平均，分散，ベータ係数のベイズ修正効果

隅田　誠[a]・今井　英彦[b]

要　旨

株式について，個別銘柄のベータ値を予測する際にベイズ修正を行うことが有効であることは過去にも検証されてきた．ここでいうベイズ修正とは，ベータ値の母数の分布を適当に仮定し，その下で標本から得られた推定値を事前情報として，その条件付期待値を得るものである．このベイズ修正は，ベータ値に限らず平均や分散といった投資決定に関わる他のパラメータの予測にも適用できる．本稿では，日本市場に上場する銘柄の約32年間にわたる月次株価データを用いて，ベータ値をはじめ，平均，分散を計算し，その際にベイズ修正を行うことによる予測精度の向上を検証した．このとき，過去の検証で用いられてきたベイズ修正には，ベータ値の予測を前提としているからこそ妥当となる仮定が設けられており，とくに分散の予測においてそれは非現実的となる．本稿では，その仮定を緩めた場合の検証も行った．この検証により，ベータ値だけでなく，平均や分散の予測に関してもベイズ修正によって少なからず予測精度が向上するという実証結果が得られた．

JEL Classification Codes：C11, G11，G17
キーワード：ベイズ修正，Vasicek，MSE，平均・分散モデル，マーケットモデル

1.　はじめに

現代ポートフォリオ理論を実際の投資決定に適用する際，リスクやリターンに関するパラメータの予測が欠かせない．しかし，予測に利用できるデータは当然ながら投資決定以前のものに限られ，そこから推定されるパラメータは過去の値となるので，将来，すなわち投資の運用期間中に実現する実際の値との間には多少なりとも誤差が生じる．それによって投資収益の発生が当初の想定と大きく乖離することになるため，投資が計画通りに運用されるには，必要なパラメータの正確な予測が必要不可欠である．

それらのパラメータは，投資対象ごとに個別に予測することが一般的であるが，投資対象を株式に限れば，各パラメータの母数（期待値）の分布がある程度想像できる．たとえば，株価の月次変化率の平均について考えれば，その値が0から大きく離れる銘柄は非常にまれであろう．実際，過去120か月程度の期間で見た場合，最近のデータであれば –0.01から0.02の間にほとんどの銘柄が含まれ，その分布は単峰型であるといえる[1]．この実証事実がある下で仮に株価変化率の平均に –0.1などというサンプルが得られたならば，その予測精度は疑わしい．もちろん，株式以外の投資対象であっても同じことがいえるが，株式投資はパラメータの母数の分布を推定しやすい．

本稿では，この分布を仮定してパラメータを推定するという，いわゆるベイズ推定の予測能力を，一般的な推定方法と比較して検証する．予測対象とするパラメータは，株価変化率の（期待収益率の推定量としての）平均，分散，およびベータ値である．

パラメータの母数の分布を考慮せずに推定した値を修正するという意味で，この推定方法はベイズ修正，あるいはベイジアン修正とも呼ばれる．ベイズ修正の詳細については後述するが，簡潔にいえばパラメータの母数の分布を適当に仮定し，その下である推定値が得られた場合の条件付期待値として推定値を修正するものである．ゆえに，パラメータの母数の分布と，その推定値が従う分布とをいかに設定するかが重要となり，その設定によって修正値の推定量や値が異なる．本稿では，それら

a　武蔵大学大学院経済学研究科博士前期課程　　〒176–8534　東京都練馬区豊玉上1–26–1
b　武蔵大学経済学部　　〒176–8534　東京都練馬区豊玉上1–26–1

1　2005年10月から2015年9月までの120か月間において，月次株価変化率の平均が –0.01から0.02の間にない銘柄は，後述する対象銘柄のうちのこの期間にデータの欠損のない全1851銘柄の中でわずか51銘柄であり，3% にも満たないことが確認できる．

の設定が異なる修正モデルをいくつか用いて，ベイズ修正自体が有効であるか否かとともに，どのような設定が適しているかについても検証する．

また，予測精度は，修正値や将来のパラメータの値（以下，事後の値と呼ぶ）を得るために用いる標本数とサンプルによって変化することが予想できる[2]．具体的にはそれらの標本数が大きいほど，予測精度が高まることが想像できる[3]．そこで本稿では，標本数を変えた場合の予測精度の変化も検証する．

2. 先行研究

ベイズ修正に関する類似した研究は，とくにベータ値，すなわち市場ポートフォリオの収益率に対する個別銘柄の収益率の感応度の予測を目的としたものが過去にも多く行われている．以下ではその例をいくつか挙げる．

まず，ベータ値の予測に対するベイズ修正の効果に関する研究は，Vasicek（1973）に端を発するといえる．Vasicek（1973）はベータ値の母数と推定値がともに正規分布に従うことを仮定した場合のベイズ修正の推定量を非常に単純な関数として示しており，これ以降の（ベータ値の予測精度に関する）実証研究の多くで，この Vasicek（1973）の修正方法（以下，Vasicek モデル）が一つの手法として比較対象に採用されている．

Klemkosky and Martin（1975）もその一例であり，この Vasicek モデルをほかのいくつかのモデルとともに比較している．ここで Vasicek モデルとともに検証されたモデルは，推定値を異なる2時点で得，新しい時点の推定値を古い時点の推定値で回帰した線形回帰モデルに新しい時点の推定値を当てはめて，次時点，すなわち将来の値の予測値を得ようとするものであり，いわゆる Blume モデルと呼ばれるモデルなどである（Blume（1971））．Klemkosky and Martin（1975）では個別銘柄のベータ値とともに，10 銘柄から成るポートフォリオのベータ値の予測も行っているが，いずれの場合でも Vasicek モデルはもちろん，他の修正モデルにおいてもその効果があることが示されている．また，3 つの異なる時点で検証しており，そのうち 2 つの時点では，個別の予測でもポートフォリオの予測でも Vasicek モデルの予測精度が最も高くなる結果であった．

Elton and Gruber（1978）は，ベータ値の予測に Vasicek モデルや Blume モデルを用い，その修正されたベータ値を使って銘柄間の相関係数の予測を行っている[4]．これらの 2 モデルの他にも，すべてのベータ値の推定値を平均化する Overall Mean モデルと呼ばれるモデル（以下で Mean モデルとして定義するモデルと同等のモデルであり，これと同等のモデルを以下，Mean モデルと呼ぶ）や，すべてのベータ値の予測値を 1 とするモデルも検証対象としている．検証された 2 つの時点の両方において，Mean モデルが最も予測精度が高い結果であったが，後述するように Mean モデルもベイズ修正の特殊な例として捉えることができる．Vasicek モデルも，修正しないベータ値を用いた場合に比べれば修正の効果が認められたが，Mean モデルほどではなかった．いずれにせよ，ベイズ修正によって予測したベータ値を相関係数の予測に用いても，依然としてその効果があることが確認されている．

Eubank and Zumwalt（1979）は Klemkosky and Martin（1975）と同じように，ベータ値の予測に対する Vasicek モデルと Blume モデルの効果を検証している．彼らは基本的にポートフォリオのベータ値を予測対象としており，その構成資産数が異なる場合や，（推定値や事後の値を得るために用いる）標本数が異なる場合の検証も行っている．結果として，標本数や資産数が多いほど予測精度が高まり，また，いずれの場合でも各修正モデルに予測精度を向上させる効果があることが示されている．修正モデルを比較すると，推定値を得るために用いる標本数が少ない場合は Blume モデルのほうが優れ，多い場合は僅差で Vasicek モデルのほうが優れる結果であった．

Eun and Resnick（1984）は Elton and Gruber（1978）と同じように，ベイズ修正を用いて予測したベータ値

2　本稿では，単に推定値や観測値といった場合，基本的に修正前のパラメータの値を指すものとし，何らかの修正モデルによって修正した後のパラメータの値を指す場合はそれらを修正値と呼び，区別することとする．これらとは別に，予測値という呼称も用いるが，これは修正の有無を問わず事後の値を予測するための値を総称するものとする．たとえば，ベータ値の予測に際して，推定値を全く計算に含まずに将来の値を 1 と予測するならば，それは予測値であっても，修正値とはいい難い．しかし，修正値は事後の値を予測するために修正を施しているものであるから，予測値であるともいえる．

3　パラメータの推定の基となる標本，すなわち株価変化率が独立に同一の分布に従っており，なおかつパラメータの推定量が一致推定量ならば，データ数が大きくなるほど異なる標本から得られる推定値の差は小さくなりやすくなる（差の絶対値の期待値が 0 に近づく）．

4　ファクターモデルに基づく場合，リスク要因 k に対する銘柄 i の収益率の感応度を $\beta_{i,k}$，リスク要因 k の分散を σ_k^2，リスク要因の数を F とし，更にリスク要因間および誤差項間で相関がないと仮定すれば，銘柄 i と銘柄 j の共分散は $\sum_{k=1}^{F} \beta_{i,k}\beta_{j,k}\sigma_k^2$ で表され，相関係数はこれを銘柄 i と銘柄 j の収益率の標準偏差で除した値である．

