.04	1.00
roil	ln1d	1.39	2.71	1.43	1.02	1.00	4.50	1.23	1.02	0.83	3.17	1.04	1.00
lcoil	ln1d	1.16	1.76	1.36	1.12	1.64	2.12	1.26	1.00	0.92	2.20	1.01	1.07
rtcoil	ln1d	1.18	1.70	1.36	1.12	1.61	2.06	1.26	1.00	0.91	2.22	1.01	1.07
commed	ln1d	0.77	3.74	0.92	1.02	0.79	3.54	1.01	1.11	0.57	2.61	0.72	1.52
rcommed	ln1d	0.77	3.69	0.93	1.03	0.80	3.49	1.02	1.14	0.57	2.68	0.77	1.52
reued	ln1d	1.64	1.49	0.96	1.21	1.69	2.93	0.98	1.33	0.97	4.52	1.03	1.11
rreued	ln1d	1.93	1.32	0.95	1.22	1.97	2.60	0.96	1.42	1.06	3.89	1.02	1.12
lcreued	ln1d	0.99	1.97	1.31	0.89	1.24	2.57	1.36	0.87	0.92	1.45	1.44	1.55
rlcreued	ln1d	0.98	2.02	1.29	0.89	1.19	2.58	1.35	0.86	0.92	1.65	1.43	1.55
rcrbed	ln1d	1.87	2.19	1.02	1.44	2.35	2.72	1.20	0.96	1.26	2.57	1.21	1.00
rcrbed	ln1d	2.48	2.13	1.03	1.20	3.07	2.48	1.29	1.17	1.52	2.44	1.24	1.04
lccrbed	ln1d	1.22	1.70	1.41	1.67	1.39	1.70	1.49	1.22	1.06	2.53	1.46	1.89
rlccrbed	ln1d	1.26	1.68	1.41	1.69	1.37	1.72	1.50	1.25	1.00	2.75	1.28	1.84
glled	ln1d	1.78	0.99	0.94	1.80	1.14	1.58	0.97	1.97	0.74	3.26	1.01	1.94

図表4 続き

		h=2				h=4				h=8			
Indicator	Trans.	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99
AR	RMSFE	1.38	0.91	0.72	0.55	1.66	0.76	0.69	0.59	2.98	0.62	0.88	0.77
rglded	ln1d	1.81	0.99	0.92	1.78	1.22	1.56	1.01	1.97	0.81	3.14	1.00	2.00
lcglded	ln1d	1.52	1.24	1.18	1.05	0.77	1.58	1.16	0.96	0.89	2.31	1.11	0.97
rlcglded	ln1d	1.54	1.25	1.20	1.06	0.78	1.52	1.12	0.95	0.89	2.35	1.12	0.99
mon0	ln1d	0.94	3.23	1.00	1.30	1.11	7.47	0.91	1.36	0.63	14.04	0.47	0.68
mon0	gap	3.32	0.94	1.74	1.57	3.45	2.70	1.43	1.54	1.85	15.19	0.68	1.57
rmon0	ln1d	1.06	2.56	0.99	1.40	1.09	6.43	0.92	1.33	0.63	11.91	0.86	0.57
rmon0	gap	1.19	1.65	2.80	1.32	1.19	5.32	3.13	1.31	0.54	10.72	1.79	2.10
mon1	ln1d	0.81	1.05	1.19	1.03	0.84	3.13	1.64	1.03	0.39	8.95	1.13	0.96
mon1	gap	3.66	0.47	1.46	1.14	2.99	0.50	2.00	1.74	1.94	5.44	1.32	2.39
rmon1	ln1d	0.77	1.09	1.21	1.03	0.78	3.74	1.54	1.02	0.40	9.91	1.12	0.96
rmon1	gap	1.08	1.38	1.62	1.37	0.95	9.13	2.03	2.35	0.68	27.32	1.71	2.43
mon2	ln1d	1.12	1.52	1.02	1.09	1.09	3.58	0.77	0.76	0.22	9.55	0.81	0.42
mon2	gap	2.32	1.65	1.26	1.34	2.21	5.41	0.93	1.49	1.23	26.86	0.79	1.77
rmon2	ln1d	1.14	1.17	0.90	0.84	1.01	3.45	0.70	0.71	0.22	7.51	0.75	0.36
rmon2	gap	0.71	2.63	2.33	1.15	0.63	5.31	3.70	1.05	0.09	10.28	4.24	1.50
mon4	ln1d	0.59	0.80	1.50	57.19	0.59	0.92	3.60	54.57	0.16	1.77	2.77	33.32
mon4	gap	2.15	1.52	1.02	58.63	2.03	5.07	0.82	92.11	1.38	28.56	0.96	52.17
rmon4	ln1d	0.74	1.34	1.19	50.90	0.48	1.67	1.67	47.13	0.22	3.42	1.54	33.41
rmon4	gap	0.42	1.26	2.01	43.98	0.42	1.52	2.78	80.89	0.09	2.76	2.68	45.50
lended	ln1d	1.54	1.48	1.01	0.92	1.99	1.05	0.67	0.58	0.99	1.68	0.48	0.16
lended	gap	2.87	2.11	1.37	1.28	3.48	3.42	1.12	1.50	2.33	5.28	2.65	1.41
rlended	ln1d	1.35	1.68	0.97	0.83	1.83	1.19	0.60	0.53	0.87	1.67	0.64	0.13
rlended	gap	1.15	2.43	2.76	1.08	0.54	8.20	5.21	1.07	0.17	22.08	7.48	1.03
mlp	lev	1.19	4.65	1.75	1.05	1.95	15.43	1.51	1.42	1.14	44.30	2.83	0.84
mlp	gap	0.77	1.00	1.89	1.59	0.75	0.67	2.13	1.81	0.52	19.87	0.87	1.71
velo	ln1d	0.88	1.05	1.05	1.25	0.70	1.07	1.12	1.68	0.42	1.72	1.09	1.14
velo	gap	0.89	1.72	1.29	1.22	1.18	2.98	0.94	1.49	0.43	6.87	0.89	1.73
note	ln1d	0.93	1.53	1.17	1.47	1.00	3.96	0.95	1.42	0.57	8.93	0.45	1.04
note	gap	2.61	1.29	2.01	1.58	3.71	2.47	1.17	1.76	1.79	16.50	0.50	1.90
rnote	ln1d	0.96	1.06	1.00	1.55	0.89	3.75	0.95	1.18	0.53	6.80	0.42	0.79
rnote	gap	1.05	1.46	2.55	1.31	1.11	3.72	2.59	1.39	0.64	7.33	2.08	1.83
jgbed	lev	0.33	2.27	1.28	0.92	0.36	1.72	1.26	0.64	0.87	9.92	0.93	0.22
rjgbed	lev	0.52	1.20	1.08	0.83	0.32	1.32	1.04	0.61	0.81	2.64	0.85	0.64
sprded	lev	0.45	1.15	1.26	1.40	0.30	1.26	1.83	1.42	0.43	1.61	1.97	1.35
alndav	lev	1.00	2.27	1.14	0.96	1.06	4.06	0.68	0.72	1.30	10.81	0.87	1.72
raindav	lev	0.96	0.95	1.24	1.55	1.15	3.69	1.88	1.31	1.23	8.21	2.21	1.34
rateav	ln1d	1.00	1.45	1.06	2.21	0.61	1.13	1.00	2.02	1.22	2.39	1.13	2.11
rateav	gap	1.11	1.74	1.02	2.86	1.08	4.38	1.15	3.81	1.96	7.53	1.20	4.37
rateed	ln1d	1.17	1.94	1.08	2.24	0.35	1.62	1.10	2.43	0.6	3.04	1.11	2.57
rateed	gap	1.20	1.90	1.04	2.76	0.67	2.77	1.13	3.90	1.83	7.19	1.18	4.27
ratest	lev	1.31	0.81	1.09	1.53	1.00	0.81	1.02	1.22	0.87	1.14	0.92	1.51
ratesk	lev	1.03	1.03	1.00	0.94	0.60	2.93	1.07	0.90	1.33	2.75	1.01	0.98
ratekr	lev	2.37	1.01	1.04	1.20	1.36	0.96	1.13	1.12	2.10	0.98	1.21	1.05
rrateav	ln1d	1.19	1.34	1.05	2.26	0.82	0.99	0.97	2.16	1.00	2.73	1.06	2.05
rrateav	gap	1.01	1.79	0.96	2.93	0.81	3.48	1.03	3.86	1.42	6.54	1.05	4.41
rrateed	ln1d	0.97	1.46	1.02	2.41	0.61	0.99	0.97	2.60	0.86	2.57	1.06	2.36
rrateed	gap	0.94	1.81	0.95	3.11	0.64	3.01	1.04	4.01	1.48	5.75	1.00	4.41
efratav	ln1d	0.94	1.49	0.94	1.94	0.72	1.26	0.99	1.71	0.62	1.92	0.94	2.70
efratav	gap	0.82	1.57	0.82	2.34	0.88	2.76	0.53	4.12	1.20	6.62	0.58	5.30
efrated	ln1d	1.06	1.67	0.92	1.90	0.70	1.40	1.06	1.95	0.53	2.01	0.98	2.76
efrated	gap	0.99	2.07	0.79	2.53	0.80	2.43	0.72	3.12	1.11	5.72	0.60	4.72
refrated	ln1d	0.94	1.41	0.99	1.81	0.67	1.17	1.03	1.49	0.53	0.91	1.04	2.46
refrated	gap	0.93	1.75	0.85	2.17	0.83	1.84	0.80	3.18	1.41	3.71	0.70	4.09
refratav	ln1d	1.15	1.48	0.90	1.64	0.71	1.26	1.12	1.56	0.46	0.95	1.09	2.66
refratav	gap	1.02	1.83	0.82	2.23	0.75	1.74	0.92	2.95	1.07	3.67	0.70	4.41
nikav	ln1d	1.54	1.11	1.26	1.39	1.67	0.95	1.22	0.70	0.98	1.12	1.23	0.95
niked	ln1d	1.07	1.79	1.22	1.36	1.56	1.16	1.21	1.12	1.18	1.50	1.11	1.09
nikst	lev	1.01	1.27	1.02	1.23	1.14	13.98	0.99	0.70	1.12	28.38	1.03	0.56
niksk	lev	1.01	2.58	1.17	1.15	0.98	2.89	2.55	1.00	1.01	3.74	2.94	0.91
nikkr	lev	1.07	1.32	1.05	1.04	0.96	1.25	1.04	1.00	0.95	1.13	1.01	0.92
rnikav	ln1d	1.67	1.11	1.25	1.39	1.63	0.95	1.22	0.81	0.94	1.13	1.24	0.95
rniked	ln1d	1.03	1.85	1.22	1.36	1.42	1.20	1.21	1.12	1.17	1.54	1.12	1.09
tpxav	ln1d	1.18	1.37	1.15	1.32	1.49	1.06	1.15	0.98	1.03	1.10	1.23	1.11
tpxed	ln1d	0.82	1.71	0.97	1.36	1.26	1.36	1.15	1.16	1.20	1.56	1.10	1.12
rtpxav	ln1d	1.23	1.38	1.15	1.32	1.45	1.06	1.15	0.98	1.05	1.11	1.72	1.12
rtpxed	ln1d	0.82	1.77	0.97	1.36	1.29	1.39	1.14	1.17	1.19	1.62	1.10	1.13
toshoh	ln	1.54	1.03	1.00	0.94	1.05	1.15	0.46	1.38	0.84	3.21	0.35	0.69
toshoh	ln1d	1.45	0.97	1.08	0.98	1.29	1.04	1.21	1.26	0.97	1.19	1.15	1.33
land	ln1d	1.41	0.79	2.11	1.26	2.44	3.27	2.96	1.05	1.19	56.48	2.57	0.54
rland	ln1d	1.48	0.75	2.13	1.24	2.17	1.01	3.18	1.02	1.04	61.47	3.38	0.52

備考：1行目の計数は平方平均自乗予測誤差（RMSFE: root mean squared forecast error）。2行目以降は、相対MSFE。シャドローは相対MSFEが1未満の場合。なお、経済活動水準、資産価格で始期が1971~73年の系列については、1983~85年の予測値が欠損値となるため、1983~86年はこれらを含まない平均。

ロ．実質GDP成長率予測

実質GDPについても、CPIと同様に各系列による2変数予測について相対MSFEを算出し、それらを予測期数、サンプル期間別に図示した（図表5）。各列・各行をそれぞれ比べると、バブル生成期・崩壊期からポスト・バブル期の1年先・2年先予測が、最もパフォーマンスがよいように見受けられる。また、情報変数の予測パフォーマンスもあわせてみると（図表6）、マネー、資産価格において予測改善効果が高い²¹。しかし、サンプル期間の違いに対する頑健性という観点からみると、CPI予測と同様、すべての予測期数、サンプル期間を通じて、予測パフォーマンスを改善させるわけではない²²。

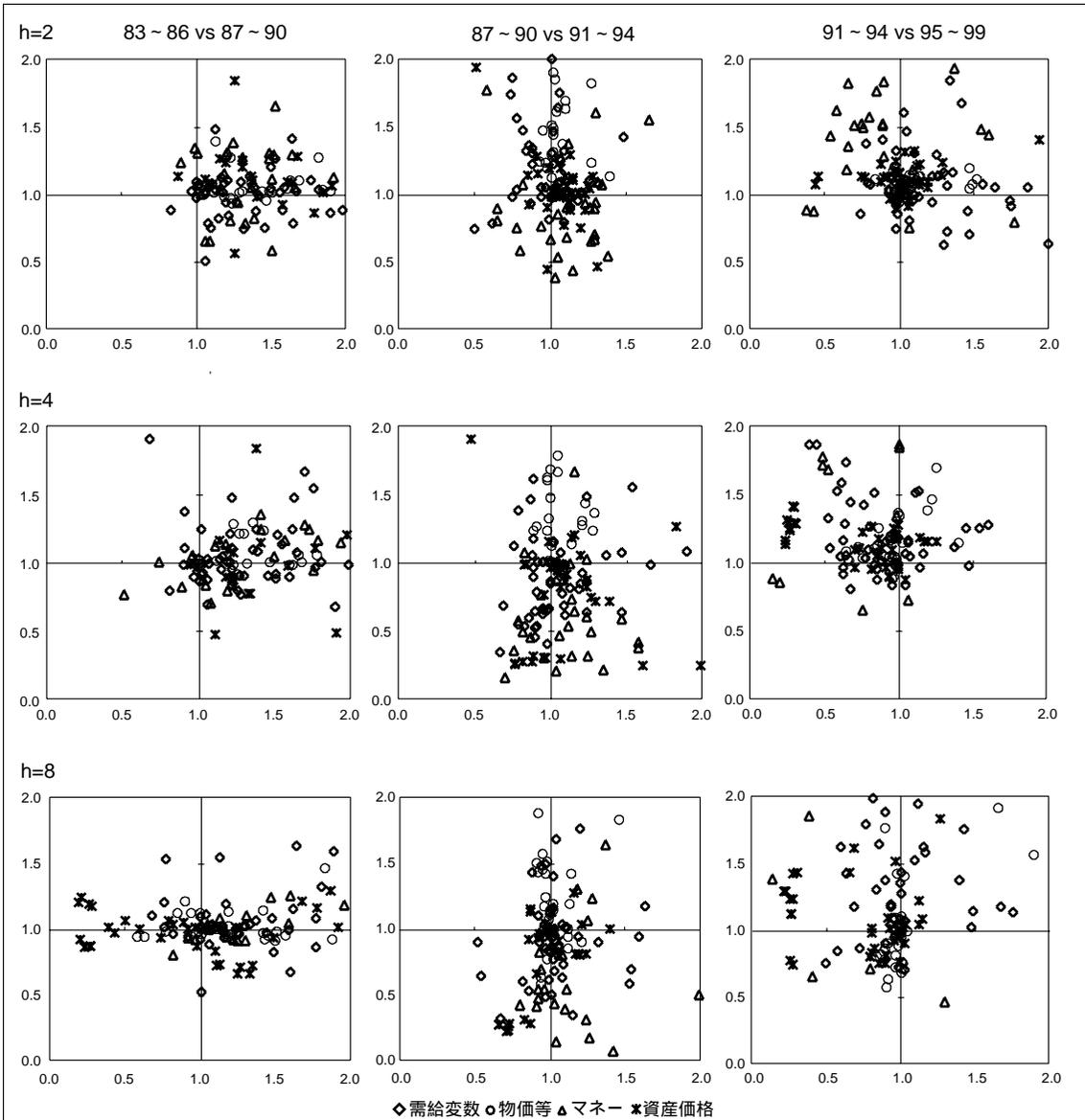
ハ．まとめ：2変数予測

以上の結果を整理すると、CPI予測、実質GDP予測いずれの場合においても、どのような2変数予測もAR予測のパフォーマンスを安定的に改善するとは言えない。言い換えれば、あるサンプル期間ではパフォーマンスを改善したとしても、別のサンプル期間ではその情報変数が予測を改善するかどうかはわからない。この結果は、Stock and Watson [2001] 等の結果を、わが国について追認するものと言えよう。