を使って相関係数を予測しているが，Elton and Gruber（1978）と異なり，相関係数を予測する際，各銘柄の国と業種という特性を考慮し，その特性によって銘柄をグループ化し，そのグループごとに異なる方法で予測を行うことも試みている．具体的には，各銘柄に対応する国あるいは業種の株価指数を説明変数に加えたマルチファクターモデル，および国あるいは業種ごとに推定値を平均化したモデルを検証対象に加えている．なお，ベイズ修正（Vasicek モデル）は，Elton and Gruber（1978）と同様に，シングルファクターモデルにおけるベータ値に対してのみ適用しており，マルチファクターモデルにおいては用いていない．もちろん，マルチファクターモデルにおける各感応度にベイズ修正を適用すること自体は可能である．

結果としては，国ごとに平均化したモデルが最も高い予測精度を示している．これは Elton and Gruber（1978）の研究において Mean モデルが最も優れていたことと似ている．また，Vasicek モデルによるベータ値を用いる場合と修正を行わないベータ値を用いる場合とを比較すれば，他の研究と同様に修正の効果が表れているものの，ベータ値を用いずに（ファクターモデルによらずに）相関係数を推定する場合と比べるとそれに劣る結果であった．

Karolyi（1992）も Eun and Resnick（1984）と似て，各銘柄の業種と規模（時価総額）という情報を活用してベイズ修正を行っている．具体的には，業種または時価総額，あるいはその両方で銘柄をグループ化し，そのグループごとに Vasicek モデルに基づく修正を行っている[5]．また，月次データだけではなく，週次データや日次データでも検証しているが，そのいずれであっても，業種あるいは時価総額のどちらかの情報だけによってグループ化する場合は予測精度があまり改善されず，その両方を用いてグループ化する場合に予測精度が向上するという結果になっている．もちろん，どちらの情報も使わない，すなわちグループ化しない普通の Vasicek モデル自体，予測精度の向上に大きく貢献することが示されている．

日本市場を対象とした同様の研究もいくつか行われており，たとえば，金崎（1987）は Vasicek モデルと Mean モデルとともに，予測誤差の 2 乗和の期待値を最小化するという James and Stein（1961）が示したモデル（以下，James-Stein モデル）などを用いて予測精度を比較している．その結果，Vasicek モデルと James-Stein モデルは，複数の時点で行われた検証のすべてで修正の効果が認められた．また，個別銘柄のベータ値だけでなく，ポートフォリオのベータ値に関しても検証しており，その場合でも修正が有効であって，さらに個別銘柄のベータ値の予測に比べて改善の効果が大きいことが示されている．Vasicek モデルと James-Stein モデルとを比較すると，前者のほうが優れていると結論づけている．なお，金崎（1987）は James-Stein モデルの推定量が，Vasicek モデルにある仮定を加えたものに等しいことについても言及しているが，修正方法の基本的な考え方の点においてベイズ修正とは異なるため，本稿では検証しない[6]．

吉原（1990）も金崎（1987）と同じように，Vasicek モデル，James-Stein モデルを検証対象としているが，これらとともに Vasicek モデルに用いるパラメータにある調整を加えたモデルも検証している．その調整とは本来 Vasicek モデルでは推定値全体の平均として推定するパラメータを時価総額による加重平均で置き換え，推定値全体の分散として推定するパラメータを平均偏差平方の時価総額による加重平均で置き換えるものである．つまり，Vasicek モデルではすべての銘柄を等しく扱っていたものを，この調整によって各銘柄の重要度が時価総額の相対的な大きさに一致するように調整していることになる．このモデルは金崎（1987）がその論文の最後に今後の課題として提案していたモデルに等しい．結果としては，やはりすべてのモデルにおいてその修正が効果的であることが示されているが，この Vasicek モデルの調整が Vasicek モデルの予測精度を必ずしも改善する結果にはなっておらず，むしろ複数の時点で検証したうちの多くの場合で Vasicek モデルのほうが効果的な結果となっている．

芳野（2011）も金崎（1987）や吉原（1990）と同様に，日本市場を対象として Vasicek モデルや James-Stein モデルの予測精度を検証している．芳野（2011）では 2011 年 9 月までの 261 か月間という比較的新しく，かつ長期間の月次データを用いて検証している．ゆえに，他の多くの先行研究のように，ある 1 時点ないし少数の時点における優劣によって逐一判断するのではなく，多数の時

5　Karolyi（1992）が Multiple Shrinkage と呼ぶモデルは，業種と時価総額の両方の情報を利用して修正値を求めているが，これは後述する sector モデルのように業種と時価総額水準がともに予測対象の銘柄と一致する銘柄だけを集め，その銘柄群だけを対象として Vasicek モデルを適用しているわけではなく，それよりもやや技巧的な方法を用いている．詳細は Karolyi（1992）を参照されたい．

6　ベイズ修正は母数の分布を仮定し，その下で特定の推定値が得られたときの条件付期待値として修正するものであるが，James-Stein モデルは誤差の 2 乗和の期待値を最小化する値として修正するものである．

点で検証した結果を，その平均等によって総合的に判断することとなっている．この点は本稿における検証方法と似ているといえる．また，個別銘柄のベータ値の予測と，ポートフォリオのベータ値の予測の両方を行っているが，そのいずれにおいてもやはりベイズ修正の効果が認められている．Vasicek モデルと James-Stein モデルとを比較すると，比較的 Vasicek モデルのほうが優れる結果となっており，最近のデータを用いても金崎（1987）や吉原（1990）の研究と整合的な結果となることが示されている．

ベイズ修正は行っていないが，日本市場におけるベータ値の予測精度に関する研究として安藤・久保田（2003）がある．安藤・久保田（2003）では推定値や事後の値を得るために用いる標本数に複数のパターンを設定し，その組み合わせによる予測精度の変化について検証している．やはり予測精度は標本数が多くなるほど小さくなる傾向にあるものの，標本数が多くなるに従って予測精度向上の効果も小さくなり，個別銘柄のベータ値を予測する場合は，その数が 36 程度，すなわち月次データで 36 か月分（安藤・久保田（2003）では月次データを用いている）のデータを用いれば充分であるという結論に至っている．

3. 検証手法

3.1 標本の設定

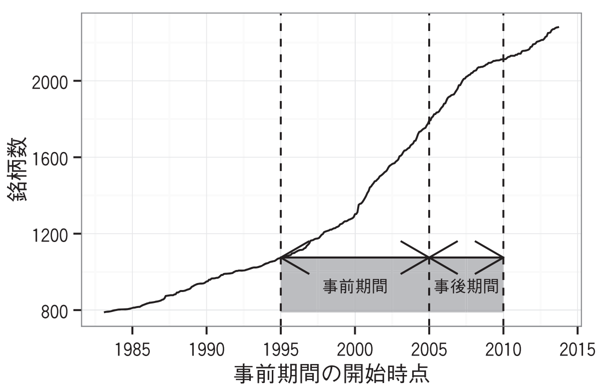
本稿では統計量の予測に使うデータを得る期間を事前期間，事後の値を計算するために使うデータを得る期間を事後期間と呼ぶ．そして，事前期間の長さを N_a（ex-ante），事後期間の長さを N_p（ex-post）とし，それらに次の値を与え，それぞれを組み合わせたすべての場合で検証を行う．本稿では月次データを扱うため，N_a および N_p の単位は月である．

$$N_a \in \{12, 36, 60, 120, 180\}$$
$$N_p \in \{12, 36, 60\}$$

なお，これらの値のすべての組み合わせに関して検証するため，N_a が N_p と比較して極端に小さい場合，つまりたとえば $\{N_a, N_p\} = \{12, 60\}$ のような場合も検証することになるが，このような予測は事前のデータが不足しており現実的であるとはいい難い．

また，事後期間の開始時点を基準時点と呼ぶ．たとえば，基準時点が 2005 年 1 月で $N_a = 120$，$N_p = 60$ のとき，事前期間は 1995 年 1 月から 2004 年 12 月までの 120 か月間，事後期間は 2005 年 1 月から 2009 年 12 月までの 60 か月間である[7]．なお，この場合の事前期間と事後期間の設定のイメージを図 1 に示した．本稿では使用したサンプルについて計算可能なすべての基準時点で検証を行う[8]．

図 1　対象銘柄数の推移



本稿で用いる株価のデータは，1983 年 1 月から 2015 年 9 月までの 393 か月間の月次株価をである[9]．ただし，実際には変化率に関して検証するので[10]，利用できる時系列でみたデータ数は最大でも $393 - 1 = 392$ となる．ゆえに，N_a と N_p の組み合わせごとに計算の対象となる基準時点の数は表 1 の通りであり，その値は $393 - N_a - N_p$ に等しい．

7　より正確にはこの 120 か月間（あるいは 60 か月間）に含まれる 60（あるいは 120）の時点の株価変化率を用いるのであって，実際には事前期間で 1994 年 12 月の，事後期間で 2004 年 12 月の株価データも必要となる．

8　事前期間と事後期間で充分な標本を得るには，基準時点から $N_a + 1$ か月前から，基準時点から $N_p - 1$ か月後までの株価が存在する必要がある．

9　月次株価のデータは Yahoo!ファイナンスから入手し，調整後終値の値を用いる．この調整とは株式分割による株価変動の調整であって，配当落ちを調整するものではない．また，2015 年 9 月以前に上場廃止となった銘柄は計算対象としないため，検証結果には多少なりとも生存バイアスがかかることが懸念されるが，少なくともモデル間の予測能力の優劣に関して論じる限り，必ずしも上場廃止となる銘柄とそうでない銘柄との間に大きな差があるとはいえないだろう．

10　銘柄 i の t 時点の変化率を $R_{t,i}$ とし，次式から得る．

$$R_{t,i} = \frac{V_{t,i}}{V_{t-1,i}} - 1$$

ここで，$V_{t,i}$ は銘柄 i の t 時点の株価である．また，時点 t は 1983 年 1 月からの経過月数とする．

表 1　N_a と N_p の組み合わせごとの基準時点の数

$N_p \backslash N_a$	12	36	60	120	180
12	369	345	321	261	201
36	345	321	297	237	177
60	321	297	273	213	153

また，検証の対象とする銘柄は 2015 年 9 月時点で日本の株式市場に上場し，かつ株価データが取得できる銘柄，計 3598 銘柄のうち，次の 3 つの条件のいずれかを満たす銘柄に限る（括弧内の数字はその条件に合致する銘柄数を示す）.

- 2015 年 9 月時点で時価総額が 100 億円以上である（2101）.
- 2014 年 10 月から 2015 年 9 月までの 12 か月間の単元当たり出来高が 50000 以上である（2250）.
- 東証 33 業種分類において同業種にあたる銘柄が，2015 年 9 月時点で全銘柄（株価データが得られた 3598 銘柄）の中に 30 銘柄以下しか存在しない（132）.

これらの条件は市場で価格のつき難い銘柄を除外するために設けたものである．この 3 つの条件のうちの少なくとも 1 つ以上を満たす銘柄は 2724 銘柄存在する．ただし，上の 3 つの条件を満たしていても，次の 4 つの条件のいずれかを満たす場合は検証対象から除外する（括弧内の数字は / の右側がその条件に合致する銘柄数，左側がそのうち上の 3 つの条件の少なくとも 1 つ以上を満たす銘柄数を示す）.

- 1983 年 1 月（それ以降に上場した銘柄ならばその上場時点）から 2015 年 9 月までの間に株価データが得られない月が 1 度でもある（98/197）.
- 1983 年 1 月から 2015 年 9 月までの間のいずれかの連続する 12 か月間において，株価変化率に変動がないことが 1 度でもある（6/9）.
- 1983 年 1 月から 2015 年 9 月までの間に株価の対数変化率の絶対値が 1 を超える月が 1 度でもある[11]（196/234）.
- 2013 年 9 月の時点で上場していない（156/208）.

これらの条件のうち，上の 2 つは修正値が正常に計算できることを保証するためであり，3 つ目の条件は異常な推定値や修正値，あるいは事後の値の発生を防ぐため[12]，最後の条件は事前期間と事後期間の設定によって，これに該当する銘柄はいずれにせよ計算の対象とならないためである．この 4 つの条件のうち少なくとも 1 つ以上を満たす上場銘柄は 630 銘柄存在し，これによって除外される銘柄数は 496 である．この結果，本稿で検証対象となる銘柄数は $2724 - 496 = 2282$ となる．この銘柄群に関して，業種や市場ごとの内訳を表 2 に示す.

事前期間と事後期間の取り方によって，そのとき計算

表 2　業種，市場ごとの対象銘柄総数

業種 \ 市場	東証 1 部	東証 2 部	マザーズ	東証 JQS	東証 JQG	他	計
水産・農林業	4	1	0	3	0	0	8
鉱業	6	0	0	0	0	0	6
建設業	85	10	0	7	0	3	105
食料品	67	13	0	6	0	1	87
繊維製品	31	6	0	0	0	0	37
パルプ・紙	11	6	0	6	0	1	24
化学	118	15	0	12	2	3	150
医薬品	32	0	6	2	2	0	42
石油・石炭製品	10	2	0	0	0	0	12
ゴム製品	11	5	0	1	0	0	17
ガラス・土石製品	27	6	0	1	0	0	34
鉄鋼	27	3	0	1	0	1	32
非鉄金属	21	3	0	1	0	1	26
金属製品	33	9	0	5	0	0	47
機械	109	16	1	19	0	0	145
電気機器	147	23	2	29	0	2	203
輸送用機器	62	8	0	3	0	0	73
精密機器	25	2	0	5	3	0	35
その他製品	44	8	0	10	1	1	64
電気・ガス業	17	2	0	0	0	1	20
陸運業	35	3	0	0	0	1	39
海運業	7	4	0	0	0	0	11
空運業	3	1	0	0	0	0	4
倉庫・運輸関連	20	1	0	1	0	2	24
情報・通信業	122	15	23	43	4	1	208
卸売業	135	22	3	30	1	4	195
小売業	155	19	7	40	0	6	227
銀行業	75	0	0	1	0	0	76
証券，商品先物取引業	22	2	0	7	0	0	31
保険業	7	0	1	2	0	1	11
その他金融業	22	0	1	2	1	0	26
不動産業	42	7	2	10	1	0	62
サービス業	125	18	14	39	2	3	201
計	1657	230	60	286	17	32	2282

11　銘柄 i の t 時点の対数変化率とは，$R_{t,i}$ を株価の変化率としたとき，$\ln(R_{t,i}+1)$ の値である.

12　株価変化率に変動がない場合，後述する一部のモデルの修正値の計算に必要なパラメータである標準誤差の推定値が 0 となり，計算過程に 0 除算が生じることになる．また，標本 R_i に巨大な値があるとき，本稿で予測するパラメータのうちとくに分散が異常な値となり易く，その場合の予測精度がその銘柄 i にあまりにも極度に依存することになる．そのような銘柄における予測精度も興味深いが，そうでない銘柄を含めて総合的に判断する場合，その銘柄による大きな予測誤差が（大多数の銘柄に対していえるような）本来あるべき判断を狂わせる危険がある.

に使われる銘柄数が異なる．これは上記 2282 銘柄のうちに 1983 年 1 月以降に上場した銘柄も含まれるためである．なお，1983 年 1 月時点ですでに上場していた銘柄は上記 2282 銘柄中 788 銘柄だけである．この結果，対象となる銘柄数の推移を図 1 に示す．

3.2　予測精度の評価

次に述べるような予測精度の評価方法は，予測対象となる統計量を問わず，複数のモデルによる複数の確率変数の予測精度の比較に対して適用できる．よって，以下では N_m 個のモデルによる N_n 個の確率変数の予測に関して[13]，モデル j による i 番目の確率変数の予測値を ${}^j\hat{\theta}_i$，事後の値を $\dot{\theta}_i$，誤差を ${}^jD_i = {}^j\hat{\theta}_i - \dot{\theta}_i$ として評価方法を記述する．

予測精度の評価基準としては，多くの先行研究において MSE（Mean Squared Error）が用いられている（Klemkosky and Martin（1975），Eubank and Zumwalt（1979），Eun and Resnick（1984），金崎（1987），吉原（1990），Karolyi（1992）など）．一方で，Elton and Gruber（1973）や安藤・久保田（2003）では MAE（Mean Absolute Error）を用いている．しかし，Eun and Resnick（1984, pp.1322）は MAE を用いても検証結果に大きな差はないと述べており，実際，後に述べるように結果が全く異なることは考え難い．ゆえに，本稿では基本的に MSE だけについて検証することとする．

まず，MSE は誤差の 2 乗の平均であり，モデル j の MSE は次式のように求める．

$$MSE = \frac{1}{N_n}\sum_{i=0}^{N_n} {}^jD_i^2 = \overline{{}^jD^2} \tag{1}$$

もちろん，この値が小さいほど予測精度が高いといえる．また，この正の平方根は RMSE（Root Mean Squared Error），あるいは RMS（Root Mean Square）と呼ばれる（本稿では RMSE と呼ぶ）．${}^j\hat{\theta}$ の標本標準偏差を $s_{{}^j\hat{\theta}}$，$\dot{\theta}$ の標本標準偏差を $s_{\dot{\theta}}$，${}^j\hat{\theta}$ と $\dot{\theta}$ の標本共分散を $s_{{}^j\hat{\theta},\dot{\theta}}$ とすると，MSE は次のように分解できる．

$$MSE = \left(\overline{{}^j\hat{\theta}} - \overline{\dot{\theta}}\right)^2 + \left(s_{{}^j\hat{\theta}} - s_{\dot{\theta}}\right)^2 + 2\left(s_{{}^j\hat{\theta}}s_{\dot{\theta}} - s_{{}^j\hat{\theta},\dot{\theta}}\right) \tag{2}$$

本稿では，この右辺第 1 項を「平均の誤差」，右辺第 2 項を「標準偏差の誤差」，右辺第 3 項を「相関の誤差」と呼び，区別することにする．つまり，MSE は，${}^j\hat{\theta}$ と $\dot{\theta}$ の平均の差，${}^j\hat{\theta}$ と $\dot{\theta}$ のばらつきの違い，${}^j\hat{\theta}$ と $\dot{\theta}$ との相関の度合いといった 3 つの観点を総合的に表す指標であるといえる．また，${}^j\hat{\theta}$ 全体の平均を調整することは平均の誤差にだけ影響を与えることになる．Elton and Gruber（1978）や Eun and Resnick（1984）は相関係数の予測に際して，予測値全体の平均を推定値全体の平均に一致するように調整することで予測精度が変化するか否かを検証している．これは，推定値と事後の値が独立で同一分布に従うならば，推定値はたとえ予測能力が低いとしても，全体的な平均の水準を予測する能力は比較的優れると考えられるためである．