21 マネーについては、4、8期先予測では、マネタリー・ベース（mon0 ln1d）、M2+CD（mon2 ln1d）が他変数より長期間にわたり予測を改善している。また、資産価格については、4、8期先予測について、名目為替（rateav ln1d）、実効為替（efrateav gap）、日経平均（nikav lev）、TOPIX（tpxav ln1d）などで、2つの連続したサンプル期間で予測を改善している。

22 個別変数の中では、バブル生成期からポスト・バブル期まで継続して予測を改善しているのはM2+CD（mon2 ln1d）のみである。加えて、この変数も1983～86年の期間では一貫して相対MSFEが1を上回っているほか、4期先予測では1987～90年でも相対MSFEが1以上となっている。

図表5 実質GDP変化率予測のパフォーマンス



備考：各グラフの（横軸・縦軸）の組み合わせは、左列から右列について、それぞれ（1983～86年・1987～90年）、（1987～90年・1991～94年）、（1991～94年・1995～99/2Q）の相対MSFE。

相対MSFEが1以下であれば、ARモデル対比で予測パフォーマンスが改善していることになるため、各グラフの第3象限にある情報変数は、2つのサンプル期間を通じてARモデルを上回る予測パフォーマンスを示したことになる。また、グラフを横方向に比較すると、予測期数（(1)式中の h ）を一定としたうえで、サンプル期間別に予測パフォーマンスがどう変化するかを、また、縦方向に比較すると、サンプル期間を固定したうえで、予測期数によって予測パフォーマンスがどう変化するかを確認できる。

図表6 実質GDP変化率予測のパフォーマンス

		h=2				h=4				h=8			
Indicator	Trans.	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99
AR	RMSFE	1.06	2.39	1.81	2.65	0.94	2.04	2.34	1.90	0.86	1.72	2.61	1.66
ip	ln1d	0.97	1.02	1.22	0.94	1.02	0.91	0.78	1.03	1.14	0.99	0.84	1.30
ip	gap	1.00	1.09	1.32	1.06	1.55	1.13	0.81	1.26	1.26	0.96	0.94	2.70
sanji	ln1d	1.53	1.26	1.10	0.99	0.91	1.37	1.05	1.06	0.68	1.10	1.01	0.98
sanji	gap	1.52	2.53	1.13	0.98	0.68	1.90	1.08	1.04	0.43	6.53	0.96	1.00
capu	ln1d	1.00	0.97	1.05	0.99	1.23	0.89	0.89	1.01	1.05	0.99	0.96	1.08
capu	gap	1.09	1.06	1.26	1.13	1.65	1.07	1.07	1.16	1.24	1.00	1.12	1.94
tnksei	lev	1.22	0.84	1.32	0.72	1.52	1.20	0.84	0.98	1.81	1.32	0.90	0.90
tnkhi	lev	1.01	1.06	1.75	0.91	1.22	1.47	1.07	1.14	1.79	2.38	1.24	2.69
tnkzen	lev	1.15	0.82	1.47	0.70	1.38	1.24	0.86	0.87	1.89	1.59	0.94	1.17
unemp	lev	1.98	0.88	1.34	1.84	2.01	0.95	0.66	3.60	1.77	0.86	0.53	6.85
unemp	gap	1.02	1.07	1.25	1.29	1.07	0.99	0.66	2.93	1.29	1.02	0.77	1.79
kyujin	lev	0.83	0.88	1.30	0.62	0.96	2.59	0.63	1.16	0.77	5.74	0.25	2.31
kyujin	gap	1.08	0.78	1.03	1.60	1.22	0.83	0.53	2.85	1.16	0.97	0.48	3.71
kijmi	ln1d	1.10	0.75	0.98	0.74	1.77	0.96	0.65	1.05	1.37	1.06	0.86	1.64
kijmi	gap	1.90	0.86	1.36	1.16	1.99	0.98	0.65	1.73	1.44	0.97	0.85	3.60
kijse	ln1d	1.27	0.94	1.05	1.07	1.81	1.00	0.86	1.07	1.17	1.19	0.94	1.19
kijse	gap	1.58	0.88	1.22	1.13	1.60	0.89	1.17	1.06	1.06	0.88	1.43	1.75
kijhi	ln1d	1.06	0.50	0.74	0.85	1.30	0.90	0.64	1.28	1.49	0.82	0.60	1.62
kijhi	gap	2.15	0.62	0.78	1.37	1.90	0.67	0.34	2.20	1.60	0.67	0.32	5.05
kenjal	ln1d	1.06	1.02	0.99	0.85	1.16	1.04	0.86	0.95	1.00	1.00	0.97	0.96
kenjal	gap	1.06	1.00	0.98	0.95	1.01	0.86	0.59	1.52	1.33	1.03	0.68	2.23
kenjmi	ln1d	1.21	1.10	1.01	1.08	1.22	1.10	0.61	1.04	1.30	1.02	0.94	1.04
kenjmi	gap	1.13	1.03	1.05	1.46	1.27	0.90	0.45	1.86	1.48	1.08	0.63	2.85
kenjhi	ln1d	1.50	1.20	1.01	0.97	1.02	1.24	0.63	0.96	1.13	1.54	0.69	1.17
kenjhi	gap	1.36	1.13	0.89	1.40	0.90	0.98	0.40	1.86	1.62	1.15	0.34	2.75
ukeall	ln1d	1.50	1.04	1.04	1.03	1.15	0.95	1.00	1.18	1.07	0.99	1.01	1.43
ukeall	gap	1.46	0.75	1.86	1.05	0.97	0.89	1.61	1.27	0.80	1.04	1.68	1.17
ukekun	ln1d	1.11	1.07	1.09	0.95	1.03	0.92	1.00	0.95	1.04	0.99	1.00	1.04
ukekun	gap	1.32	0.74	1.74	0.95	0.81	0.79	1.38	1.11	0.92	0.94	1.48	1.02
ukechi	ln1d	2.03	1.01	0.97	1.02	1.19	1.00	1.00	0.99	1.08	1.03	1.00	1.35
ukechi	gap	1.65	0.78	1.56	1.07	1.06	0.87	1.46	1.25	0.76	1.20	1.76	1.13
juckko	ln1d	1.30	1.09	0.79	1.08	1.39	1.08	0.77	1.42	1.16	0.93	0.90	1.37
juckko	gap	1.42	0.99	0.81	1.07	1.51	0.91	0.53	1.32	2.22	0.54	0.64	1.42
juckme	ln1d	1.17	1.12	0.97	1.09	1.14	1.02	1.14	1.52	1.01	1.02	1.16	1.62
juckme	gap	1.40	0.87	0.93	1.03	1.06	0.69	0.68	1.44	1.01	0.52	0.90	1.88
ckhime	ln1d	1.18	1.01	1.30	1.16	1.22	0.95	0.62	1.58	1.09	1.00	0.82	1.98
ckhime	gap	1.06	1.01	0.98	1.32	1.46	0.90	0.52	3.19	1.59	0.99	0.61	4.40
car	ln1d	1.82	1.03	0.98	0.96	1.22	0.89	0.96	0.83	1.00	1.00	1.00	1.08
car	gap	1.27	0.94	0.98	0.94	1.26	0.79	0.54	1.10	1.04	1.11	1.01	1.27
kouri	ln1d	1.67	1.02	1.46	0.87	1.06	1.02	1.15	0.96	1.25	0.97	1.49	1.14
kouri	gap	1.64	1.41	2.12	1.08	1.76	1.54	1.55	1.25	1.96	2.72	1.74	2.03
hyaka	ln1d	1.39	1.05	1.64	1.05	1.21	1.00	0.84	1.51	1.15	1.02	1.40	1.37
hyaka	gap	1.13	1.48	1.42	1.67	1.70	1.66	0.98	2.18	3.12	2.65	1.20	4.91
lcexpt	ln1d	2.35	1.04	1.01	0.99	1.51	0.88	1.05	0.83	0.82	0.96	1.03	0.70
lcexpt	gap	2.05	1.18	0.95	1.11	1.36	1.09	0.68	0.80	0.76	1.01	0.50	0.75
lcimpt	ln1d	1.77	1.10	1.10	1.30	3.19	1.19	0.80	1.09	1.00	1.09	0.73	0.86
lcimpt	gap	2.07	1.32	1.07	0.80	1.63	1.47	0.63	0.91	0.77	1.53	0.58	0.84
expt	ln1d	1.20	0.94	1.05	1.07	0.91	1.10	0.98	1.21	1.07	1.02	0.99	0.72
expt	gap	1.59	1.05	1.08	1.08	1.28	0.76	1.12	1.51	1.00	0.92	1.10	1.52
impt	ln1d	1.19	1.08	2.05	0.88	1.21	1.21	0.92	1.15	1.77	1.08	1.03	1.07
impt	gap	1.56	1.01	2.00	0.63	1.62	1.24	1.48	0.97	1.64	1.63	1.17	1.58
ngdp	ln1d	1.13	1.39	1.52	1.11	1.36	1.29	1.20	1.38	2.17	1.35	1.60	2.21
pgdp	ln1d	1.23	1.27	1.47	1.04	1.43	1.23	1.41	1.14	1.83	1.46	1.66	1.91
cpi	ln1d	1.23	0.97	1.47	1.19	1.26	0.91	1.26	1.69	2.03	1.34	1.44	3.01
wpi	ln1d	1.51	1.01	0.84	1.11	1.23	0.89	0.65	1.08	1.88	0.92	0.92	0.90
wpiinav	ln1d	1.82	1.27	0.97	1.20	1.62	0.98	0.66	1.02	1.57	0.95	0.90	0.78
wpiined	ln1d	2.05	1.20	0.96	1.17	1.68	1.00	0.67	1.01	1.45	0.95	0.89	0.84
ipiav	ln1d	3.18	1.24	1.05	1.10	4.02	1.06	0.73	1.11	1.01	0.96	0.97	0.72
ipisoav	ln1d	4.48	1.23	1.03	1.11	5.36	1.22	0.76	1.03	1.43	0.92	0.91	0.57
ipisoed	ln1d	4.23	1.27	1.01	1.11	5.16	1.20	0.86	1.12	1.50	0.91	0.92	0.63
earn	ln1d	1.37	1.08	1.49	1.07	1.12	1.01	1.23	1.46	1.42	1.14	1.90	1.56
rearn	ln1d	1.12	1.01	0.98	0.92	1.01	0.99	1.03	1.14	0.99	0.98	0.90	1.76
oil	ln1d	5.37	1.16	1.02	1.04	2.95	1.26	0.98	1.04	0.77	0.98	1.00	0.99
roil	ln1d	5.03	1.16	1.02	1.04	2.87	1.29	0.91	1.04	1.07	0.99	1.05	0.99
lcoil	ln1d	5.19	1.26	1.00	1.02	4.61	1.17	0.99	1.02	0.63	0.94	1.02	0.75
rcoil	ln1d	4.91	1.27	1.00	1.02	4.48	1.20	0.99	1.02	0.58	0.94	1.01	0.76
commed	ln1d	1.44	1.02	1.00	1.11	1.27	1.21	0.87	1.25	1.19	1.02	0.97	0.81
rcommed	ln1d	1.47	0.95	1.08	1.05	1.23	1.28	0.84	1.10	0.94	1.12	0.99	0.88
reued	ln1d	1.85	1.03	1.00	0.94	1.32	0.99	0.98	0.98	1.24	0.97	0.95	1.09
rreued	ln1d	1.90	1.02	1.00	0.94	1.23	0.98	0.99	0.98	1.17	0.96	0.97	1.09
lcreued	ln1d	1.61	1.04	1.07	1.07	1.13	1.14	0.95	1.16	1.19	1.13	1.01	0.95
lcreued	ln1d	1.49	1.02	1.08	1.06	1.30	1.21	0.95	1.16	0.90	1.21	1.01	1.10
crbed	ln1d	1.19	0.98	1.07	1.14	1.66	1.05	0.96	1.03	1.51	0.98	0.98	1.42
rcrbed	ln1d	1.24	0.93	1.10	1.14	1.78	1.05	0.94	1.04	1.42	0.97	1.03	1.40
lccrbed	ln1d	2.81	0.98	1.11	1.10	3.54	1.15	0.95	1.18	1.00	1.12	1.01	0.71
lccrbed	ln1d	2.81	0.99	1.12	1.10	3.49	1.16	0.95	1.19	0.85	1.12	1.01	0.68