また，MSE は次のように分解することもできる．

$$MSE = \left(\overline{{}^j\hat{\theta}} - \overline{\dot{\theta}}\right)^2 + \left(1 - \frac{s_{{}^j\hat{\theta},\dot{\theta}}}{s^2_{{}^j\hat{\theta}}}\right)^2 s^2_{{}^j\hat{\theta}} + \left(1 - \frac{s^2_{{}^j\hat{\theta},\dot{\theta}}}{s^2_{{}^j\hat{\theta}}s^2_{\dot{\theta}}}\right)s^2_{\dot{\theta}} \tag{3}$$

ここで，$\dfrac{s_{{}^j\hat{\theta},\dot{\theta}}}{s^2_{{}^j\hat{\theta}}}$ は $\dot{\theta}$ の ${}^j\hat{\theta}$ に対する回帰係数，$\dfrac{s^2_{{}^j\hat{\theta},\dot{\theta}}}{s^2_{{}^j\hat{\theta}}s^2_{\dot{\theta}}}$ は $\dot{\theta}$ の ${}^j\hat{\theta}$ による線形回帰モデルの決定係数となる．この右辺の各項は Klemkosky and Martin（1975）や Eubank and Zumwalt（1979）などで，それぞれ bias, inefficiency, random error と呼ばれている．ただし，この分解した各成分については本稿では基本的に検証しない．

一方で，MAE は誤差の絶対値の平均であり，モデル j の MAE は次式のように求める．

$$MAE = \frac{1}{N_n}\sum_{i=0}^{N_n} |{}^jD_i| = \overline{|{}^jD|} \tag{4}$$

MAE も値が小さいほど予測精度が高いといえる．なお，$MSE \geq MAE^2$ であり，MSE から MAE の 2 乗を引いた値は誤差 ${}^jD_i$ の絶対値の標本分散に等しい．

誤差の大きさを比較するには MSE や MAE が有用であり，あるモデルの MSE や MAE が他のあるモデルのそれに対して小さければ，そのモデルのほうが相対的に予測精度が高いと判断できる．しかし，適当な 2 モデル間で MSE の大小関係と MAE の大小関係が逆転する場合もないではない．MSE は MAE よりも絶対値が大きな誤差の影響を受けやすいので，たとえば，あるモデルが他のあるモデルに対して MSE は小さいが MAE は大きかった場合，その原因としては誤差のばらつきが小さかったものの，その平均的な水準がその効果を相殺しない程度に 0 から乖離していた，つまり予測値が全体的に大きすぎた，あるいは小さすぎた可能性があると考えられる．

実際，${}^jD \sim \mathrm{N}(\mu, \sigma^2)$ であると仮定すれば，

$$\mathrm{E}(MSE) = \mu^2 + \sigma^2$$

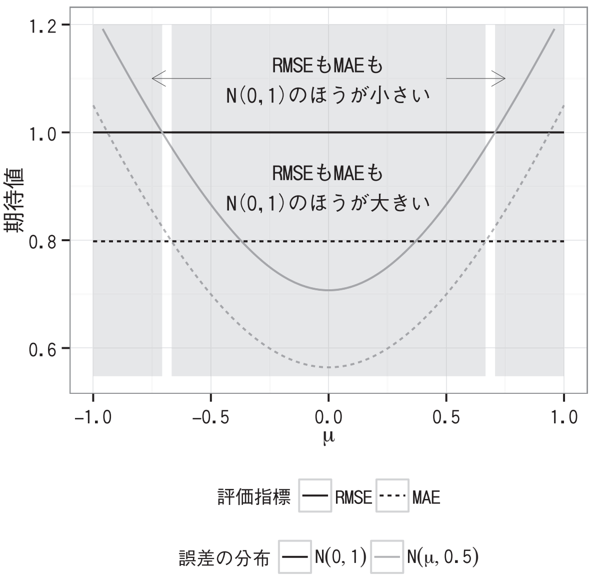
$$\mathrm{E}(MAE) = \mathrm{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right)\mu + \exp\left(-\frac{\mu^2}{2\sigma^2}\right)\sqrt{\frac{2}{\pi}}\sigma$$

13　N_m は比較するモデル数に，N_n は対象銘柄数に対応する．

であるから[14]，$\sigma^2 = 0.5$ であるとき，RMSE と MAE の期待値は図 2 のように推移する．この場合，誤差が標準正規分布に従うと仮定されるモデル（図 2 の水平線）に対して，MSE ならば $\mu = \pm\sqrt{0.5}$ のときに，MAE ならば $\mu \simeq \pm 2/3$ のときに期待値が一致することがわかる[15]．これはおよそ $-\sqrt{0.5}$ から $-2/3$ の区間および $2/3$ から $\sqrt{0.5}$ の区間では MSE と MAE の期待値の大小関係が反対になることを意味する（誤差が標準正規分布に従うモデルと比べて，MSE は小さいが MAE は大きくなる）．しかし，この範囲は非常に狭く，たいていの場合は期待値でみれば MSE と MAE の大小関係は一致する．

多くの先行研究（とくに Klemkosky and Martin（1975）や Elton and Gruber（1978），Eun and Resnick（1984）などのように時代の古いもの）では，基準時点を 1 つないしは少数に限って検証を行っているため，MSE そのものを 1 時点ごとに比較してモデルの優劣を順序付けることが容易であるが，本稿では表 1 に示したように（各 N_a と N_p の組み合わせにおいて）多数の基準時点で検証を行うため，（N_a や N_p が一定であっても）得られる MSE の数が膨大となり，そのすべてにおいてモデルの順位が一貫する保証はなく総合的な判断が難しい．MSE の中央値や平均でもってモデル間の優劣を決定することもできるが，本稿ではそれとともにカイ 2 乗検定による対比較を行う．これは MSE の母代表値に関して，任意の 2 モデル間で有意な差があるか否かを検定するものである[16]．ただし，この検定はあくまで母代表値に差があるか否かを検定するものであるため，検定結果と MSE の中央値だけをみて，いずれのモデルの母代表値が小さいかを判断することは早計である．ゆえに，本稿では検定結果とともに MSE の十分位数の大小も比較する．なお，有意水準は 1% とする．

図 2　RMSE と MAE の期待値の変化のイメージ



前述のとおり，本稿では基本的に MSE に基づいてモデルの優劣を比較するが，誤差の水準やモデルの優劣が業種によって異なる可能性が考えられる．そこで，業種ごとに MSE を計算した結果についても一部の業種と特定の N_a および N_p の組み合わせに限り検証する．

3.3　修正モデル

以下に述べる修正モデルは，あくまで一般的な推定量を用いて任意の確率変数（としての統計量）の推定値に修正を加えるものであって，その元となる推定値を得る手順（すなわち推定量）そのものを変えるものではないため，複数の推定値が観測できるならば予測対象となる確率変数を問わず適用可能である．ゆえに，以下では予測対象とする統計量を明示せずに述べる．

14　$\mathrm{erf}(x)$ は誤差関数であり，次式のように定義される．

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt$$

15　より正確には MAE の期待値は μ が

$$\int_0^\mu \exp(-t^2)dt = \frac{\sqrt{2} - \exp(-\mu^2)}{2\mu}$$

を満たすときに，標準正規分布に従うモデルの MAE の期待値一致する．これはおよそ $\mu \simeq \pm 0.66665$ のときである．

16　「モデル a の MSE の代表値とモデル b の MSE の代表値は等しい」という帰無仮説の下で，次の検定統計量 z は自由度 $N_m - 1$ のカイ 2 乗分布に従う．

$$z = \frac{6(N_m - 1)}{N_t(N_m^3 - N_m)} \left(\sum_{i=1}^{N_t} r_{i,a} - r_{i,b} \right)^2$$

ここで，N_t は得られる MSE の数，つまり $393 - N_a - N_p$，$r_{i,k}$ は i 番目の基準時点（N_t 個ある基準時点のうち i 番目に古いものである必要はない）におけるモデル k による MSE の順位とする．検定は N_a と N_p の組み合わせごとに行う．なお，この検定方法は青木（2009）に基づく．

まず，$\tilde{X}$ を確率変数とし，i 番目の観測値（としての統計量の推定値）$\hat{x}_i$ が従う分布自体も確率的に決定されるものと仮定する．その分布の期待値（統計量の母数）を θ_i，θ_i の予測値を $\hat{\theta}_i$ と表す．また，この確率的に決まる分布の期待値の分布を θ_i が従う分布と呼び，任意の θ_i の下において観測値が従う分布を $\tilde{X}$ が従う分布と呼ぶことにする．以下に述べる修正モデルでは，この観測値を N_n 個得るものとして θ_i の予測，すなわち $\hat{\theta}_i$ の推定を行う．

ベイズの定理に基づけば，θ_i が従う分布の確率密度関数を $f(\hat{\theta}_i)$，f の定義域を K とし，$\theta_i = \hat{\theta}_i$ であると仮定した下で $\tilde{X}$ が従う分布の確率密度関数を $g(\hat{x}_i, \hat{\theta}_i)$ としたとき（g の定義域は K^2 である），$\tilde{X} = \hat{x}_i$ であった下での $\hat{\theta}_i$ の条件付き期待値は，次のように計算できる．

$$\mathrm{E}(\theta_i|\tilde{X} = \hat{x}_i) = \frac{\int_K f(\hat{\theta}_i)g(\hat{x}_i, \hat{\theta}_i)\hat{\theta}_i d\hat{\theta}_i}{\int_K f(\hat{\theta}_i)g(\hat{x}_i, \hat{\theta}_i)d\hat{\theta}_i} \tag{5}$$

推定値をこの値に修正する修正方法を一般にベイズ修正と呼ぶ．

ベイズ修正は θ_i が従う分布と $\tilde{X}$ が従う分布をどのように仮定するかによって修正値が大きく変わるが，本稿ではその違いによっていくつかの修正モデルを考える．

もっとも代表的なベイズ修正は，Vasicek（1973）が示したモデルである．これは f と g をそれぞれ次の関数 f_N および g_N と仮定するモデルである．

$$f_N(\hat{\theta}_i) = \frac{\exp\left(-\frac{(\hat{\theta}_i - \mu)^2}{2\mathrm{Var}(\theta)}\right)}{\sqrt{2\pi \mathrm{Var}(\theta)}}$$

$$g_N(\hat{x}_i, \hat{\theta}_i) = \frac{\exp\left(-\frac{(\hat{x}_i - \hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}\right)}{\sqrt{2\pi \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}}$$

なお，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ は $\theta_i = \hat{x}_i$ であるときの $\tilde{X}$ の分散，すなわち $\hat{x}_i$ の標準誤差の2乗であると考える[17]．つまり，標準誤差が $\hat{\theta}_i$ によらず $\hat{x}_i$ によって決定されると仮定している．また，μ は θ_i が従う分布の平均であり，この推定値 $\overline{x}$ は次式から得る．

$$\overline{x} = \frac{1}{N_n}\sum_{i=1}^{N_n} \hat{x}_i$$

$\mathrm{Var}(\theta)$ は θ_i が従う分布の分散であり，実際には次のように計算するものとする．

$$\mathrm{Var}(\theta) = \sum_{i=1}^{N_n} \frac{(\hat{x}_i - \overline{x})^2}{N_n - 1} - \frac{\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}{N_n}$$

ただし，ここから計算される値が負である場合，$\mathrm{Var}(\theta) \to +0$ とする．また，$K = \mathbb{R}$ である．Vasicek（1973）はこの場合の修正値の推定量を非常に簡単な形で示している．実際，このモデルの修正値（${}^B\hat{\theta}_i$ と表す）は次のように導出できる．

まず，

$$\begin{aligned}{}^B\hat{\theta}_i &= \frac{\int_{-\infty}^{\infty} f_N(\theta_i)g_N(\hat{x}_i, \hat{\theta}_i)\hat{\theta}_i d\hat{\theta}_i}{\int_{-\infty}^{\infty} f_N(\hat{\theta}_i)g_N(\hat{x}_i, \hat{\theta}_i)d\hat{\theta}_i} \\ &= \frac{\int_{-\infty}^{\infty} \exp\left(-\frac{(\hat{\theta}_i - \overline{x})^2}{2\mathrm{Var}(\theta)} - \frac{(\hat{x}_i - \hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}\right)\hat{\theta}_i d\hat{\theta}_i}{\int_{-\infty}^{\infty} \exp\left(-\frac{(\hat{\theta}_i - \overline{x})^2}{2\mathrm{Var}(\theta)} - \frac{(\hat{x}_i - \hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}\right)d\hat{\theta}_i}\end{aligned}$$

であるから，

$$\alpha = \frac{\mathrm{Var}(\theta) + \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}{\mathrm{Var}(\theta)\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}$$

$$\beta = \frac{\mathrm{Var}(\theta)\hat{x}_i + \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)\overline{x}}{\mathrm{Var}(\theta)\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}$$

$$\gamma = \frac{\mathrm{Var}(\theta)\hat{x}_i^2 + \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)\overline{x}^2}{\mathrm{Var}(\theta)\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}$$

とおけば，

$$\begin{aligned}{}^B\hat{\theta}_i &= \frac{\int_{-\infty}^{\infty} \exp\left(-2^{-1}\alpha\hat{\theta}_i^2 + \beta\hat{\theta}_i - 2^{-1}\gamma\right)\hat{\theta}_i d\hat{\theta}_i}{\int_{-\infty}^{\infty} \exp\left(-2^{-1}\alpha\hat{\theta}_i^2 + \beta\hat{\theta}_i - 2^{-1}\gamma\right)d\hat{\theta}_i} \\ &= \frac{\beta}{\alpha} = \frac{\mathrm{Var}(\theta)\hat{x}_i + \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)\overline{x}}{\mathrm{Var}(\theta) + \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)}\end{aligned} \tag{6}$$

である[18]．つまり，修正値は観測値 $\hat{x}_i$ とその全体の平均 $\overline{x}$ を，それぞれ $\mathrm{Var}(\theta)$ と $\hat{x}_i$ の標準誤差の2乗 $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ で重み付けした加重平均となっている．本稿ではこのモデルを Vasicek モデルと呼ぶ．

17　本稿で検証する統計量の標準誤差の推定量に関しては表4に示している．$\hat{x}_i$ の標準誤差は，必ずしも $\hat{x}_i$ の期待値 θ_i に依存するとは限らない．たとえば，平均の場合，標準誤差の推定値は標本の分散と標本数から得られる．本稿において，その場合は，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ を元の標本から推定される標準誤差の2乗であるものとする．

18　任意の定数 $\alpha, \beta, \gamma \in \mathbb{R}$（ただし，$\alpha > 0$）に関して，

$$\int_{-\infty}^{\infty} \exp\left(-\alpha x^2 + \beta x + \gamma\right)dx = \sqrt{\frac{\pi}{\alpha}}\exp\left(\frac{\beta^2}{4\alpha} + \gamma\right)$$

$$\int_{-\infty}^{\infty} \exp\left(-\alpha x^2 + \beta x + \gamma\right)x dx = \frac{\beta}{2\alpha}\sqrt{\frac{\pi}{\alpha}}\exp\left(\frac{\beta^2}{4\alpha} + \gamma\right)$$

である．

先に述べたとおり，Vasicek モデルは Klemkosky and Martin（1975）や Elton and Gruber（1978），Eubank and Zumwalt（1979），金崎（1987），吉原（1990），芳野（2011）など，多くの研究で検証されており，またそのいずれでもこのモデルの有効性が示されている．

Vasicek モデルでは期待値 θ_i に異なる値を仮定しても，観測値 $\hat{x}_i$ が同一ならば標準誤差の 2 乗 $\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{x}_i)$ が一定であると仮定した．しかし，標準誤差の推定量が期待値 θ_i から決定される関数として表わされるならば，標準誤差の 2 乗を $\theta_i=\hat{x}_i$ である下での標準誤差の 2 乗 $\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{x}_i)$ で代用し，期待値（として仮定する値）$\hat{\theta}_i$ に依存しないと仮定することは不合理である．その点の修正として，関数 g_N 中の $\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{x}_i)$ を $\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{\theta}_i)$ と置き換えることが考えられる．この場合の修正値（${}^V\hat{\theta}_i$ と表す）は次のように表せる．

$$ {}^V\hat{\theta}_i=\frac{\int_{-\infty}^{\infty} f_N(\theta_i)g_V(\hat{x}_i,\hat{\theta}_i)\hat{\theta}_i d\hat{\theta}_i}{\int_{-\infty}^{\infty} f_N(\hat{\theta}_i)g_V(\hat{x}_i,\hat{\theta}_i)d\hat{\theta}_i} \tag{7} $$

ただし，

$$ g_V(\hat{x}_i,\hat{\theta}_i)=\frac{\exp\left(-\frac{(\hat{x}_i-\hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{\theta}_i)}\right)}{\sqrt{2\pi\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{\theta}_i)}} $$

である．この値を修正値として用いるモデルを本稿では Variable モデルと呼び，検証する．なお，平均やベータ値の場合，標準誤差を期待値（として仮定する値）$\hat{\theta}_i$ の関数で表すことが出来ないため，本稿で検証する統計量のうち，この Variable モデルの対象となるものは分散だけである．

また，Vasicek モデルでは g を正規分布の確率密度関数と仮定していたが，これは必ずしも妥当ではない[19]．Vasicek（1973）はベータ値の予測に関して議論していたため，g を正規分布の確率密度関数と仮定することは妥当であったが，たとえば，分散の予測に際しては，その推定値が従う分布がカイ 2 乗分布と同じ形状（カイ 2 乗分布そのものではない）であって，決して正規分布ではない．