図表6 続き

		h=2				h=4				h=8			
Indicator	Trans.	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99
AR	RMSFE		2.39	1.81	2.65	0.94	2.04	2.34	1.90	0.86	1.72	2.61	1.66
glded	ln1d	1.29	1.01	1.04	1.09	1.00	1.01	0.96	0.88	1.15	1.01	0.87	0.82
rglded	ln1d	1.31	1.02	1.05	1.09	1.00	1.02	0.97	0.89	1.14	1.02	0.90	0.86
lcglded	ln1d	1.69	1.10	1.13	1.13	1.47	1.00	1.01	1.34	1.11	1.00	1.00	1.01
rlcglded	ln1d	1.63	1.10	1.14	1.13	1.60	0.98	1.00	1.36	1.14	1.00	1.00	1.03
mon0	ln1d	1.23	0.80	0.58	1.62	0.51	0.76	0.35	2.13	0.82	0.80	0.42	2.08
mon0	gap	1.52	1.29	0.66	1.82	2.54	1.58	0.37	3.09	3.30	1.99	0.50	4.40
rmon0	ln1d	1.33	0.78	0.75	1.52	0.89	0.82	0.49	1.77	1.22	0.92	0.47	2.41
rmon0	gap	1.20	1.29	0.70	1.51	1.41	1.35	0.21	2.57	2.98	1.42	0.07	4.02
mon1	ln1d	0.90	1.23	0.93	1.03	0.74	1.00	1.01	1.84	0.92	0.93	0.97	3.42
mon1	gap	1.49	1.30	1.60	1.44	1.73	1.24	1.02	2.67	2.17	1.28	1.23	4.79
rmon1	ln1d	0.99	1.34	1.07	0.95	0.96	1.05	1.01	1.86	1.05	1.03	0.98	3.37
rmon1	gap	1.64	1.29	0.89	1.52	1.41	1.24	0.60	2.50	2.35	1.11	0.54	4.26
mon2	ln1d	1.19	1.15	0.43	0.87	1.08	0.70	0.15	0.88	1.30	0.91	0.41	0.65
mon2	gap	1.37	1.05	0.85	1.76	2.17	1.58	0.41	2.83	3.43	2.29	0.26	4.32
rmon2	ln1d	1.16	1.03	0.38	0.88	1.50	1.04	0.20	0.85	1.59	1.04	0.14	1.38
rmon2	gap	1.03	1.00	0.66	1.35	1.79	1.16	0.64	2.20	1.60	1.25	1.06	4.22
mon4	ln1d	1.01	1.30	0.94	30.45	1.57	1.16	1.66	21.61	2.98	1.37	2.82	46.55
mon4	gap	1.02	2.11	1.26	24.02	1.66	2.49	1.02	32.60	2.03	2.58	1.29	65.50
rmon4	ln1d	1.66	1.05	0.53	44.46	2.59	1.06	0.46	49.87	3.12	0.96	0.54	63.68
rmon4	gap	2.76	1.11	0.68	12.25	4.43	0.96	0.30	25.50	2.26	1.03	0.43	48.90
lended	ln1d	1.51	0.58	1.77	0.79	1.05	0.83	1.07	0.72	1.13	1.08	0.80	0.71
lended	gap	1.92	1.12	1.37	1.93	2.69	1.14	0.73	3.21	2.49	1.37	1.64	3.24
rlended	ln1d	1.39	0.82	1.07	0.75	1.76	0.94	0.76	0.65	1.96	1.18	1.30	0.46
rlended	gap	1.25	1.38	0.54	1.43	2.07	1.25	0.31	2.17	2.21	0.93	0.82	2.40
mip	lev	1.53	1.65	1.55	1.48	1.84	2.19	1.52	2.22	2.63	3.44	2.39	2.61
mip	gap	1.21	1.31	1.09	1.18	1.18	1.13	0.99	1.34	1.25	0.91	2.60	1.21
velo	ln1d	1.31	1.27	0.65	1.18	1.11	1.12	0.53	1.68	1.31	1.10	0.39	1.85
velo	gap	1.49	1.26	0.90	1.28	1.70	1.27	0.49	1.71	1.47	1.24	0.31	3.05
note	ln1d	1.09	0.65	0.80	1.57	1.03	0.87	0.45	2.51	0.97	0.92	0.54	2.75
note	gap	1.51	1.11	0.90	1.83	2.21	1.47	0.58	3.56	2.89	2.23	0.77	4.65
rnote	ln1d	1.06	0.65	0.89	1.51	1.19	0.79	0.57	2.22	1.18	0.94	0.69	3.04
rnote	gap	1.28	0.94	0.76	1.49	1.94	1.14	0.31	2.79	2.89	1.26	0.17	4.39
igbed	lev	3.03	0.51	1.94	1.40	1.91	0.48	1.90	2.04	1.87	1.29	2.53	3.68
rigbed	lev	0.88	1.13	1.29	1.23	1.14	1.16	1.20	1.15	1.68	1.21	1.03	0.90
sprded	lev	1.11	1.07	1.19	1.05	1.04	0.99	1.15	1.18	1.92	1.01	1.13	1.22
alndav	lev	1.26	0.56	2.54	1.44	1.11	0.47	2.80	2.04	1.49	0.93	2.88	3.11
raindav	lev	1.16	1.26	1.04	1.31	1.41	1.14	1.18	1.15	1.78	1.16	1.27	1.83
rateav	ln1d	7.11	1.14	1.00	0.99	4.63	1.05	0.94	0.89	0.74	0.93	0.96	0.92
rateav	gap	10.22	1.11	0.97	1.17	5.36	0.96	0.76	1.22	0.20	0.92	0.93	1.01
rateed	ln1d	5.64	1.01	1.02	1.07	4.30	1.00	0.96	1.09	0.60	1.00	1.00	0.95
rateed	gap	6.32	0.99	1.19	1.09	5.04	1.07	0.82	1.26	0.23	0.86	0.92	1.09
ratest	lev	1.31	1.25	1.06	0.91	1.98	1.20	1.05	0.87	2.08	1.40	1.00	0.76
ratesk	lev	1.08	1.08	1.02	0.99	0.98	1.03	1.00	0.95	0.97	1.01	0.99	0.99
ratekr	lev	1.91	1.06	0.98	1.24	0.95	0.98	1.00	1.27	0.92	1.00	0.97	1.51
rrateav	ln1d	6.16	1.08	0.99	0.97	3.27	1.05	0.94	0.90	0.39	1.01	1.00	0.98
rrateav	gap	7.34	0.91	1.28	1.14	4.69	0.83	0.98	1.20	0.27	0.87	1.13	1.04
rrateed	ln1d	6.29	1.03	1.02	1.06	3.15	1.06	0.92	0.96	0.43	0.97	1.04	1.04
rrateed	gap	7.61	0.85	1.14	1.22	4.65	0.84	0.99	1.20	0.27	0.87	1.15	1.08
efratav	ln1d	4.08	1.19	1.01	1.06	2.92	1.11	0.91	1.00	0.79	1.06	0.80	0.80
efratav	gap	4.73	1.13	0.88	1.09	3.19	1.30	0.71	0.96	0.21	1.24	0.81	0.98
efrated	ln1d	4.57	1.08	1.00	1.15	3.31	1.09	0.97	1.25	0.89	1.05	0.83	0.86
efrated	gap	4.07	1.15	1.10	1.32	3.24	1.27	0.74	1.10	0.26	1.19	0.81	1.01
refratav	ln1d	5.40	1.14	1.01	1.07	3.73	1.06	0.99	1.17	0.82	1.03	0.86	0.75
refratav	gap	4.92	1.23	0.88	1.22	4.30	1.24	0.82	0.95	0.19	1.20	0.80	0.80
refratav	ln1d	4.96	1.05	1.00	1.16	4.10	1.03	0.96	1.17	0.50	1.06	0.91	0.81
refratav	gap	5.16	1.15	1.12	1.21	4.43	1.24	0.87	1.17	0.28	1.17	0.81	0.83
nikav	ln1d	1.41	0.98	0.90	1.12	1.33	0.77	0.25	1.31	1.27	0.71	0.22	1.29
niked	ln1d	1.20	1.23	1.00	1.08	1.23	1.07	0.29	1.41	0.98	0.87	0.28	1.42
nikst	lev	1.26	1.84	2.04	0.95	1.38	1.83	1.26	1.15	1.11	2.68	2.29	1.09
niksk	lev	1.37	1.13	0.94	0.96	1.19	1.09	0.89	1.10	1.24	1.01	0.90	0.75
nikkr	lev	1.06	1.11	1.06	0.99	0.80	2.23	0.98	1.00	1.11	2.14	1.02	1.02
rnikav	ln1d	1.79	0.86	0.92	1.13	1.34	0.77	0.26	1.29	1.35	0.72	0.23	1.29
rniked	ln1d	1.64	1.08	1.00	1.08	1.59	0.97	0.30	1.40	1.10	0.83	0.31	1.43
tpxav	ln1d	1.31	1.20	0.75	1.13	1.23	0.82	0.27	1.25	1.25	0.66	0.27	1.12
tpxed	ln1d	1.38	1.10	1.13	1.07	0.99	0.96	0.30	1.29	1.11	0.72	0.26	1.23
rtpxav	ln1d	1.61	1.09	0.77	1.12	1.22	0.88	0.27	1.23	1.33	0.66	0.27	1.12
rtpxed	ln1d	1.68	1.28	1.13	1.07	1.18	0.89	0.31	1.28	1.13	0.73	0.28	1.23
toshov	ln	1.58	0.92	1.15	1.09	2.67	1.39	0.71	1.09	3.68	3.21	0.69	1.61
toshov	ln1d	1.85	1.01	0.98	0.89	1.77	1.10	0.86	1.05	1.23	0.91	0.66	1.43
land	ln1d	6.78	0.98	0.44	1.07	6.34	1.99	0.24	1.16	6.62	8.06	0.26	0.77
rland	ln1d	3.44	1.31	0.46	1.13	4.07	1.61	0.24	1.13	4.63	5.97	0.28	0.74

備考：1行目の計数は平方平均自乗予測誤差（RMSFE：root mean squared forecast error）。2行目以降は、相対MSFE。シャドーは相対MSFEが1未満の場合。なお、経済活動水準、資産価格で始期が1971～73年の系列については、1983～85年の予測値が欠損値となるため、1983～86年はこれらを含まない平均。

3. 多変数予測とその結果

前節の結果からは、個別の情報変数を用いた2変数予測の枠組みでは、CPI変化率、実質GDP成長率ともに、異なった予測期数やサンプル期間を通じて常に予測パフォーマンスを改善させることはできなかった。しかし、より多くの変数から予測に有用な情報を的確に取り出せば、予測パフォーマンスを向上させられる可能性がある。この観点から、本節では、多くの情報変数を用いた予測手法を整理したうえで、それらの推計結果について比較する。

(1) 予測手法の整理

前節の(1)式において、利用可能な情報変数が1つでなく n 種類 (X^1, X^2, \dots, X^n) 存在する場合に、多くの情報変数から得られる情報を集約することで y_{t+h}^h の予測を行う手法を検討しよう。具体的な集約方法にはいくつかのバリエーションがあり得るが、それらは、1節でみたとおり、情報の集約を予測の前に行うか後に行うかに応じて、「指数アプローチ」と「予測統合アプローチ」に分けることができる。本稿では、Stock and Watson [1999] に従い、前者として主成分分析を、後者として単純・加重平均のバリエーションを用いる方法を検討する。

イ．主成分による予測

まず、多数の情報変数から情報を抽出した指標を作成し、これを用いて予測を行う指数アプローチとして、本稿では主成分分析を採用する²³。具体的には、 n 個の情報変数 X^i ($i = 1, \dots, n$) に主成分分析を適用し、抽出された第1主成分の t 時点の値を $D^{(1)}$ とすれば、それは雑多なノイズを除去し、 n 変数の共通情報を抽出した「指数」とみることができる。そして、この $D^{(1)}$ を新たな情報変数として前節の2変数予測の枠組みにおいて利用し、 y_{t+h}^h の主成分による予測 (fctr) を作成した²⁴。

ロ．2変数予測の単純・加重平均

一方、多数の情報変数から個別の予測系列を多数作成したうえで、それを加重平均する予測統合アプローチとしては、加重平均のためのウエイトの作成方法によっていくつかのバリエーションが考えられる。

まず、比較的単純な手法として、同じ時点における n 個の2変数予測の単純平均 (mean)、中央値 (median)、刈込平均を作成した。刈込平均については、最大値・最小値を除外したもの (tr.mean)、最大値・最小値それぞれ2系列ずつ除外したも

23 なお、脚注10で触れたとおり、変数の数が多いときには、主成分による予測をダイナミック・ファクター・モデルとしても解釈できる。

24 第1～第4主成分も含めた予測モデルも検討したが、それらのパフォーマンスは、第1主成分のみを用いたモデルを下回るものが多かった (同様の結果はStock and Watson [1999] でも得られている)。このため、本稿では第1主成分のみを用いたモデルの結果のみ掲載している。

の(TM₋₂)、さらにnが大きい場合については、上位・下位系列をそれぞれ15%ずつ除外した刈込平均(TM₋₁₅)の3種類を作成した。

さらに別の方法として、過去のデータからよりパフォーマンスがよいと判断される予測に高いウエイトを割り当てる方法も考えられる。具体的には、可変ウエイトをリッジ回帰²⁵により推計した予測統合(ridge)を、以下のような枠組みから作成する。

情報変数 X^i による t 時点の予測値を $f_t^i = \hat{y}_{t+h}^{h,i}$ 、 t 時点において変数 i による予測値に割り当てるウエイト w_t^i としたうえで、 $f_t = (f_t^1, f_t^2, \dots, f_t^n)'$ 、 $w_t = (w_t^1, w_t^2, \dots, w_t^n)'$ 、あるパラメータ k のもとで $c = k \times TR(n^{-1} \sum_{s=1}^t f_s f_s')$ とすると($TR(\cdot)$ は行列の対角成分の和)、リッジ回帰による可変ウエイトの予測統合は以下のように決まる²⁶。

$$y_{t+h}^{h,Ridge} = \sum_{i=1}^n w_t^i f_t^i, \quad w_t = (cI_n + \sum_{s=1}^t f_s f_s')^{-1} (\sum_{s=1}^t f_s y_{t+h}^h + c/n). \quad (2)$$

なお、以下では $k=0.25, 0.5, 1, 10, 100, 500$ の場合を計算した²⁷。

(2) 多変数予測のパフォーマンス

次に、以上で整理した多変数予測手法のパフォーマンスを、わが国のデータを使って比較・検証する。なお、ここでは、情報の集約化を行う情報変数の範囲として、経済活動水準(react)、物価・賃金等(pr_ea)、マネー(money)、資産価格(asset)の4つの種類別、種類別予測に使われた約30変数をそのまま利用した部分予測(licatall)、全系列を利用した全体予測(liall)の計6通りを採用した(対象系列のリストは図表7)²⁸。これにより、情報変数の種類別での差違のほか、変数の特性を考慮せずに情報変数を増やしたときの予測の改善度合いを比較できる。

25 なお、以下ではStock and Watson [1999] に従い、通常のリッジ回帰ではなく、(2)式のようにパラメータ k が大きくなるにつれてウエイト w_t が $1/n$ に収束するように修正されたリッジ回帰を用いている。このリッジ回帰に基づく加重平均は、 k が大きいほど単純平均に近くなり、例えば $k=1$ で単純平均におよそ50%近く。

26 なお、リッジ回帰によるウエイト推計に当たっては、直近40四半期のデータのみを用いた場合、サンプル数の減少に伴い推計結果が不安定化し、予測パフォーマンスがむしろ悪化したため、データを始期から利用している。

27 Stock and Watson [1999] では、 $k=1$ のパフォーマンスが最もよいとして、他のパラメータでの推計結果は示されていない。

28 なお、種類別予測統合、部分予測統合に利用する変数は以下の手順に従って選択している(主成分抽出の対象変数も予測統合と同じ変数のグループを採用した)。

個別情報変数による2変数予測の全系列のうち、すべての四半期で予測値が利用可能な1987~90年、1991~94年、1995~99年の各予測先行期間のいずれかにおいて、パフォーマンスが顕著に劣る予測系列(相対MSFEが、いずれかの時期で2以上)を除外。

AR予測を改善している系列(相対MSFEが平均1.0未満)は統合対象とした。相対的にパフォーマンスがよい指標(相対MSFEが平均1.5未満)も、できるだけ統合する。なお、同一の変数を用いた異なる加工系列(期末と期中平均、対数階差とレベル等)は、パフォーマンスが最もよいもの1系列に絞り込んだ。

イ．CPI変化率予測

多変数によるCPI予測のパフォーマンスについてみると（図表8、表中の太字は相対MSFEを最小化するもの）、まず種類別では、物価・賃金関連（pr_ea）とマネー（money）の予測統合系列のパフォーマンスが総じて良好である。物価・賃金関連（pr_ea）は、単純平均、中央値、刈込平均、一部のリッジ回帰（ridge $k=100, 500$ ）では、すべてのサンプル期間・予測期数で予測パフォーマンスを改善させている。また、マネー（money）は、1987～90年のバブル拡大期を除けば、安定的に予測を改善している。また、時期別では1991～94年、1995～99年、予測期数別では、4、8期先予測のパフォーマンス改善が目立つ²⁹。

図表7 多変数による予測に用いた変数

(a) CPI

経済活動水準 (react)		物価・賃金・市況 (pr_ea)		マネー (money)		資産価格 (asset)	
kijmi	ln1d	ngdp	ln1d	mon2	ln1d	sprded	lev
kijhi	ln1d	pgdp	ln1d	rmon2	ln1d	ratest	lev
kenjal	ln1d	ipiav	ln1d	lended	ln1d	ratekr	lev
kenjmi	ln1d	earn	ln1d	rlended	ln1d	nikav	ln1d
kenjhi	ln1d	rearn	ln1d	velo	ln1d	nikkr	lev
ukeall	ln1d	lcglded	ln1d	rmote	ln1d	rnikav	ln1d
ukekun	ln1d					tpxav	ln1d
ukechi	ln1d					rtpxav	ln1d
ckhime	ln1d					tosho	ln
expt	ln1d						

備考：部分予測統合（licatall）は、上記31系列を一括して統合した系列。
全体予測統合（liall）は、図表4にあるすべて（148）の2変数予測系列を統合した系列。

(b) GDP

経済活動水準 (react)		物価・賃金・市況 (pr_ea)		マネー (money)		資産価格 (asset)	
capu	ln1d	wpi	ln1d	rmon0	ln1d	rateav	ln1d
kijmi	ln1d	wpiined	ln1d	mon2	ln1d	rrateav	ln1d
kijhi	ln1d	ipiav	ln1d	rmon2	ln1d	efratav	ln1d
kenjal	ln1d	ipisoav	ln1d	lended	ln1d	nikav	ln1d
kenjmi	ln1d	reued	ln1d	rlended	ln1d	niksk	lev
juckko	gap	rreued	ln1d	velo	ln1d	rnikav	ln1d
juckme	gap	glded	ln1d			tpxav	ln1d
car	gap	rglded	ln1d			rtpxav	ln1d
lcexpt	gap					tosho	ln1d

備考：部分予測統合（licatall）は、上記32系列を一括して統合した系列。
全体予測統合（liall）は、図表6にあるすべて（147）の2変数予測系列を統合した系列。

この選別を、変数の各種類（経済活動水準、物価・賃金等、マネー、資産価格関連）について行った。この際、1種類に含まれる候補が6系列を下回る場合は、ある時期・予測先行期間の相対MSFEが2を越えていても、別の時期・先行期間には大幅な改善がみられる系列を、追加的に各種類の候補に追加。その結果、各種類につき、6～14系列の2変数予測を統合対象とした。

部分予測は、までの各種類の候補を、種類を問わずまとめて統合したものである。CPI変化率予測については148系列中31系列、実質GDP変化率予測については147系列中32系列をそれぞれ抽出している。

29 例えば、物価・賃金等の予測統合（pr_ea mean）は、 $h=2$ においてすべての2変数予測を上回る改善（ -0.280 ）を示しているほか、 $h=2,4$ の95～99年など、名目GDP・名目賃金による2変数予測より改善幅が大きい時期もある。