本稿で検証する統計量に限れば $\theta_i=\hat{\theta}_i$ の下において $\tilde{X}$ が従う分布は，その確率密度関数が明らかであるから，それが正規分布でなかったとしても，その確率密度関数（これを $g_T(\hat{x}_i,\hat{\theta}_i)$ と表す）を g と仮定して用いればよい．このときの修正値（${}^T\hat{\theta}_i$ と表す）は次のように表される．

$$ {}^T\hat{\theta}_i=\frac{\int_{-\infty}^{\infty}\frac{\exp\left(-\frac{(\hat{\theta}_i-\overline{x})^2}{2\mathrm{Var}(\theta)}\right)}{\sqrt{2\pi\mathrm{Var}(\theta)}}g_T(\hat{x}_i,\hat{\theta}_i)\hat{\theta}_i d\hat{\theta}_i}{\int_{-\infty}^{\infty}\frac{\exp\left(-\frac{(\hat{\theta}_i-\overline{x})^2}{2\mathrm{Var}(\theta)}\right)}{\sqrt{2\pi\mathrm{Var}(\theta)}}g_T(\hat{x}_i,\hat{\theta}_i)d\hat{\theta}_i} \tag{8} $$

予測する統計量によって g_T が異なるので，これ以上簡単に表すことはできない．また，g_T が $\hat{x}_i$ および $\hat{\theta}_i$ だけで決定される関数でないならば，その変数をどのように推定するかという問題が生じる．平均およびベータ値の予測における Vasicek モデルでは関数 g_N の中の分散がこれに該当しているが，観測値から推定される標準誤差 $\mathrm{Var}(\tilde{X}_i|\theta_i=\hat{x}_i)$ で代用している．ただし，本稿で検証する統計量に限っては，その問題が生じるものはない．

本稿で検証する統計量のうち，$\hat{x}_i$ が正規分布に従うと仮定し難いものは分散だけであり，その分布はカイ 2 乗分布の縮尺を変えた分布である．そこで，本稿ではこの値を修正値として用いるモデルを Theoretical モデルと呼び，分散の予測の場合に限って検証する．

Variable モデルや Theoretical モデルは Vasicek モデルと比較して極めて計算が複雑である．ゆえに，他のモデルよりも推定精度が優れていたとしても，実用的ではない．本稿ではあくまで Vasicek モデルの仮定を分散の予測に合わせることで予測精度が如何に変化するかを検証するために用いる．

次に，$\delta(x)$ をディラックのデルタ関数として，f および g として仮定する次のような関数 f_D および g_D を考える．

$$ f_D(\hat{\theta}_i)=\delta(\hat{\theta}_i-\overline{x}) $$

$$ g_D(\hat{x}_i,\hat{\theta}_i)=\delta(\hat{x}_i-\hat{\theta}_i) $$

これらに関して，まず，f を f_D と仮定するならば，g が g_D のような関数でない限りベイズ修正による修正値は $\overline{x}$ に一致する．つまり，すべての推定値を平均化するモデルであり，$\theta_i=\hat{\theta}_i$ の下で $\tilde{X}$ が従う分布が全く予測できない場合であるともいえる．

このモデルは Elton and Gruber（1978, pp.1377）や Eun and Resnick（1984, pp.1312）などで Mean モデル（原文ではそれぞれ Overall Mean，Mean Model）と呼ばれており[20]，本稿でもこの呼称を用い，検証する．また，とくに名称はつけていないが，金崎（1987）でも Mean モデルの予測能力を検証している．Mean モデルは計算

19　f を正規分布の確率密度関数と仮定することも必ず妥当であるとは言い難いが，観測値 $\hat{x}$ の分布から正規分布以外の f を推定することは容易でないため，本稿では検証しない．

20　より正確には，Elton and Gruber（1978）や Eun and Resnick（1984）における Mean モデルは，本稿でいう Mean モデルによるベータ値の修正値を相関係数の予測値を得るためのパラメータとして用いるモデルである．

が極めて簡単であるが，これらの先行研究のいずれにおいても比較的高い予測精度が示されている．

一方で，g を g_D と仮定するならば，f が f_D のような関数でない限りベイズ修正による修正値は $\hat{x}_i$ に一致する．つまり，修正を一切加えないモデルであり，Mean モデルとは反対に θ_i が従う分布が全く予測できない場合であるともいえる．このモデルを本稿では Historical モデルと呼び，検証する．Historical モデルは修正を加えないモデルであるから，予測精度がこれよりも高いか否かによって，その修正モデルが有効であるか否かを判断できる．

なお，Mean モデルと Historical モデルは，それぞれ $\mathrm{Var}(\theta) \to +0$，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i) \to +0$ であるときの Vasicek モデルと一致する．

最後に，メリルリンチが用いた修正モデルを述べる．ベータ値の予測に際して，メリルリンチは次のように修正値（${}^{ML}\hat{\theta}_i$ と表す）を計算している（Pastor（2001）などに記述がある）．

$$ {}^{ML}\hat{\theta}_i = \frac{2}{3}\hat{x}_i + \frac{1}{3} \tag{9} $$

この修正値は $\overline{x} = 1$ かつすべての i について $\mathrm{Var}(\theta) : \mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i) = 2 : 1$ であると仮定した場合の Vasicek モデルの修正値に等しい．このことは安達・池田（2016）でも，ベイズ修正の概要とともに言及されている．また，これはベータ値の予測を前提に考えられたものであって，上に述べた仮定をみればわかるように，他の統計量の予測に対しては必ずしも適しているといえない．そこで，本稿ではベータ値の予測に対してのみ，この修正方法を用いることとし，そのモデルをメリルリンチモデル（以下 ML モデル）と呼ぶ．

本稿ではこれらのモデルを用いて以下に述べる統計量を予測し，その予測精度を比較する．その際，すべての対象銘柄をまとめて予測するとともに，業種ごとの予測も試みる．これは θ_i が従う分布の特徴が業種によって異なることが想像できるためである．つまり，$f(\hat{\theta}_i)$ をすべての対象銘柄の θ_i が従う分布の確率密度関数ではなく，特定の業種に属する銘柄の θ_i が従う分布の確率密度関数であると考えて，同様の修正を行う．ただし，その業種に属する対象銘柄数が 10 を下回る場合は，それらの銘柄に限り，従来通り $f(\hat{\theta}_i)$ をすべての対象銘柄の θ_i が従う分布の確率密度関数として修正することにする．これは銘柄数が極端に少ないとき，観測された $\hat{x}$ の値に基づく $f(\hat{\theta}_i)$ の推定精度が著しく悪化する恐れがあるためである．本稿では，このように $f(\hat{\theta}_i)$ を求めるモデルを総称して sector モデルと呼び，そうでないモデルをとくに区別する必要がある場合，それらを総称して all モデルと呼ぶことにする[21]．この sector モデルと同様の考えに基づくモデルは Eun and Resnick（1984）や Karolyi（1992）でも検証されており，Eun and Resnick（1984）は本稿でいうところの Mean モデルに対して，Karolyi（1992）は Vasicek モデルに対して適用している[22]．

以上に述べたモデルの一覧を，以降の図表で用いる略称とともに表 3 に示す．また，Vasicek モデル，Variable モデル，Theoretical モデルの等確率密度曲線によるイメージを図 3 に示す．なお，これは分散の予測を前提にしており，f を正規分布の確率密度関数 f_N（ただし，その平均 μ および分散 $\mathrm{Var}(\theta)$ はそれぞれ 4，1 としている），$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{\theta}_i)$ や g_T を表 4 で示した分散の予測における関数と仮定したものである（$N_a = 5$ としている）．また，左の図が $g(\hat{x}_i, \hat{\theta}_i)$ を，右の図が $f(\hat{\theta}_i)g(\hat{x}_i, \hat{\theta}_i)$ を表しており，Vasicek モデルに関しては縦軸によらず g_N で用いる分散 $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ を $\mathrm{Var}(\tilde{X}_i|\theta_i = 3)$ としている．破線はベイズ修正の修正値を示している[23]．

3.4　予測する統計量

本稿で修正の効果を検証する統計量は，平均・分散モデルにおいて重要な統計量である平均，分散，および

21　あるモデルのうち，all モデルあるいは sector モデルだけをとくに区別する必要がある場合，モデル名の後に（all）あるいは（sector）を置き，呼び分けることとする．たとえば，Mean モデルであれば，その all モデルを Mean（all）モデル，sector モデルを Mean（sector）モデルと呼ぶ．また，とくに all モデルと sector モデルとを区別する必要がない場合は，単に Mean モデルと呼ぶ．