図表8 多変数による予測の相対MSFE CPI予測

		h=2				h=4				h=8			
		83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99	83-86	87-90	91-94	95-99
react	mean	0.95	1.12	0.94	0.92	0.95	1.08	0.86	0.64	0.98	1.05	0.90	0.62
react	median	0.98	1.05	0.95	0.96	0.92	1.10	0.92	0.77	1.03	1.07	0.94	0.71
react	tr.mean	0.93	1.10	0.94	0.98	0.96	1.10	0.92	0.67	1.03	1.06	0.94	0.65
react	TM ₋₂	0.96	1.08	0.93	0.96	0.96	1.09	0.92	0.72	1.12	1.04	0.94	0.68
react	ridge k=0.25	0.69	1.75	0.90	0.73	0.65	2.90	1.01	0.42	0.91	6.80	0.57	0.36
react	ridge k=0.5	0.69	1.70	0.89	0.72	0.65	2.88	0.94	0.41	0.91	6.54	0.58	0.35
react	ridge k=1	0.69	1.64	0.87	0.72	0.65	2.75	0.89	0.41	0.91	6.10	0.59	0.36
react	ridge k=10	0.77	1.25	0.85	0.80	0.73	1.51	0.77	0.50	0.91	2.87	0.68	0.46
react	ridge k=100	0.89	1.12	0.91	0.90	0.85	1.07	0.83	0.61	0.91	1.21	0.85	0.59
react	ridge k=500	0.92	1.12	0.93	0.92	0.88	1.07	0.85	0.63	0.91	1.07	0.89	0.61
react	fctr	1.52	1.25	1.07	1.19	1.02	1.53	0.89	0.65	0.96	1.49	0.83	0.41
pr_ea	mean	0.77	0.87	0.63	0.66	0.73	0.87	0.66	0.54	0.71	0.69	0.71	0.53
pr_ea	median	0.95	0.84	0.66	0.74	0.94	0.79	0.67	0.58	0.77	0.69	0.80	0.57
pr_ea	tr.mean	0.92	0.84	0.62	0.71	0.89	0.80	0.67	0.54	0.76	0.70	0.79	0.56
pr_ea	TM ₋₂	0.95	0.84	0.66	0.74	0.94	0.79	0.67	0.58	0.77	0.69	0.80	0.57
pr_ea	ridge k=0.25	0.32	1.38	0.75	0.54	0.14	2.22	0.88	0.33	0.41	5.25	0.43	0.09
pr_ea	ridge k=0.5	0.33	1.34	0.73	0.52	0.13	2.09	0.84	0.33	0.41	4.67	0.42	0.11
pr_ea	ridge k=1	0.33	1.28	0.70	0.52	0.14	1.90	0.79	0.33	0.43	3.99	0.39	0.13
pr_ea	ridge k=10	0.50	0.94	0.58	0.59	0.35	1.05	0.58	0.43	0.56	1.27	0.43	0.32
pr_ea	ridge k=100	0.72	0.87	0.62	0.65	0.66	0.87	0.63	0.52	0.68	0.70	0.65	0.49
pr_ea	ridge k=500	0.76	0.87	0.63	0.66	0.71	0.87	0.65	0.54	0.70	0.69	0.69	0.52
pr_ea	fctr	0.66	1.21	0.67	1.34	0.70	1.08	0.53	1.40	0.47	0.83	0.19	0.61
money	mean	0.87	1.22	0.93	0.87	0.85	1.86	0.68	0.61	0.38	3.26	0.48	0.34
money	median	0.84	1.02	0.93	0.91	0.74	1.63	0.66	0.66	0.38	2.76	0.51	0.32
money	tr.mean	0.85	1.12	0.94	0.86	0.83	1.66	0.67	0.60	0.37	2.92	0.48	0.31
money	TM ₋₂	0.84	1.02	0.93	0.91	0.74	1.63	0.66	0.66	0.38	2.76	0.51	0.32
money	ridge k=0.25	0.40	1.92	0.85	0.72	0.32	2.72	1.02	0.45	0.16	4.72	0.46	0.21
money	ridge k=0.5	0.44	1.79	0.86	0.72	0.33	2.50	0.99	0.46	0.17	3.88	0.43	0.21
money	ridge k=1	0.48	1.67	0.86	0.73	0.35	2.29	0.93	0.46	0.18	3.21	0.40	0.21
money	ridge k=10	0.66	1.31	0.84	0.80	0.55	1.71	0.67	0.54	0.28	2.32	0.36	0.27
money	ridge k=100	0.83	1.23	0.91	0.86	0.79	1.82	0.67	0.60	0.36	3.05	0.46	0.32
money	ridge k=500	0.86	1.22	0.93	0.86	0.83	1.85	0.67	0.61	0.38	3.21	0.48	0.33
money	fctr	0.81	1.68	1.02	0.84	0.78	2.71	0.71	0.62	0.33	6.47	0.54	0.28
asset	mean	1.08	1.03	1.07	1.00	1.08	0.91	1.03	0.72	0.83	0.86	1.10	0.92
asset	median	1.22	1.12	1.04	1.20	1.32	0.94	1.06	0.75	0.88	0.92	1.08	1.03
asset	tr.mean	1.13	1.06	1.08	1.05	1.16	0.93	1.08	0.75	0.86	0.94	1.11	0.97
asset	TM ₋₂	1.19	1.07	1.08	1.10	1.25	0.91	1.09	0.78	0.85	0.90	1.11	0.99
asset	ridge k=0.25	0.54	1.63	0.81	0.74	0.60	2.91	0.88	0.46	0.91	6.84	0.80	0.51
asset	ridge k=0.5	0.52	1.61	0.80	0.74	0.60	2.85	0.85	0.47	0.91	6.37	0.74	0.51
asset	ridge k=1	0.52	1.56	0.79	0.75	0.61	2.69	0.83	0.48	0.91	5.73	0.71	0.52
asset	ridge k=10	0.72	1.19	0.86	0.86	0.70	1.43	0.81	0.58	0.91	2.41	0.83	0.70
asset	ridge k=100	0.97	1.04	1.03	0.97	0.82	0.94	0.97	0.69	0.91	1.00	1.04	0.88
asset	ridge k=500	1.02	1.03	1.06	0.99	0.84	0.91	1.02	0.72	0.91	0.88	1.09	0.91
asset	fctr	1.04	1.14	1.21	1.32	0.99	0.90	1.20	0.68	0.90	0.89	1.18	0.74
licatall	mean	0.84	1.03	0.87	0.80	0.80	1.07	0.80	0.54	0.65	1.04	0.81	0.56
licatall	median	0.95	1.01	0.92	0.92	0.89	1.06	0.91	0.62	0.75	1.04	0.94	0.83
licatall	tr.mean	0.85	1.03	0.88	0.83	0.82	1.04	0.82	0.55	0.67	1.05	0.82	0.57
licatall	TM ₋₂	0.87	1.02	0.89	0.86	0.83	1.03	0.83	0.57	0.69	1.02	0.83	0.59
licatall	TM _{.15}	0.89	1.01	0.89	0.88	0.85	1.02	0.84	0.59	0.71	0.99	0.85	0.63
licatall	ridge k=0.25	0.65	1.43	0.69	0.57	0.64	2.66	0.87	0.52	0.98	7.18	0.71	0.23
licatall	ridge k=0.5	0.65	1.51	0.71	0.59	0.64	2.75	0.85	0.43	0.98	6.96	0.64	0.24
licatall	ridge k=1	0.66	1.56	0.73	0.61	0.64	2.79	0.84	0.40	0.98	6.65	0.60	0.27
licatall	ridge k=10	0.70	1.39	0.75	0.67	0.68	2.13	0.76	0.41	0.98	4.39	0.55	0.36
licatall	ridge k=100	0.83	1.07	0.81	0.75	0.80	1.11	0.74	0.50	0.98	1.49	0.70	0.49
licatall	ridge k=500	0.89	1.03	0.86	0.79	0.86	1.06	0.78	0.53	0.98	1.10	0.78	0.54
licatall	fctr	1.13	1.12	1.29	0.96	1.10	0.92	1.22	0.48	1.34	1.22	1.11	0.35
liall	mean	0.53	1.01	0.85	1.04	0.55	1.22	0.76	1.07	0.62	1.72	0.77	1.22
liall	median	0.72	1.01	0.89	0.97	0.74	1.10	0.84	0.92	0.80	1.15	0.87	1.03
liall	tr.mean	0.54	1.01	0.85	1.02	0.56	1.23	0.76	1.01	0.64	1.67	0.77	1.18
liall	TM ₋₂	0.55	1.01	0.85	1.02	0.57	1.22	0.77	0.97	0.65	1.66	0.78	1.15
liall	TM _{.15}	0.63	1.01	0.87	1.01	0.65	1.18	0.80	0.92	0.72	1.39	0.82	1.09
liall	ridge k=0.25	0.63	1.55	0.71	0.74	0.42	2.40	1.31	1.25	0.78	7.16	0.70	0.46
liall	ridge k=0.5	0.60	1.47	0.70	0.79	0.42	2.40	1.28	1.14	0.78	7.14	0.67	0.29
liall	ridge k=1	0.56	1.41	0.71	0.81	0.42	2.40	1.21	1.00	0.78	7.10	0.63	0.23
liall	ridge k=10	0.49	1.32	0.74	0.84	0.42	2.33	0.82	0.75	0.78	6.32	0.55	0.50
liall	ridge k=100	0.49	1.10	0.75	0.91	0.45	1.52	0.66	0.86	0.78	3.82	0.64	0.92
liall	ridge k=500	0.53	1.01	0.81	0.99	0.50	1.20	0.70	0.99	0.78	2.29	0.72	1.12

備考：1. シャドーは相対MSFEが1未満の場合。

2. 太字の計数は、あるサンプル期間・予測先期数・統合変数の種類のもとで、単純平均・各加重平均（中央値、刈込平均、リッジ帰帰）・主成分分析による統合を比較し、予測誤差が最小となった統合方法を意味する。なお、1、2、3、4列目が1983～86、1987～90、1991～94、1995～99年に対応している。
3. 1983～86年については、経済活動水準、資産価格、部分統合、全体統合では欠損値を含む時点を除いて統合。

しかしながら、利用する情報変数を単に増やしても、予測パフォーマンスが改善するとは限らない。事実、140余りの全変数を用いた全体予測 (liall) の予測パフォーマンスは、ほぼすべてのサンプル期間について、約30系列からなる部分予測 (licatall) を下回っている。また、予測の統合 (および主成分の抽出) に際し変数の種類を区別していない部分予測 (licatall) は、多くのサンプル期間で、物価・賃金関連の種類別予測 (pr_ea) を下回る予測パフォーマンスしか示していない。

この間、統合方法別のパフォーマンスの違いをみると、サンプル期間によって改善度合いが異なり、どれが望ましいということは言えない。例えば、最もAR予測を改善しているサンプル期間が多い物価・賃金 (pr_ea) について、単純平均、中央値、刈込平均およびリッジ回帰 (ridge $k=100, 500$) による予測統合を比べると、各サンプル期間および予測期先数によって相対MSFEをもっとも改善する統合方法 (図表8の太字のケース) は異なっており、どれか1つが他より常に優れたパフォーマンスを示しているわけではない。他方、リッジ回帰 (ridge $k=0.25, 0.5, 1$) は、バブル崩壊期・ポストバブル期の2期間は連続して他より優れたパフォーマンスを示しているものの、その前のバブル生成期ではAR予測を大幅に悪化させており、サンプル期間の変更に対して不安定である³⁰。

ロ．実質GDP成長率予測

次に、実質GDPの多変数による予測のパフォーマンスをみると (図表9) まず、種類別予測 (react, pr_ea, money, asset) の中では、すべてのサンプル期間、予測期数で相対MSFEが1以下となった系列はみられなかった。しかし、マネーによる予測は、ほとんどの予測期数で1987年以降3時期を通じて予測を改善しているほか、改善幅も大きい³¹。また、資産価格による予測 (asset mean) も、ほとんどの予測期数で1987年以降3時期を通じて予測を改善しており、2変数予測に比べて、より安定かつ大幅に予測パフォーマンスを改善させている。

他方、CPI同様、利用する情報変数の増加は必ずしも予測改善につながらないことも確認される。すなわち、部分予測 (licatall) については、全体予測 (liall) より改善幅が大きいほか、特に $h=8$ ではすべての時期で予測を改善している。また、1987～99年では、マネー、資産価格による予測の方が全体予測に比べ総じて改善幅が大きい。

次に、多変数による予測方法ごとのパフォーマンスの違いについてみると、GDP予測の場合はCPI予測に比べて、主成分 (fctr) やリッジ回帰 (ridge)、単純平均 (mean) で持続的にほかより優れたパフォーマンスを示す場合もあった。具体的には、

30 なお、リッジ回帰は、経済活動水準 (react)、マネー (money) 全体統合 (liall) の2期先予測・8期先予測において、1991～94年、1995～99年で他の予測統合よりパフォーマンスがよいが、1987～90年ではいずれもAR予測に比べて大幅に悪化している。

31 1987～90年の8期先予測では、一部の系列 (mean, tr.mean) で相対MSFEは1.00～1.01となっており、予測を改善させていないケースもみられる。

図表9 多変数による予測の相対MSFE GDP予測

		h=2				h=4				h=8			
		83~86	87~90	91~94	95~99	83~86	87~90	91~94	95~99	83~86	87~90	91~94	95~99
react	mean	1.07	0.82	0.82	0.87	1.07	0.84	0.60	0.93	0.96	0.83	0.76	1.03
react	median	1.13	0.84	0.85	0.92	1.02	0.88	0.58	0.94	1.05	0.96	0.79	0.91
react	tr.mean	1.11	0.86	0.83	0.88	1.04	0.88	0.59	0.92	0.97	0.85	0.77	1.02
react	TM_-2	1.11	0.85	0.84	0.90	1.05	0.88	0.59	0.92	1.01	0.90	0.79	1.00
react	ridge k=0.25	1.66	0.84	0.84	0.80	1.86	1.00	0.76	0.91	0.77	1.06	1.17	0.98
react	ridge k=0.5	1.69	0.87	0.86	0.82	1.84	0.97	0.74	0.90	0.77	1.03	1.14	0.92
react	ridge k=1	1.66	0.89	0.87	0.83	1.78	0.96	0.73	0.90	0.78	1.01	1.10	0.90
react	ridge k=10	1.22	0.86	0.86	0.86	1.34	0.89	0.66	0.91	0.81	0.90	0.93	0.95
react	ridge k=100	1.07	0.83	0.83	0.87	1.10	0.85	0.61	0.93	0.85	0.84	0.79	1.01
react	ridge k=500	1.07	0.82	0.82	0.87	1.07	0.85	0.60	0.93	0.85	0.83	0.77	1.02
react	fctr	1.33	0.76	0.85	1.00	1.66	0.73	0.47	1.04	1.58	0.65	0.37	1.90
pr_ea	mean	1.56	1.05	0.99	1.05	1.32	1.00	0.81	0.97	0.87	0.95	0.90	0.76
pr_ea	median	1.34	1.07	1.02	1.03	1.15	1.00	0.88	0.96	1.03	0.96	0.94	0.78
pr_ea	tr.mean	1.46	1.07	1.01	1.04	1.25	1.00	0.83	0.96	0.92	0.96	0.91	0.77
pr_ea	TM_-2	1.37	1.08	1.02	1.04	1.17	1.00	0.85	0.95	1.00	0.96	0.93	0.79
pr_ea	ridge k=0.25	2.73	1.53	0.78	0.94	2.79	1.48	0.76	0.88	1.91	1.36	1.08	0.60
pr_ea	ridge k=0.5	2.50	1.50	0.78	0.94	2.65	1.45	0.75	0.87	1.87	1.34	1.07	0.62
pr_ea	ridge k=1	2.24	1.46	0.79	0.95	2.46	1.42	0.75	0.87	1.78	1.31	1.06	0.63
pr_ea	ridge k=10	1.41	1.23	0.89	1.00	1.45	1.19	0.78	0.91	1.18	1.12	0.97	0.69
pr_ea	ridge k=100	1.48	1.08	0.97	1.04	1.28	1.03	0.81	0.96	0.90	0.98	0.91	0.74
pr_ea	ridge k=500	1.54	1.06	0.98	1.05	1.31	1.01	0.81	0.96	0.87	0.96	0.90	0.75
pr_ea	fctr	1.55	1.12	1.01	0.95	2.29	1.08	1.02	1.26	1.13	0.96	1.09	0.90
money	mean	1.04	0.82	0.38	0.80	0.99	0.85	0.16	0.80	1.26	1.01	0.29	0.78
money	median	1.16	0.79	0.37	0.85	1.06	0.87	0.19	0.93	1.23	0.99	0.27	0.73
money	tr.mean	1.09	0.82	0.36	0.81	1.05	0.87	0.17	0.80	1.25	1.00	0.27	0.71
money	TM_-2	1.16	0.79	0.37	0.85	1.06	0.87	0.19	0.93	1.23	0.99	0.27	0.73
money	ridge k=0.25	2.46	1.08	0.39	0.80	2.55	1.00	0.16	0.84	1.97	1.36	0.39	0.91
money	ridge k=0.5	2.34	1.08	0.38	0.79	2.58	1.05	0.16	0.82	1.94	1.35	0.39	0.87
money	ridge k=1	2.16	1.06	0.38	0.79	2.46	1.06	0.16	0.81	1.88	1.32	0.38	0.84
money	ridge k=10	1.25	0.92	0.38	0.79	1.35	0.95	0.16	0.80	1.47	1.13	0.33	0.79
money	ridge k=100	1.05	0.83	0.38	0.80	1.02	0.86	0.16	0.80	1.29	1.02	0.29	0.78
money	ridge k=500	1.04	0.82	0.38	0.80	1.00	0.85	0.16	0.80	1.26	1.01	0.29	0.78
money	fctr	1.24	0.94	0.35	1.60	1.03	0.88	0.10	1.34	1.11	0.99	0.17	1.03
asset	mean	1.83	0.99	0.59	1.00	1.37	0.83	0.35	0.99	1.06	0.69	0.42	0.81
asset	median	1.53	1.02	0.96	1.04	1.24	0.95	0.68	1.05	1.24	0.75	0.60	0.98
asset	tr.mean	1.81	1.00	0.61	1.01	1.36	0.83	0.37	0.99	1.17	0.69	0.44	0.80
asset	TM_-2	1.68	1.00	0.63	1.02	1.30	0.84	0.40	1.01	1.19	0.70	0.47	0.80
asset	ridge k=0.25	2.82	0.94	0.78	1.04	1.87	0.90	0.42	1.21	0.81	0.86	0.85	1.25
asset	ridge k=0.5	2.87	0.97	0.75	1.04	1.80	0.93	0.43	1.16	0.81	0.84	0.84	1.10
asset	ridge k=1	2.89	0.99	0.73	1.03	1.73	0.95	0.45	1.11	0.81	0.81	0.82	0.98
asset	ridge k=10	2.40	1.00	0.66	1.01	1.47	0.89	0.41	1.02	0.81	0.68	0.62	0.83
asset	ridge k=100	2.01	0.99	0.60	1.00	1.34	0.83	0.36	1.00	0.81	0.67	0.45	0.81
asset	ridge k=500	1.96	0.99	0.59	1.00	1.32	0.83	0.35	0.99	0.81	0.68	0.43	0.81
asset	fctr	3.27	1.05	0.65	1.10	2.38	0.66	0.25	1.08	0.85	0.55	0.24	1.02
licatall	mean	1.24	0.90	0.60	0.90	1.07	0.85	0.41	0.84	0.95	0.82	0.52	0.68
licatall	median	1.13	0.97	0.83	0.99	1.02	0.94	0.52	0.93	1.06	0.94	0.71	0.75
licatall	tr.mean	1.22	0.92	0.63	0.91	1.07	0.86	0.44	0.84	0.97	0.84	0.55	0.69
licatall	TM_-2	1.19	0.92	0.65	0.92	1.06	0.87	0.46	0.86	0.99	0.86	0.59	0.71
licatall	TM_15	1.15	0.93	0.70	0.94	1.05	0.89	0.49	0.90	1.02	0.89	0.62	0.73
licatall	ridge k=0.25	2.91	0.73	0.50	0.81	1.81	1.06	0.38	0.93	0.86	1.36	1.06	0.96
licatall	ridge k=0.5	2.72	0.78	0.52	0.83	1.75	1.06	0.40	0.90	0.86	1.33	1.01	0.86
licatall	ridge k=1	2.58	0.83	0.57	0.85	1.69	1.06	0.44	0.87	0.86	1.31	0.99	0.76
licatall	ridge k=10	2.21	0.91	0.68	0.90	1.48	1.01	0.51	0.84	0.86	1.12	0.88	0.64
licatall	ridge k=100	1.64	0.90	0.63	0.90	1.21	0.89	0.45	0.84	0.86	0.86	0.63	0.66
licatall	ridge k=500	1.50	0.90	0.61	0.90	1.14	0.86	0.42	0.84	0.86	0.83	0.55	0.67
licatall	fctr	2.72	0.95	0.66	0.83	2.88	0.76	0.27	0.64	0.58	0.50	0.28	0.56
liall	mean	1.12	0.95	0.81	1.11	1.10	0.95	0.58	1.19	1.04	1.05	0.70	1.14
liall	median	1.07	0.99	0.94	1.01	1.08	0.96	0.74	1.04	1.07	0.99	0.85	1.00
liall	tr.mean	1.13	0.95	0.81	1.07	1.09	0.95	0.58	1.14	1.04	1.03	0.70	1.08
liall	TM_-2	1.13	0.95	0.82	1.04	1.09	0.95	0.58	1.11	1.04	1.02	0.71	1.05
liall	TM_15	1.06	0.96	0.86	1.01	1.07	0.96	0.65	1.07	1.07	0.99	0.79	1.02
liall	ridge k=0.25	2.85	0.74	0.85	2.86	1.94	1.33	0.69	1.41	0.96	1.76	1.20	1.82
liall	ridge k=0.5	2.67	0.79	0.84	2.50	1.88	1.27	0.67	1.41	0.96	1.73	1.25	1.66
liall	ridge k=1	2.51	0.83	0.86	2.06	1.79	1.24	0.67	1.37	0.96	1.71	1.29	1.48
liall	ridge k=10	2.24	0.93	0.97	1.27	1.52	1.16	0.76	1.19	0.96	1.64	1.35	1.04
liall	ridge k=100	1.79	0.93	0.93	1.13	1.28	1.05	0.71	1.16	0.96	1.31	1.09	1.01
liall	ridge k=500	1.43	0.93	0.85	1.12	1.11	0.98	0.63	1.17	0.96	1.10	0.83	1.09

備考：1. シャドローは相対MSFEが1未満の場合。

2. 太字の計数は、あるサンプル期間・予測先期数・統合変数の種類のもとで、単純平均・各加重平均（中央値、刈込平均、リッジ回帰）・主成分分析による統合を比較し、予測誤差が最小となった統合方法を意味する。なお、1、2、3、4列目が1983~86、1987~90、1991~94、1995~99年に対応している。
3. 1983~86年については、経済活動水準、資産価格、部分統合、全体統合では欠損値を含む時点を除いて統合。

部分予測統合 (licatall) において、主成分 (fctr) の4期先・8期先予測、リッジ回帰 (ridge $k=0.25$) の2期先予測において、1987～99年までの3期間連続でほかに比べて相対MSFEが小さくなっている。また、単純平均も、マネー (money) の4期先予測、および物価等 (pr_ea) の8期先予測では、4つのサンプル期間すべてにおいてAR予測を改善すると同時に、4期間のうち3期間でもっとも優れたパフォーマンスとなっている。しかしながら、これらの多変数予測の統合方法も、統合 (主成分抽出) の対象となる変数や期先予測数の変化に対しては不安定なものであることから、全体で見ればCPI予測同様、どの方法が望ましいかは一意的には決められないと考えられる。

八．まとめ：多変数予測

以上から、CPI予測、GDP予測ともに、多変数による予測は2変数予測と比べて安定的に予測パフォーマンスを改善することが確認できた。ただし、統合方法 (単純平均、刈込平均、リッジ回帰、主成分) 間での優劣は、対象変数の種類や予測期先数によって異なり、一意的には決めることができないと考えられる³²。また、利用する情報変数を単に増やすだけでは予測パフォーマンスは必ずしも改善しないこともわかった。

4．予測統合によるパフォーマンス改善の背景とその限界

前節では、多くの変数を利用した予測のパフォーマンスを検討し、これらが2変数予測に比べ全体として高いパフォーマンスを示すことを確認した。また、サンプル期間をいくつかに分けてパフォーマンスを評価し、多くの変数を利用した予測のパフォーマンスは、サンプル期間を通じて安定していることをみた。

こうした結果は、多くの変数を利用した予測ではより多くの情報が反映されているため、予測パフォーマンスが向上すると直感とも整合的である。しかし、統合対象系列 (もしくは主成分分析の対象変数) の拡大が常に予測パフォーマンスの向上につながっているわけではない。

そこで本節では、どのようなメカニズムによって、多変数による予測のパフォーマンス改善がもたらされているのかを検討する。なお、前節の結果のとおり、情報の集約方法として単純平均、加重平均、主成分分析のうちどれが最も望ましいかは、必ずしも明らかではない。このため、ここでは、最も分析の行いやすい、単純平均

32 この見解は、主成分による予測のパフォーマンスが最もよいとしたStock and Watson [1999] の結論と対照的にみえる。もっとも、彼らの分析結果でも統合方法間でのパフォーマンスの差はそれほど大きくないこと、また分析対象がインフレ率の1年先予測のみであることについて注意が必要である。また、Stock and Watson [2001] は、外れ値の影響を受けにくい刈込平均に焦点を当てているが、本稿の推計結果は、刈込平均が予測の統合方法として常に望ましいとは限らないことを示唆している。この点は、次節で詳しく検討するとおり、予測統合によるパフォーマンスの改善効果が高々2～4変数で出尽くしてしまうとの予測統合の限界とも整合的であるように思われる。

による予測統合に焦点を当てて検討を行うことにする³³。

(1) 予備的考察

まず、単純平均による予測統合と統合対象となる原予測のMSFE（平均自乗予測誤差）をいくつかの要因に分解し、単純平均による予測統合のパフォーマンスの高さと安定性が生み出されるメカニズムを整理しておく。

イ．MSFEの分解

個別変数による予測のMSFEは、予測バイアスによって生み出される部分と、予測誤差が変動することによって生み出される部分から構成される。すなわち、個別変数 i ($i=1, \dots, n$) を用いた予測モデルの t 期 ($t=1, \dots, T$) における（サンプル期間外）予測誤差を e_{it} とし、その t に関する平均値（すなわち予測バイアス）と分散をそれぞれ $\mu_i = 1/T \sum_{t=1}^T e_{it}$ 、 $\sigma_i^2 = 1/T \sum_{t=1}^T (e_{it} - \mu_i)^2$ とすると³⁴、この予測モデルの平均自乗予測誤差 $MSFE_i$ は、

$$MSFE_i = \frac{1}{T} \sum_{t=1}^T e_{it}^2 = \mu_i^2 + \sigma_i^2, \quad (3)$$

と書ける。

一方、単純平均による予測統合のMSFEも、予測バイアスによって生み出される部分と、予測誤差が変動することによって生み出される部分から構成されるが、後者をさらに、各原予測の予測誤差が変動することによって生み出される部分と、原予測の予測誤差同士が互いの動きを打ち消しあう部分に分けることができる。すなわち、単純平均による予測統合の t 期における予測誤差が $1/n \sum_{i=1}^n e_{it}$ である³⁵ことに注意すると、単純平均による予測統合の平均自乗予測誤差 $MSFE_{com, n}$ は、次のように表すことができる。

$$\begin{aligned} MSFE_{com, n} &= \frac{1}{T} \sum_{t=1}^T \left(\frac{1}{n} \sum_{i=1}^n e_{it} \right)^2 \\ &= \left(\frac{1}{n} \sum_{i=1}^n \mu_i \right)^2 + \left(\frac{1}{n} \sum_{i=1}^n \sigma_i \right)^2 - \sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{\sigma_i}{n} \frac{\sigma_j}{n} (1 - \rho_{ij}) \right\}. \end{aligned} \quad (4)$$

33 ただし、単純平均以外の多変数予測においても、基本的に以下でみるのと同じようなメカニズムが働いていることは容易に想像できよう。主成分分析による方法については、情報の集約が予測系列作成の前になるため以下の議論は直接適用できないが、ノイズが平準化されるという意味では、基本的に以下でみるのと同じようなメカニズムが働いていると考えられる。

34 e_{it} はサンプル期間外の予測誤差であるから、その平均値は必ずしもゼロとはならないことに注意されたい。

35 個別変数 i を用いた2変数予測の予測値を f_{it} とすると、単純平均による予測統合の予測値は $1/n \sum_{i=1}^n f_{it}$ である。予測対象変数の実現値を y_t とすると、 $e_{it} = f_{it} - y_t$ であるから、単純平均による予測統合の予測誤差は、 $1/n \sum_{i=1}^n f_{it} - y_t = 1/n \sum_{i=1}^n (f_{it} - y_t) = 1/n \sum_{i=1}^n e_{it}$ であることがわかる。

ただし、 $\rho_{ij} = 1/(\sigma_i \sigma_j T) \sum_{t=1}^T (e_{it} - \mu_i)(e_{jt} - \mu_j)$ ($i, j = 1, \dots, n$)は e_{it} と e_{jt} の間の相関係数である。

ここで、単純平均による予測統合の予測バイアスは $1/n \sum_{i=1}^n \mu_i$ であるから、右辺第1項は、予測バイアスによって生み出される平均自乗予測誤差に相当する。また、右辺第2項と第3項は、予測誤差が変動することによって生み出される部分であり、(3)式の右辺第2項と対応している。このうち、第3項は、原予測系列間の相関係数がすべて1の場合に0をとり、原予測誤差系列の変動に、1マイナス相関係数を乗じたものの総和になっている。1から相関係数を差し引いた値は、予測誤差系列が異なる動きをしている度合いと解釈でき、第3項は異なる動きをする原予測を統合することによって打ち消される平均自乗予測誤差を表している。これに対して第2項は、第3項がゼロの場合、つまり原予測誤差がすべて同じ動きをしている場合においても生じる予測誤差の変動であり、各原予測系列の予測誤差の変動によって生み出される平均自乗予測誤差を示している。

ロ．予測統合のMSFEと原予測系列の平均的なMSFEとの比較

次に、上で得たMSFEの分解式を用いて、単純平均による予測統合のMSFEを統合対象となる原予測系列の平均MSFEと比較してみよう。

(3)式の両辺を i について平均し、(4)式との差をとると、

$$MSFE_{com,n} - \frac{1}{n} \sum_{i=1}^n MSFE_i = - \left[\frac{1}{n} \sum_{i=1}^n \mu_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \mu_i \right)^2 \right] - \left[\frac{1}{n} \sum_{i=1}^n \sigma_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \sigma_i \right)^2 \right] - \sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{\sigma_i \sigma_j}{n} (1 - \rho_{ij}) \right\}, \quad (5)$$

となる。ここで、(5)式右辺の各項はすべて常に負の値をとる。まず、相関係数は必ず1以下となるため、第3項が常に負の値をとることは明らかである。一方、(5)式右辺第1項、第2項はそれぞれ、

$$\left[\frac{1}{n} \sum_{i=1}^n \mu_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \mu_i \right)^2 \right] = \frac{1}{n} \sum_{i=1}^n \left(\mu_i - \frac{1}{n} \sum_{j=1}^n \mu_j \right)^2, \quad (6)$$

$$\left[\frac{1}{n} \sum_{i=1}^n \sigma_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \sigma_i \right)^2 \right] = \frac{1}{n} \sum_{i=1}^n \left(\sigma_i - \frac{1}{n} \sum_{j=1}^n \sigma_j \right)^2, \quad (7)$$

と変形でき、これらは常に正となるため、両者の符号はともに負である。

したがって、

$$MSFE_{com,n} \leq \frac{1}{n} \sum_{i=1}^n MSFE_i, \quad (8)$$

が成立する³⁶。すなわち、単純平均による予測統合のパフォーマンスは、原予測の平均的なパフォーマンスを常に上回ることになる^{37、38}。また、パフォーマンス改善の程度は、(5)式の第1項から第3項によって決まる。

八．統合によるMSFE改善メカニズム

では、予測統合のパフォーマンスはなぜ、原予測の平均的なパフォーマンスを上回るのでしょうか。換言すれば、(5)式の第1項から第3項が表す、MSFEの縮小効果は、それぞれどのようなメカニズムによって生じているのだろうか。

まず、第3項は、上述のとおり、異なる動きをする原予測を統合することによって「予測誤差の変動が打ち消される効果」を表している。この効果は、原予測間の相関係数が小さいほど、すなわち、原予測の動きが互いに異なっているほど、大きくなる。

次に、第1項と第2項は、それぞれ、複数の原予測を平均することにより、「予測バイアスが平準化される効果」と「予測誤差の変動幅が平準化される効果」を表している。このことは、(4)式の右辺第1項、第2項が、それぞれ、原予測の予測バイアスの平均、原予測の予測誤差の標準偏差の平均を2乗する形になっていることからわかる。これらの項は、(6) (7)式から明らかなように、原予測のMSFEにおける対応する項の平均値を下回る。つまり、平均によって統合することで、予測バイアスと予測誤差の変動幅が平準化され、MSFEを縮小させることになる。

なお、平準化効果は、以下の理由から、予測誤差の変動幅に関するものよりも、予測バイアスに関するものの方が大きくなる可能性が高いと推測される。すなわち、(6) (7)式が示しているとおり、平準化することによる平均自乗予測誤差の縮小幅は、各原予測の予測バイアスや予測誤差の変動幅のばらつきが大きければ大きいほど大きくなる。他方、予測バイアス μ_i は正の値も負の値もとり得る一方で、予測誤差の標準偏差 σ_i は正の値しかとり得ないから、予測バイアスの方がばらつきが大きくなる場合が多いと予想される。したがって、平準化効果は、予測誤差の変動幅に関するものよりも、予測バイアスに関するものの方が大きくなる可能性が高いと考えられる。

36 等号成立条件は、(5)式右辺のすべての項がゼロとなることであり、この条件は、各期における原予測の予測値がすべて等しい ($e_{1t} = \dots = e_{nt}$ $t=1, \dots, T$) ことと同値である。

37 このため、個々の原予測のパフォーマンスが不安定であっても、それらを統合した予測のパフォーマンスは安定化する可能性が考えられる。

38 ちなみに、最適なウエイトを用いて統合した場合には、

$$MSFE_{com,n} \leq \min \{ MSFE_1, \dots, MSFE_n \},$$

が成立し (例えば、Clements and Hendry [1998] を参照) 予測統合はどの原予測のパフォーマンスをも上回る。ただし、「最適なウエイト」を予測時点では知ることは、前節の加重平均による予測統合の結果からわかるとおり、困難である。上で述べたことは、「最適なウエイト」を計算することができなくとも、単純に平均するだけで、原予測の平均的なパフォーマンス以上のパフォーマンスが保証される、ということの意味する。