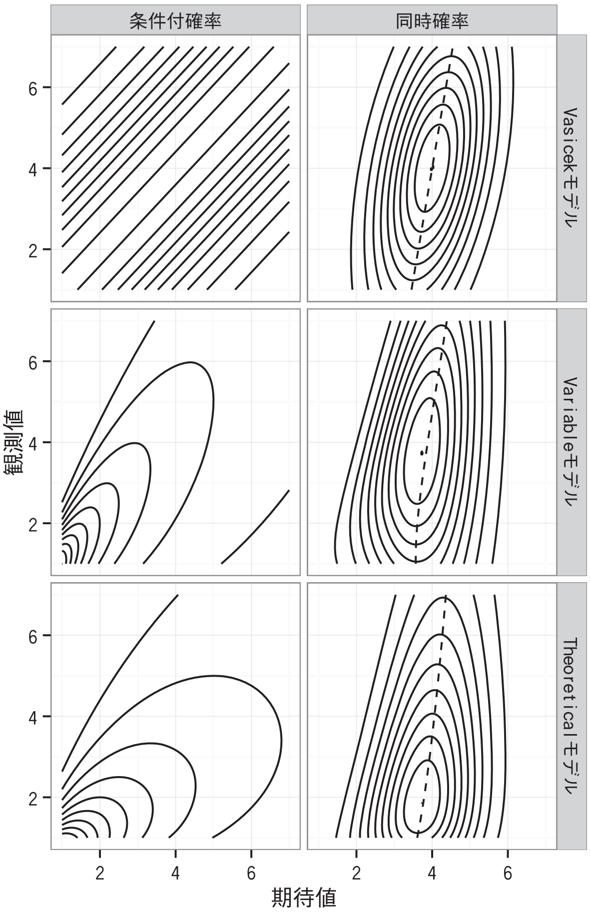
22　Eun and Resnick（1984）において Industry Mean Model と呼ばれるモデルが本稿でいうところの Mean（sector）モデルに対応するモデルであるといえる．ただし，Industry Mean Model は，より正確には Mean（sector）モデルによって得たベータ値の予測値を，相関係数の予測値を計算するためのパラメータとして用いるモデルである．Eun and Resnick（1984）では似たようなモデルとして National Mean Model というモデルも検証されており，これは業種ごとではなく，国ごとに予測するモデルである．本稿ではすべての銘柄が日本市場に上場する銘柄であるため，このようなモデルは適用できない．また，Karolyi（1992）において Industry Shrinkage と呼ばれるモデルが本稿でいうところの Vasicek（sector）モデルに対応するモデルであるといえる．Karolyi（1992）では銘柄を時価総額でグループ化し，そのグループごとに Vasicek モデルを適用する Size Shrinkage と呼ばれるモデルも検証しているが，本稿では検証期間が極めて長いため，このようなモデルも適用しなかった．

23　Vasicek モデルにおいて，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ を $\mathrm{Var}(\tilde{X}_i|\theta_i = 3)$ とするならば，ベイズ修正の修正値は図中の破線上において縦軸の値が 3 である点に対応する横軸の値しか，図 3 で示すことが出来ない．縦軸の値が 3 以外の点に対応する修正値は，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ の値が異なるために，図全体が変わり，同一破線上にないためである．

表３　モデルの一覧[24]

モデル	略称		仮定される関数	
	all	sector	$f(\hat{\theta}_i)$	$g(\hat{x}_i, \hat{\theta}_i)$
Historical モデル	H	—	任意	デルタ関数
Mean モデル	M	MS	デルタ関数	任意
Vasicek モデル	B	BS	正規分布	正規分布
Variable モデル	V	VS	正規分布	分散が $\hat{\theta}_i$ に依存する正規分布
Theoretical モデル	T	TS	正規分布	縮尺を変えたカイ 2 乗分布
ML モデル	ML	—	正規分布	正規分布

図３　各モデルのイメージ



マーケットモデルにおいて重要な統計量であるベータ値の 3 つである．

本節ではそれらの統計量ごとに推定値 $\hat{x}_i$ の推定量等を簡単にまとめる．本稿において $\hat{x}_i$ の推定量は株価変化率 $R_{t,i}$ の関数であり，$R_{t,i}$ は正規分布に従うと仮定する．また，ベータ値の推定には $R_{t,i}$ の他に説明変数も必要であるが，本稿では TOPIX の変化率を用い，その標本を M_t と示し，その標本平均を $\overline{M}$ と記す．なお，M_t も正規分布に従うと仮定する．

このときの各統計量の推定値 $\hat{x}_i$ の不偏推定量と $\theta_i = \hat{\theta}_i$ であるときの $\hat{x}_i$ の標準誤差の 2 乗の不偏推定量 $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{\theta}_i)$，および $\theta_i = \hat{\theta}_i$ であるときに $\tilde{X}$ が従う分布の確率密度関数 $g_T(\hat{x}_i, \hat{\theta}_i)$ を表 4 に示す．なお，基準時点を T 時点として示している．また，推定のために用いる標本数は N_a となる．

ここで，分散の予測の場合の $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{\theta}_i)$ には $\hat{\theta}_i$ を含み，なおかつそれ以外に $R_{t,i}$ 等，元の標本を含まない．一方で平均やベータ値の場合は $R_{t,i}$ や M_t を含む．ゆえに，Variable モデルを計算するのは分散の予測の場合だけである．また，平均やベータ値の場合の $g_T(\hat{x}_i, \hat{\theta}_i)$ は正規分布の確率密度関数であるため，Theoretical モデルを検証するのも分散の予測の場合だけであり，平均やベータ値の場合には Theoretical モデルを適用してもその修正値が Vasicek モデルに一致することになる[25]．

なお，本稿における事後の値も事後期間のデータを用いて表 4 に示した $\hat{x}_i$ の不偏推定量に準じて得ることになるが，その場合は N_a が N_p，$T - N_a$ が T，$T - 1$ が $T + N_p - 1$ となる．

4.　検証結果

以上に述べた方法に従って検証した結果を以下にまとめる．

4.1　MSE の推移

まず，$\{N_a, N_p\} = \{60, 60\}$ の場合の MSE の推移を図 4 および図 5 に示す．図 4 は MSE そのものの推移を，

24　Theoretical モデルに関しては分散の予測を前提にしており，そうでない場合は必ずしも表で示すような分布であるとは限らない．

25　表 4 をみると，平均やベータ値の $g_T(\hat{x}_i, \hat{\theta}_i)$ は $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{\theta}_i)$ を含んでおり，また，$\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{\theta}_i)$ は実際には $\hat{\theta}_i$ の関数で表せないため，Theoretical モデルを計算するとすればその推定の問題も生じるが，Vasicek モデルではこれを単に元の標本から計算される $\hat{x}_i$ の標準誤差の 2 乗，すなわち $\mathrm{Var}(\tilde{X}_i|\theta_i = \hat{x}_i)$ と表しているものとして推定しており，そのような推定を行う場合に限り，Theoretical モデルの修正値が Vasicek モデルの修正値と一致する．

表４　各統計量とその標準誤差の２乗の不偏推定量および推定値の確率密度関数

統計量	$\hat{x}_i$	$\mathrm{Var}(\tilde{X}_i \mid \theta_i = \hat{\theta}_i)$	$g_T(\hat{x}_i, \hat{\theta}_i)$
平均	$\overline{R_i} = \sum_{t=T-N_a}^{T-1} \frac{R_{t,i}}{N_a}$	$\sum_{t=T-N_a}^{T-1} \frac{(R_{t,i} - \overline{R_i})^2}{N_a(N_a-1)}$	$\frac{\exp\left(-\frac{(\hat{x}_i - \hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i \mid \theta_i = \hat{\theta}_i)}\right)}{\sqrt{2\pi \mathrm{Var}(\tilde{X}_i \mid \theta_i = \hat{\theta}_i)}}$
分散	$\sum_{t=T-N_a}^{T-1} \frac{(R_{t,i} - \overline{R_i})^2}{N_a - 1}$	$\frac{2\hat{\theta}_i^2}{N_a - 1}$	$\sqrt{\frac{\left(\hat{x}_i \frac{N_a - 1}{\hat{\theta}_i}\right)^{N_a - 3}}{2^{N_a - 1} \exp\left(\hat{x}_i \frac{N_a - 1}{\hat{\theta}_i}\right)} \frac{\frac{N_a - 1}{\hat{\theta}_i}}{\Gamma\left(\frac{N_a - 1}{2}\right)}}$
ベータ値	$\frac{\sum_{t=T-N_a}^{T-1} (R_{t,i} - \overline{R_i})(M_t - \overline{M})}{\sum_{t=T-N_a}^{T-1} (M_t - \overline{M})^2}$	$\frac{\frac{\sum_{t=T-N_a}^{T-1} (R_{t,i} - \overline{R_i})^2}{\sum_{t=T-N_a}^{T-1} (M_t - \overline{M})^2} - \hat{x}_i^2}{N_a - 2}$	$\frac{\exp\left(-\frac{(\hat{x}_i - \hat{\theta}_i)^2}{2\mathrm{Var}(\tilde{X}_i \mid \theta_i = \hat{\theta}_i)}\right)}{\sqrt{2\pi \mathrm{Var}(\tilde{X}_i \mid \theta_i = \hat{\theta}_i)}}$

図５はMSEを平均の誤差，標準偏差の誤差，相関の誤差に分けたときのそれぞれの推移を示している．ただし，すべてのモデルのグラフを表示すると個々を区別しづらくなるため，Historicalモデル，Mean（all）モデル，Vasicek（all）モデルだけを表示している．