以上より、単純平均による予測統合は、予測バイアスの平準化効果、予測誤差の変動幅の平準化効果、および各原予測の予測誤差同士が互いの挙動を打ち消しあう効果によって、パフォーマンスを向上させていることがわかった。これらの効果によって、予測統合のパフォーマンスは、原予測の平均的なパフォーマンスを常に上回ることになる。なお、平準化効果については、変動幅に関する効果よりも、予測バイアスに関する効果の方が大きいと予想される。

(2) 各要因の実際的な重要性

次に、上述した予測統合のMSFE縮小効果のメカニズムがどの程度働いているのか、また、3つのMSFE縮小効果のうち実際にはどの効果が支配的なのかを検証する。そのために、図表10として、(5)式の両辺をAR予測のMSFEで除した関係を使い、単純平均による予測統合系列の相対MSFE縮小効果を要因分解した結果を示した。

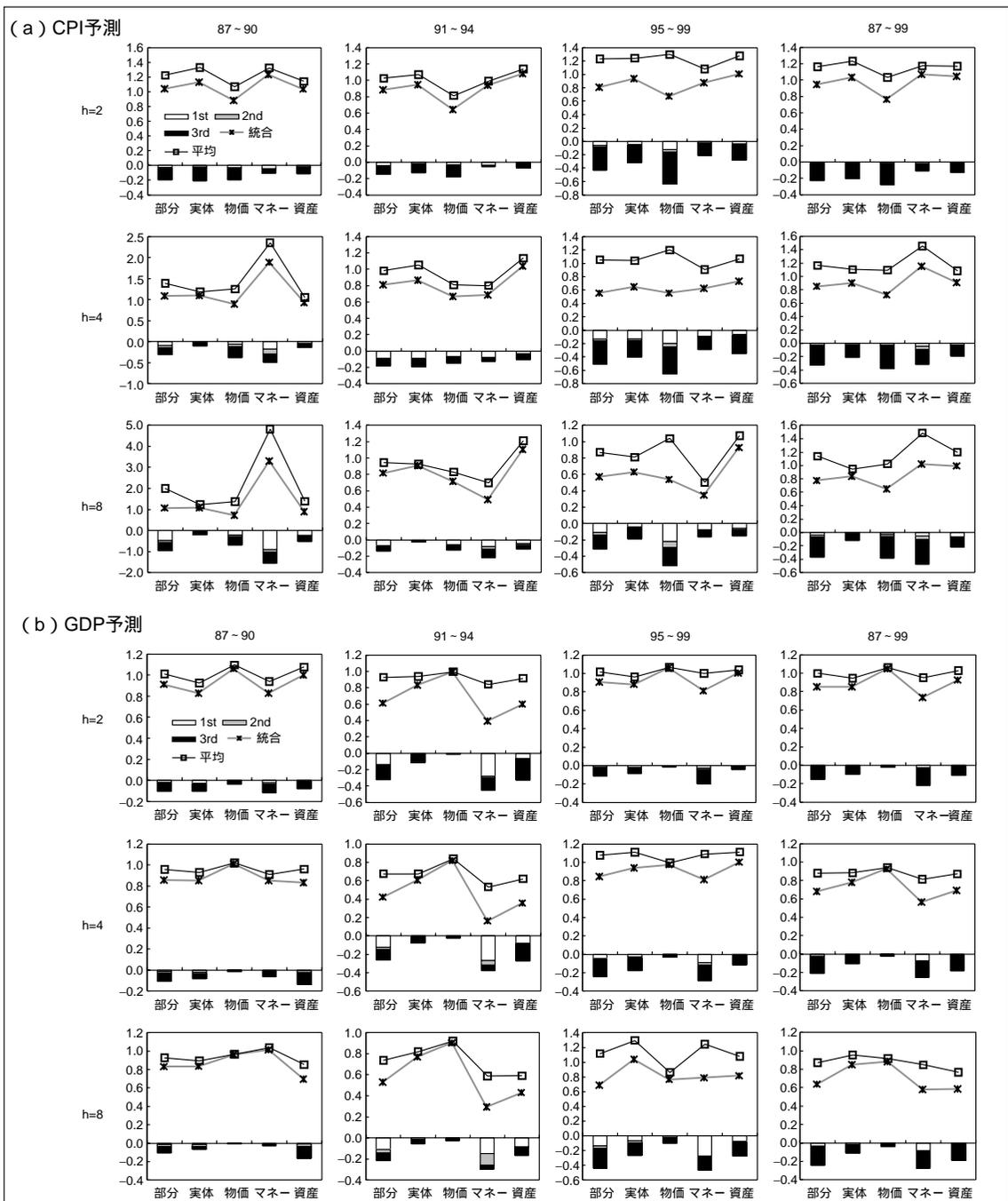
まず、通期での結果をみると(図表10の最右列)、統合による相対MSFEの縮小効果が大きいことを確認できる。予測統合の相対MSFEは、予測期数や変数の種類により大きさの違いがあるものの、原予測の相対MSFEの平均値を0.1~0.5程度下回っている。また、原予測の相対MSFEの平均値が1を上回っている場合にも、多くの場合、予測統合の相対MSFEは1を下回っている。例えば、4期先CPI予測では、部分統合、実体指標、物価関連指標のすべてが、原予測の相対MSFE平均値が1を上回っているが、予測統合の相対MSFEはいずれも1を下回っている。

また、MSFE縮小効果の内訳をみると、こうしたMSFE縮小効果のほとんどが(5)式の第3項の効果によってもたらされていることがわかる。すなわち、単純平均による予測統合に関しては、MSFE縮小効果の大部分が、各原予測系列の予測誤差の変動が互いに打ち消しあう効果によってもたらされている。この間、第1項と第2項の効果は、第3項の効果と比べて、ともにかなり小さいが、両者のうちでは第1項の効果、すなわち予測バイアスの平準化効果の方が相対的に大きくなっている。この結果は、前小節でたてた予想と整合的である。

次に、時期別の結果をみると(図表10の1~3列目)、まず、時期によって予測統合の相対MSFEと原予測の相対MSFEの平均値との差がかなり変動し、統合によるMSFE縮小効果の大きさは、時期によって異なることがわかる。例えば、CPI予測においては、1987~90年の期間や1995~99年の期間で大きなMSFE縮小効果が発生しているが、1991~94年の期間の縮小効果はやや小さい。また、GDP予測においては、MSFE縮小効果は1991~94年の期間と1995~99年の期間で大きい一方、1987~90年の期間においては小さい。

また、縮小要因の分解結果からは、第3項の効果が依然として大きいものの、第1項も大きな改善効果を生み出していることがわかる。さらに、通期ではほぼゼロであった第2項の効果も、ところどころに表れている。

図表10 予測統合の相対MSFE縮小効果の要因分解



備考：統合は各予測統合の相対MSFE、平均は各予測統合を構成する個別予測の相対MSFEの平均。1st、2nd、3rdは、本文（5式）にあるように、それぞれ予測バイアスの平準化、予測誤差変動幅の平準化、予測誤差変動が相互に打ち消しあう効果。

このように、通期ではみられない第1項、第2項の効果が時期別にみると表れてくる背景は、次のように考えることができる。すなわち、比較的長い期間でみると、原予測の予測バイアスや予測誤差の変動幅は、原予測間でそれほど大きくはばらつかなくなると考えられる。このため、第1項や第2項による平準化効果は、長いサンプル期間においては小さくなる。しかし、短い期間では、予測バイアスや予測誤差の変動幅の散らばりは大きくなり、第1項や第2項による平準化効果が大きくなる。

(3) 最適な予測統合：ステップワイズな統合対象の選択

最後に、以上のようなパフォーマンス改善のメカニズムが、どの程度の統合対象数まで効果を発揮するのかを検討しておこう。これまでの結果は、統合対象の変数が多ければ多いほどよいわけではないことを示している。そこで、こうした予測パフォーマンスの改善メカニズムの限界を探るために、「(単純平均による)最適な予測統合」がどのような系列から構成され、またどの程度のパフォーマンスを示すのかを検証する。

具体的には、ステップワイズ(逐次的)に1つずつ変数を取り込み、最良の予測パフォーマンスを示す組み合わせのみ残す方法によって、最適な統合対象の組み合わせを探索する³⁹。また、統合対象としては、種類別統合および部分予測統合で用いた系列(CPI予測では31系列、実質GDP予測では32系列)を候補とする。

イ．ステップワイズな最適化の手続き

ステップワイズな最適化による最善の予測統合の組み合わせの探索方法は、以下のとおりである。

- 1) まず、 K 個の変数($K=31, 32$)の中から、ある変数を「初期変数」として選ぶ。次に、それ以外の変数を1つだけ追加して、2変数の予測統合を作る。統合対象候補となる変数の数を K とすると、2変数の予測統合は $K-1$ 通り作ることができるが、このうち予測誤差が最も小さいものを保存する。
- 2) 次に、この2変数予測にそれ以外の変数を1つだけ追加して、3変数の予測統合を作る。この作り方は $K-2$ 通りあるが、このうち予測誤差が最も小さいものを保存する。同様に4変数の統合、5変数の統合と繰り返し、すべての候補系

39 本稿で採用しているステップワイズの最適化は、すべての統合の可能性を悉皆的に検証しておらず、最善の組み合わせに到達しているかは、必ずしも保証されていない。しかしながら、統合の組み合わせはCPIでは約21億($=2^{31}-1$)通り、GDP予測では約43億($=2^{32}-1$)通り存在し、これらをすべて検討することも計算量の大きさから現実的ではない。もっとも、以下の分析によって示されるとおり、ステップワイズによって探索された最適な統合系列は、高々2~4系列程度の予測統合である。そこで、計算量としても現実的な最大8系列までの統合の組み合わせについて、悉皆的に予測パフォーマンスをチェックしてみたところ、最善のパフォーマンスを示す組み合わせは、ステップワイズな探索結果と完全に一致することが確認できた。

列を統合対象に追加するまで続ける。

- 3) このプロセスでは、ある初期変数を所与として、変数を追加することに 初期変数、 統合に含まれる変数の数、 予測統合の誤差の対応関係が求まる。これらの対応関係の中で、予測誤差が最も小さくなる組み合わせを「初期変数を所与とした最適」とする。
- 4) 初期変数を所与とした最適な統合の組み合わせは初期変数の数だけ存在するが、そのうち相対MSFEが最小となる組み合わせを、「ステップワイズで最適」とする。以下では、これを「最適な統合の組み合わせ」と呼ぶ。

ロ．最適な統合の組み合わせと予測パフォーマンス

上述したステップワイズな最適化の手続きに従って、すべての予測期数・時期について、各初期変数ごとに最適な統合の組み合わせを図表11に示した。図表11上段のグラフには、横軸には初期変数を通期での予測パフォーマンス順に並べ替えた順に並べてある。また、折れ線は各初期変数を所与としたときの最適な組み合わせの相対MSFEを、各点（もしくは ）は統合対象となった系列数を示している。なお、ステップワイズで最適な統合に至った初期変数については、その最適統合系列数を示す点を黒くしている。図表11下段には、最適となる変数の構成を示している。

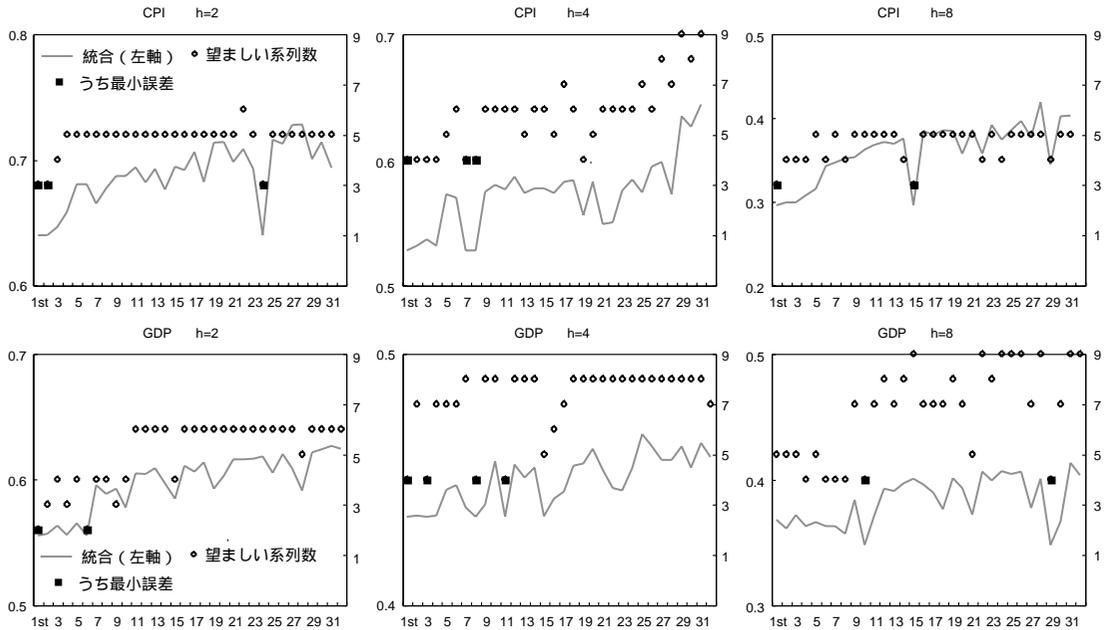
まず、通期（1987～99年）についてみると、最適な統合の組み合わせは（図表11上段のグラフ中の点 ）CPI、実質GDP予測ともに、すべての予測期数、時期の組み合わせで2～4つの原予測系列を統合したものとなっている。また、統合対象となっている原予測系列は、パフォーマンスの優れたものとなっているケースが多いが、必ずしもパフォーマンスの優れたもののみから構成されているわけではなく、むしろ中程度のパフォーマンスの原予測が含まれていることも多い。実際、CPI予測では、名目GDP（ngdp）が3つのサンプル期間すべてに共通して含まれているが、実質GDPでは、すべての期間に共通する指標はみられない。同様の傾向は、サンプル期間を分割した結果からも確認される（図表12）⁴⁰。

では、最適な組み合わせは、どの程度のパフォーマンスを示すのであろうか。まず通期についてみると（図表13）、最適な予測統合は、統合による改善効果だけでなく、原予測の単純平均自体が小さくなったこともあり、相対MSFEは0.2～0.6と、AR予測に比べれば4～8割の改善を示している。このうち、統合による改善効果は、種類別統合、部分統合と同様、原予測の相対MSFEの平均からみれば0.1～0.5ポイント程度となっている。この縮小効果のほとんどは(5)式の第3項、すなわち各原予測系列の予測誤差の変動が互いを打ち消しあう効果によってもたらされている。

40 CPI予測における1987～90年の4期先予測が、最もパフォーマンスのよい個別予測をどの予測統合も改善できない唯一のケースである。

図表11 最適な統合の組み合わせ

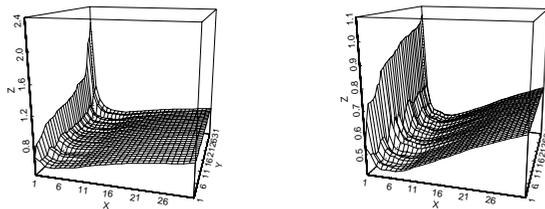
(a) 初期変数を所与とした最適な統合数と予測誤差



備考：横軸は、初期系列に選んだ個別予測のパフォーマンスの良さ（その初期系列が全系列中何番目にパフォーマンスが良いか）。右の縦軸は統合系列数、左の縦軸は相対MSFE。系列数（右軸）の各点（・）は、ある初期系列のもと、誤差が最小となる組み合わせの数。また、系列数のうちは、ステップワイズで最適となる統合数。また、予測誤差（左軸・相対MSFE）は、初期系列を所与とした場合に最小となる組み合わせを用いた場合の値。

なお、最善の予測組み合わせの探索過程は、 $(X, Y, Z) = (\text{統合に含まれる系列の数, 初期変数のパフォーマンス順位, 予測統合の誤差})$ の空間において曲面（下記参考図表）として表すことができる。このとき、ステップワイズで最適な (X, Y, Z) の組み合わせ（上図）は、曲面全体での「谷底」、「初期変数を所与とした最適」は、あるYで曲面を切断したX-Z平面における「谷底」に対応する。

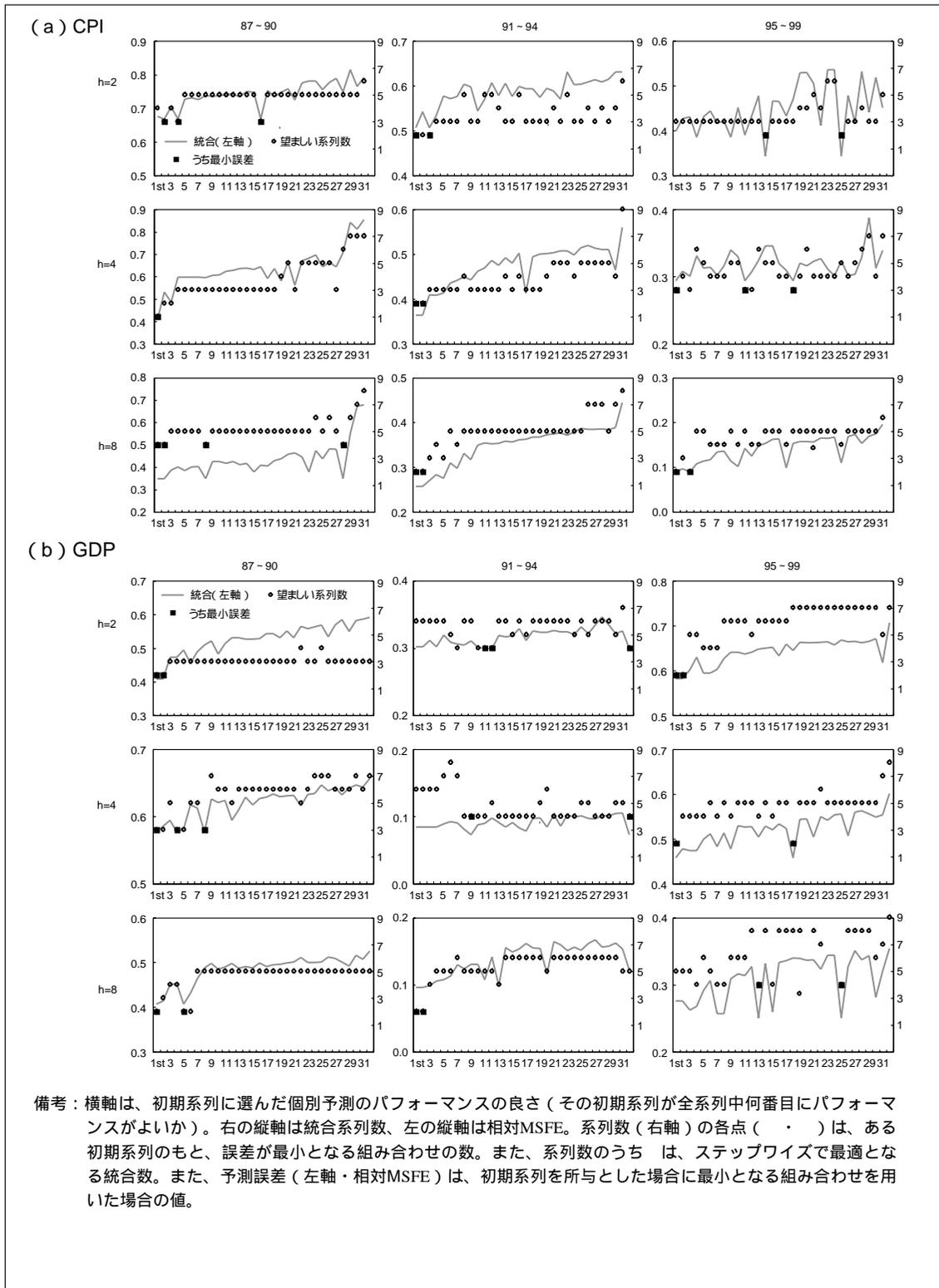
(参考図表) ステップワイズな最適化のイメージ



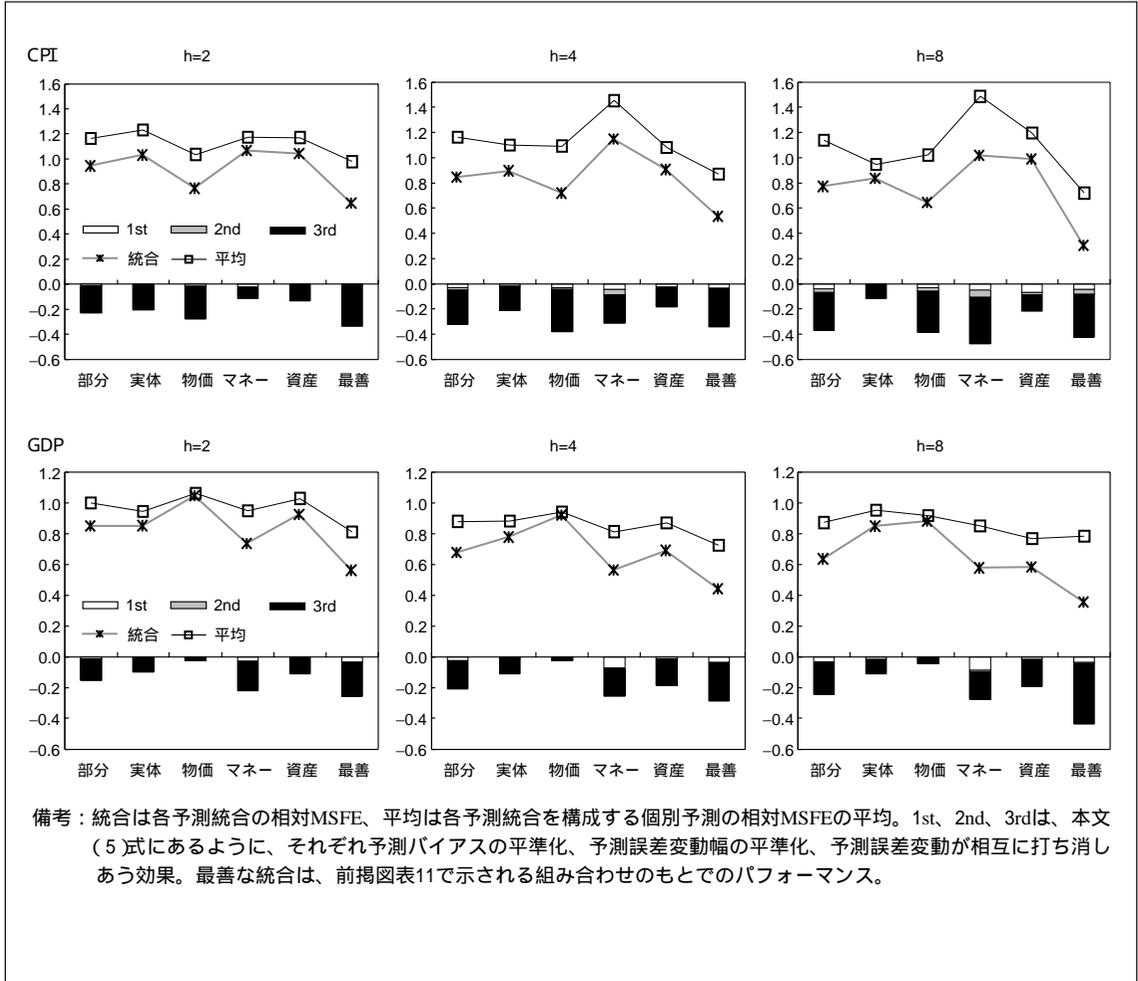
(b) 組み合わせの構成

		統合数	相対MSFE	統合系列の内訳 (インデクス)			個別系列のパフォーマンス順位				
CPI	h=2 87~99	3	0.64	ngdp	rearn	sprded	1	2	24		
	h=4 87~99	4	0.527	ngdp	earn	ratest	tosho	1	4	8	7
	h=8 87~99	3	0.295	ngdp	earn	tosho	1	3	15		
GDP	h=2 87~99	2	0.555	ipisoav	lended		1	6			
	h=4 87~99	4	0.435	kijmi	mon2	rlended	nikav	11	1	8	3
	h=8 87~99	4	0.348	juckme	rmon2	lended	nikav	29	6	10	4

図表12 最適な組み合わせ

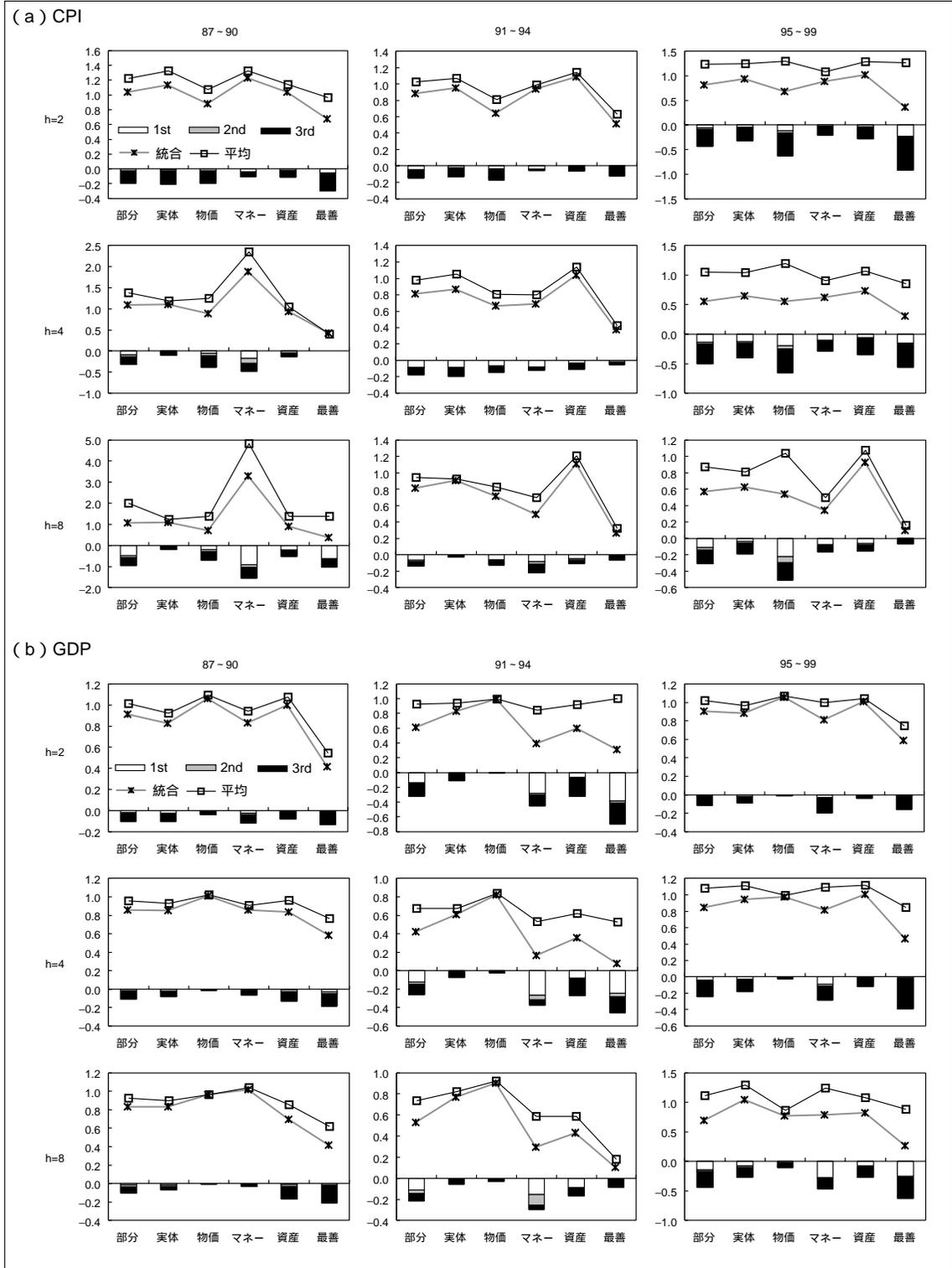


図表13 最善な統合のパフォーマンス (1987~99年)



また、時期別についても(図表14) (5)式の第3項の改善効果が高いのは通期と同様であるが、その大きさは時期によって異なっている。例えば、CPIの1991~94年をみると、相対MSFEは他より小さい一方、改善効果は他の統合方法に比べて小さくなっている。また、GDPの1991~94年では、第1項つまり予測バイアスの平準化がある程度機能していることが確認できる。

図表14 最善な統合とそのパフォーマンス（時期別）



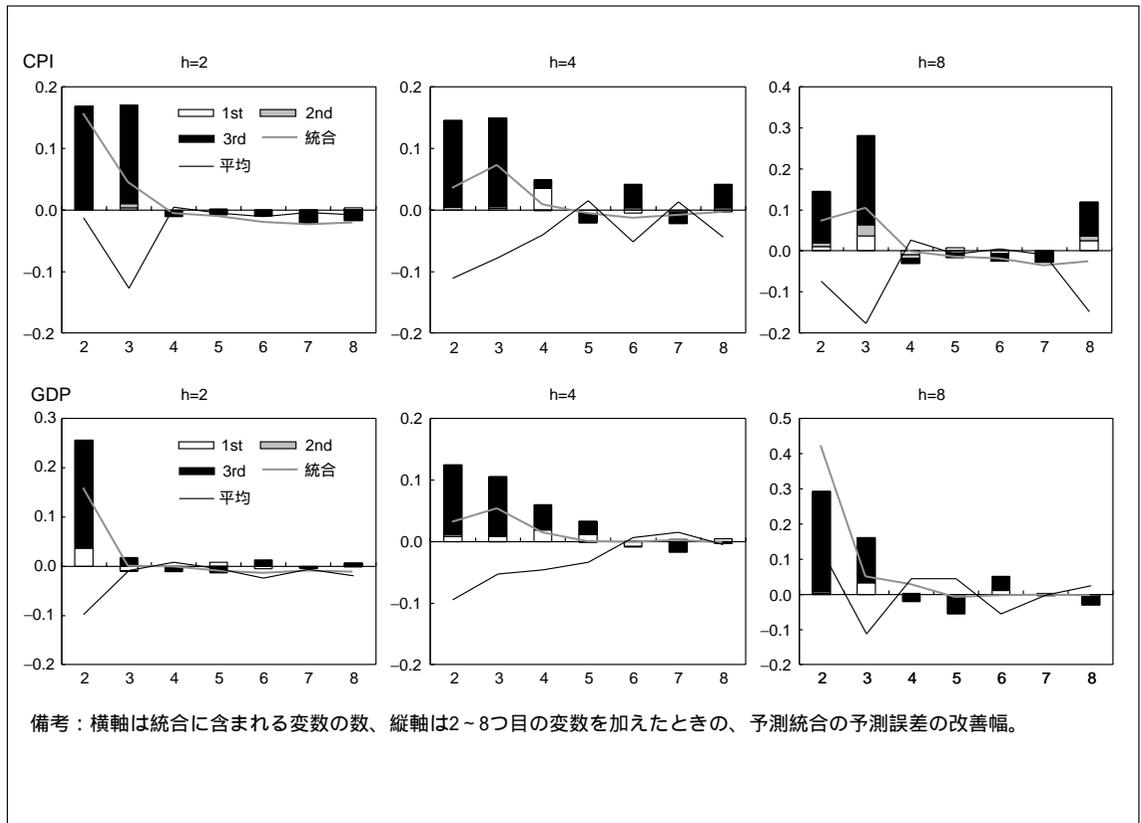
八．ステップワイズな変数追加とMSFE改善の要因

次に、なぜ統合対象系列数が5程度以上に増えたときに予測パフォーマンスが改善しなくなるのかを検討しておこう。

ステップワイズに統合対象系列を追加したときの相対MSFEの変化を要因分解してみると（図表15）、予測統合の改善幅（実線）が正になっている系列数2~4のときには、主として(5)式の第3項によって予測の改善がもたらされていることがわかる。しかし、こうした効果が期待できるのは、最適な組み合わせではせいぜい4変数が限度で、変数を5つ以上追加しても、こうした効果は期待できない⁴¹。

時期別についてみると、（図表16）、時期によって第1項による改善幅が大きくなるケースもみられるが、変化幅でみる限りでは、通期と同様第3項の改善幅の方が支配的なケースがほとんどである。実際のデータ分析結果から、すべての予測期数、時期についても、CPI、実質GDPともに上限となる統合系列数は2~4となっている。

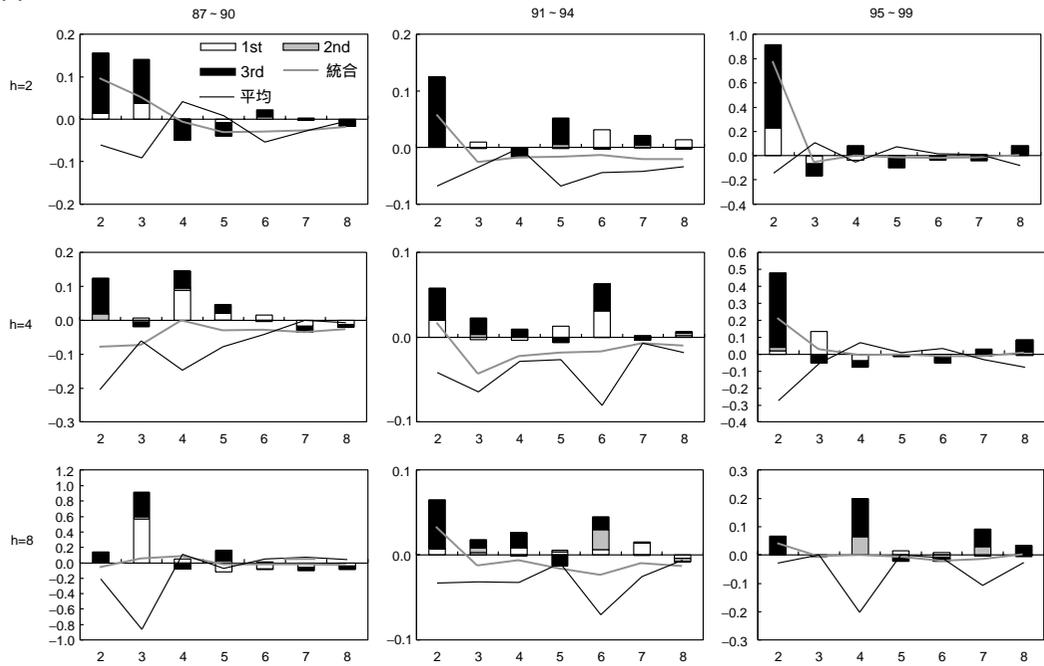
図表15 変数追加による改善とその要因（1987~99年）



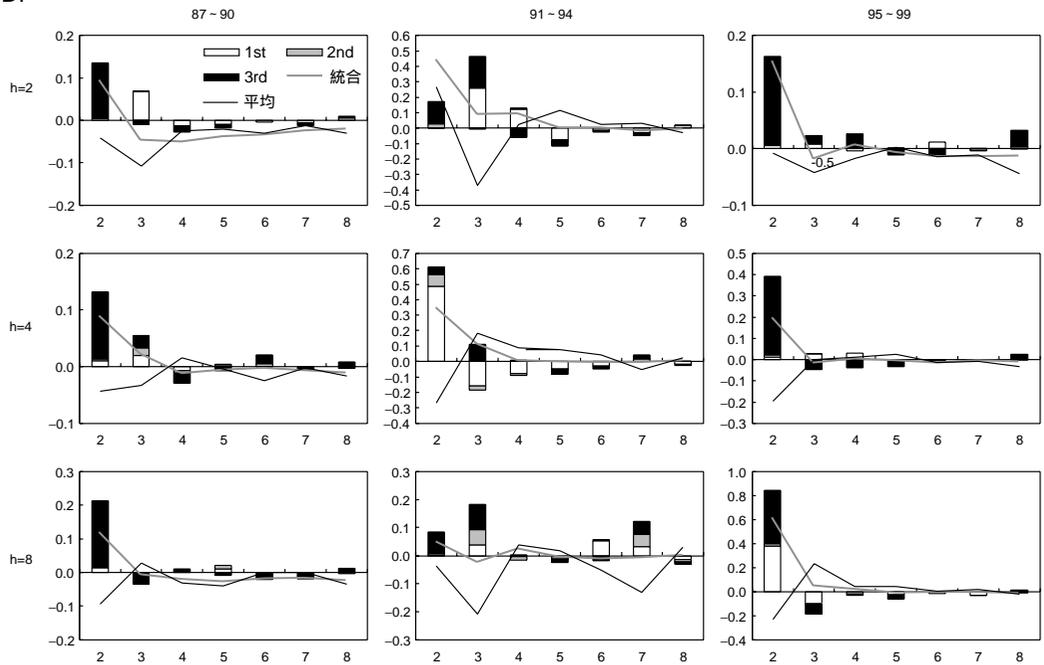
41 ここではステップワイズで最適となるような統合に対象を絞っている。初期変数の選択によっては、8~9変数の追加まで改善する場合もあるが、そのような場合は必ずステップワイズで最適な統合に劣る。

図表16 変数追加による改善とその要因（時期別）

(a) CPI



(b) GDP



備考：横軸は統合に含まれる変数の数、縦軸は2～8つ目の変数を加えたときの、予測統合の予測誤差の改善幅。

以上の結果を整理すると、現実には、予測誤差の相関が低い情報変数を多数みつけだすことは困難であるため⁴²、ステップワイズに変数を追加していくと、統合による改善効果が減衰する一方で、パフォーマンスの劣る原予測が追加され、相対MSFEの平均が上昇する効果が大きくなっていくと考えられる。このため、統合による予測パフォーマンス改善の効果は、高々4変数程度の統合で出尽くしてしまうことになる。

二．サンプル期間別の最適化との比較

最後に、最適な統合の組み合わせがサンプル期間を通じてどの程度頑健性をもっているのかを検討しておく。そのために、図表17として、通期でみて最適な組み合わせについて、各期間ごとの相対MSFEを算出し、各期間で最適な組み合わせのパフォーマンスとを比較した。なお、ここでは時期ごとに最適な組み合わせを、より長期の通期で近似することを考えているので、通期で最適な組み合わせを「次善」、各期で最適な組み合わせを「最善」としている⁴³。

まず、予測誤差の乖離をみると、乖離が大ききときはCPIで0.2ポイント、GDPで0.3ポイント程度となった（図表17上段）。このように予測誤差が拡大しているときは、最善と次善の間に共通する変数が少ない。例えば、1995～99年の期間におけるCPIの4半期先予測では、乖離が0.2程度生じているが、最善の組み合わせと次善の組み合わせで、共通していない変数の方が多い。また、同じ期間のGDPの8四半期先予測でも、共通している変数は1つだけである⁴⁴。

次に、最善と次善の統合による予測パフォーマンスの改善を、それぞれの要因分解に基づき比較してみると（図表17下段）、両者の差が生じている主要因は、改善効果よりは原予測の相対MSFEの平均である場合が多い。つまり、次善の予測統合では、よりパフォーマンスに劣る予測を統合しているが、改善効果はさほど変わらないため、相対MSFEはいずれも1.0を下回り、AR予測対比ではパフォーマンスを改善させている。

以上から、通期で最適な組み合わせと、各期で最適な組み合わせとのパフォーマンスの乖離幅は相対MSFEで0.2～0.3ポイント程度であること、その原因は、予測の統合による改善幅でなく2変数予測のパフォーマンスによって生じていることがわかった。とは言え、上でみたとおり、通期で最適な組み合わせの相対MSFEは0.2～0.6といずれの場合でも1を十分大きく下回っている。したがって、次善（通期で最適）な組み合わせは、AR予測の改善に役立つ程度には頑健であるように思われる。

42 ちなみに、予測誤差の主成分を計算すると、CPI、実質GDPともに、ほとんどの予測期数。時期で第2主成分までにより全変動の9割以上が説明できる。これは、統合対象系列数の上限が2～4程度であるとの結果とも整合的であろう。

43 サンプル期間を分割すれば、当然最適化における制約は緩和されるため、「最善」の組み合わせのパフォーマンスは「次善」のそれを必ず上回ることに注意されたい。

44 もちろん、変数をさほど共有していなくても、最善と次善との乖離が大きくないケース（例えば、1991～94年のGDP4半期先予測）もあり、変数が共通しているか否かは絶対的な基準ではない。

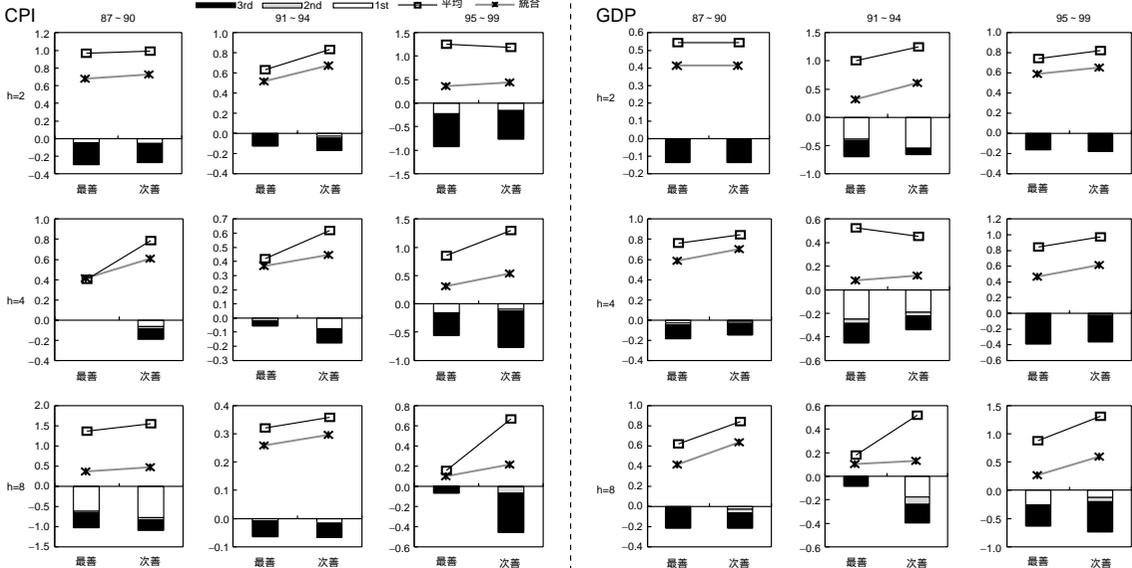
図表17 時期別の最適な組み合わせ

(a) 組み合わせ

h=2_87~90年		CPI				h=2_91~94年		CPI				h=2_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.667		ratest	ngdp	sprded		0.506	0.563	ngdp	earn		0.342	1.114	1.403	rearn	sprded					
(0.05)	0.714	0.814	0.934	1.146	(0.16)	0.663	0.563	0.670	1.257	(0.08)	0.421	1.037	1.114	1.403	rearn	sprded				
h=4_87~90年		CPI				h=4_91~94年		CPI				h=4_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.405		earn				0.363	0.379	ngdp	tosh		0.293	0.526	1.073	0.957	rlended	rearn	lcljded			
(0.19)	0.598	0.794	0.405	1.148	0.813	(0.08)	0.441	0.379	0.613	0.462	1.022	(0.23)	0.527	0.730	1.853	1.375	1.222	ratest		
h=8_87~90年		CPI				h=8_91~94年		CPI				h=8_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.347		earn	okhime	tosh	ngdp	0.257	0.289	0.355	earn	tosh	CPI	87-90	0.089	0.129	0.186	95-99				
(0.11)	0.457	1.044	0.408	3.209	(0.04)	0.293	0.432	0.289	0.355	(0.12)	0.206	0.186	1.116	0.693						
h=2_87~90年		GDP				h=2_91~94年		GDP				h=2_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.408		kijhi	lended			0.299	0.918	0.382	1.769	0.931	0.581	0.736	0.752	kijmi	rlended					
(0.0)	0.408	0.502	0.584		(0.30)	0.594	0.738	1.769		(0.06)	0.645	0.855	0.789	kijhi	lended					
h=4_87~90年		GDP				h=4_91~94年		GDP				h=4_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.576		juckme	nikav	lended		0.072	0.529	1.067	0.246	0.265	0.456	0.651	1.046	rlended	kijmi					
(0.12)	0.696	0.686	0.773	0.830	(0.04)	0.112	0.147	0.246	0.764	0.654	(0.15)	0.603	0.875	1.307	0.651	1.046	mon2	nikav	rlended	kijmi
h=8_87~90年		GDP				h=8_91~94年		GDP				h=8_95~99年								
最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測				最善	統合	含まれる変数とその個別予測						
0.406		juckme	nikav			0.096	0.143	0.216	rmon2	nikav		0.249	0.924	1.382	0.457	lcexpt	0.753			
(0.22)	0.627	0.524	0.713		(0.03)	0.122	0.803	0.216	0.902	0.143	(0.33)	0.578	0.706	1.286	1.882	1.382	lended	nikav	juckme	rmon2

備考：最善は各期で最適な組み合わせ、次善は通期（1987～99年）で最適な組み合わせ。数値は相対MSFE。「次善」下の（）内は、最善と次善の予測統合の差。シャドーは、最善と次善の組み合わせが共有していない変数。

(b) 改善幅の比較



5 . 結び

本稿では、幅広い金融・経済指標について、CPI、実質GDPの変動を予測するうえで、個別の指標による予測を統合したり各指標の共通成分を抽出することで、予測パフォーマンスを改善できるか、またその背後で何が起きているかを検討した。

分析結果を総括すると、情報変数の指標性は、時期や予測期数に依存して変化しているが、予測を統合したり変数の主成分を用いたりすることで、より安定した予測パフォーマンスが得られる。もっとも、統合による予測誤差の縮小効果は、変数の追加に伴って急速に減衰する一方、予測パフォーマンスの劣る系列が追加されるため、統合による予測パフォーマンスの改善効果は、高々2~4変数で出尽くしてしまう。このため、予測統合のパフォーマンスを向上させるうえでは、予測パフォーマンス自体が優れているだけでなく、他の情報変数を使った予測との相関係数の小ささをも兼ね備えた指標を探し出すことが重要と考えられる。

参考文献

- 大山慎介、「Forecast Combinationによる実質GDPの予測」、日本銀行調査統計局ワーキング・ペーパー、No.01-3、日本銀行調査統計局、2001年
- 翁 邦雄・白塚重典、「資産価格バブル、物価の安定と金融政策：日本の経験」、『金融研究』第21巻第1号、日本銀行金融研究所、2002年、71～115頁
- 粕谷宗久・真木和彦、「物価変動の転換点予測について」、日本銀行調査統計局ワーキング・ペーパー、No. 01-20、日本銀行調査統計局、2001年
- 加藤健吾、「金融政策遂行上の中間目標と情報変数について　米国における議論のサーベイとわが国への若干の応用」、『金融研究』第9巻第4号、1990年、51～91頁
- 北川源四郎・川崎能典、「時系列モデルによるインフレ率予測誤差の分析」、日本銀行調査統計局ワーキング・ペーパー、No. 01-13、日本銀行調査統計局、2001年
- 伴 金美・齊藤 誠、「マクロ計量モデルによるインフレ率予測誤差の分析」、日本銀行調査統計局ワーキング・ペーパー、No. 01-12、日本銀行調査統計局、2001年
- 肥後雅博・中田(黒田)祥子、「経済変数から基調的変動を抽出する時系列的手法について」、『金融研究』第17巻第6号、日本銀行金融研究所、1998年、39～98頁
- 福田慎一・慶田昌之、「インフレ予測に関する実証研究の展望　フィリップス曲線の日本における予測力を中心に」、日本銀行調査統計局ワーキング・ペーパー、No. 01-21、日本銀行調査統計局、2001年
- 本多祐三・松岡幹裕、「景気指標へのいくつかの統計学的接近　先行指標を中心として」、『フィナンシャル・レビュー』、財務省財務総合政策研究所、2001年
- 三尾仁志、「基調的なインフレ率とフィリップス曲線」、『金融研究』、第19巻第2号、日本銀行金融研究所、2000年、73～100頁
- Artis, Michael J, Anindya Banerjee and Massimiliano Marcellino, “Factor Forecasts for the UK,” CEPR Working Paper No. 3119, 2002.
- Bates, J.M. and C.W.J. Granger, “The Combination of Forecasts,” *Operations Research Quarterly*, Vol.20, 1969, pp. 319-25.
- Bernanke, Ben and Jean Boivin, “Monetary Policy in a Data-Rich Environment,” NBER Working Paper No. 8379, 2001.
- Cecchetti, Stephen G., Rita S. Chu and Charles Steindel, “The Unreliability of Inflation Indicators,” *Current Issues in Economics and Finance*, 6 (4), Federal Reserve Bank of New York, 2000.
- Clements, M.P. and David F. Hendry, *Forecasting Economic Time Series*, Cambridge University Press, 1996.
- Forni, Mario, Marc Hallin, Marco Lippi and Lucrezia Reichlin, “Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?” CEPR Working Paper No. 3146, 2002.
- Hirata, Hideaki and Kazuo Ueda, “The Yield Spread as a Predictor of Japanese Recessions,” Bank of Japan Research and Statistics Department Working Paper 98-3, 1998.
- Marcellino, Massimiliano, “Forecast Pooling for Short Time Series of Macroeconomic Variables,” CEPR Working Paper No. 3313, 2002.

- Marcellino, Massimiliano, James H. Stock and Mark W. Watson, "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information," mimeo, 2000.
- Stock, James H. and Mark W. Watson, "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business and Economic Statistics*, 14 (1), 1996, pp. 11-30.
- , and , "Diffusion Index," NBER Working Paper No. 6702, 1998.
- , and , "Forecasting Inflation," *Journal of Monetary Economics*, 44 (2), 1999, pp. 293-335.
- , and , "Forecasting Output and Inflation: The Role of Asset Prices," NBER Working Paper No. 8180, 2001.
- Watson, Mark W., "Macroeconomic Forecasting Using Many Predictor", mimeo, 2000.