図４　MSEの推移

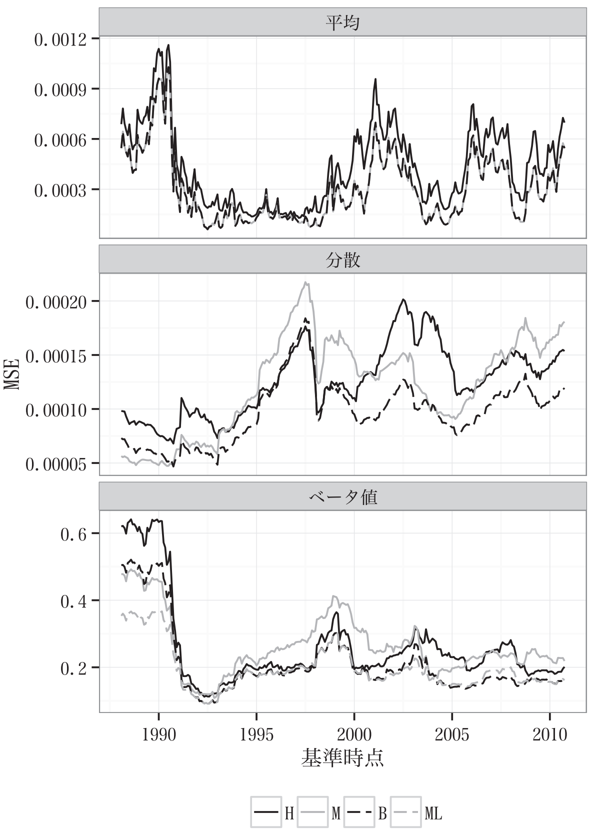


図４をみると，平均の予測に関してはMSEの変動幅に対してモデル間の差が小さいことがわかる．ほかの統計量に関しても，モデル間の差よりはMSEの変動のほうが大きい．ゆえに，MSEの平均で評価する場合，その水準の高かった時期の結果を強く反映することが懸念される．

MSEの水準は平均の予測ならば1990年代の間，低い水準にあるが，これは図５をみればわかるように，推定値全体の平均，すなわち市場全体でみた株価変化率の平均的な水準の変動が，この期間に小さかったことによるところが強い．同様のことが2005年前後や2008年ごろにもいえる．それ以外の時期では平均の誤差の割合が極めて高く，平均の予測におけるMSEの水準を左右する要因は，推定値全体の平均の水準にあると考えられる．一方で，分散やベータ値の予測の場合は相関の誤差によるところが大きい．しかし，とくにベータ値の予測では，その水準の変動が安定的ないしは傾向的であり，MSEの水準の急激な変動は平均の誤差に起因するようである．ベータ値の予測におけるMSEの水準は，1980年代ごろに極めて高いが，このときどのモデルでもMSEの半分以上を平均の誤差が占めている．

なお，Mean（all）モデルのMSEに占める標準偏差の誤差の割合は，とくに分散の予測において極めて大きくその変動も激しいが，この変動は単に事後の値の分散の推移を表していることになる．Mean（all）モデルのMSEは，事後の値の分散に平均の誤差を加えたものに等しいため，Mean（all）モデルでは他のモデルに比べて推定値のばらつき具合の変動がMSEの水準の安定性に強く影響することになる．

図４をみると，平均の予測ではHistoricalモデルが，分散やベータ値の予測ではHistoricalモデルかMean（all）モデルのいずれかが最も大きい時期が多いことがわかる．一方で，Vasicek（all）モデルはいずれの統計量でもほとんどの時点で最も低い水準を推移していることがわかる．とくに，分散の予測の場合，2000年代以降のVasicek（all）モデルのMSEはHistoricalモデルのそれ

図 5　MSE の各成分の推移

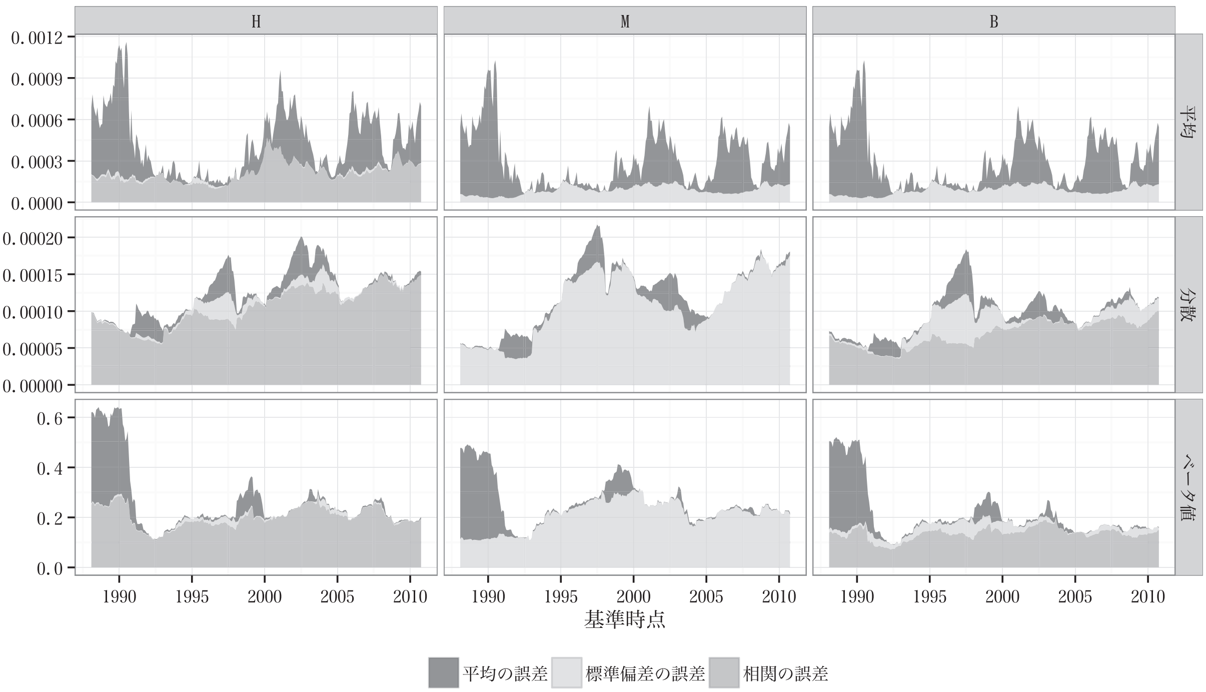


表 5　平均の予測における MSE の中央値，平均，標準偏差

$N_p \backslash N_a$	中央値 (10^{-2})					平均 (10^{-2})					標準偏差 (10^{-2})				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.18911	.12091	.10241	.08412	.07947	.22502	.14670	.12888	.10438	.10312	.11907	.08094	.07961	.06499	.06169
M	.09785	.09185	.08480	.07600	.07508	.13834	.11738	.11277	.09765	.09839	.09810	.07598	.07788	.06399	.06038
MS	.11028	.09505	.08686	.07687	.07541	.14901	.12113	.11482	.09862	.09895	.10164	.07697	.07856	.06415	.06056
B	.09847	.09185	.08480	.07600	.07508	.13830	.11724	.11277	.09765	.09839	.09775	.07582	.07788	.06399	.06038
BS	.10987	.09505	.08691	.07687	.07541	.14903	.12110	.11482	.09861	.09895	.10170	.07690	.07854	.06415	.06056
36															
H	.13385	.06496	.04823	.03599	.03077	.15381	.07911	.06097	.04288	.03897	.08266	.05535	.03975	.02494	.02516
M	.04968	.03610	.03267	.02965	.02777	.07031	.05284	.04730	.03671	.03536	.06270	.04903	.03712	.02401	.02471
MS	.05969	.03961	.03353	.03104	.02799	.08145	.05636	.04910	.03776	.03555	.06496	.05050	.03769	.02424	.02458
B	.05020	.03610	.03267	.02965	.02777	.07020	.05264	.04730	.03671	.03536	.06163	.04815	.03712	.02401	.02471
BS	.06094	.03948	.03353	.03104	.02799	.08137	.05625	.04909	.03776	.03555	.06440	.05009	.03766	.02424	.02458
60															
H	.12474	.05514	.03923	.02642	.01922	.14211	.06375	.04357	.02939	.02069	.07910	.04075	.02410	.01374	.00941
M	.04259	.02682	.02465	.02090	.01551	.05963	.03892	.03043	.02342	.01759	.05694	.03363	.02058	.01220	.00920
MS	.05583	.03088	.02753	.02161	.01610	.07109	.04243	.03244	.02455	.01773	.05960	.03496	.02142	.01277	.00895
B	.04356	.02682	.02465	.02090	.01551	.05950	.03871	.03043	.02342	.01759	.05513	.03297	.02058	.01220	.00920
BS	.05585	.03088	.02753	.02161	.01610	.07089	.04230	.03243	.02455	.01773	.05857	.03464	.02140	.01277	.00895

を大きく下回る位置にある．これらのことから，Vasicek（all）モデルはどの統計量の予測に対しても修正の効果が認められることがうかがえる．

ベータ値の予測では 1980 年代にモデル間の差が明白になっており，とくに ML モデルの MSE は相対的にみて極めて小さい．図 5 をみると，この時期の MSE に関しては平均の誤差の影響が強いことがわかるが，ML モデルでは推定値の平均をあらかじめ定数として設定しているため，その影響が軽減されていると考えられる．

図 5 に関して，平均の予測における Vasicek モデルでは相関の誤差がほぼ 0 である．これは $\mathrm{Var}(\theta)$ が 0 であるかそれに近いことを示している．$\mathrm{Var}(\theta) \to +0$ であれば Vasicek モデルの修正値は Mean モデルの修正値と一致し，事後の値との標本共分散と修正値の標本分散がともに 0 になるためである．

4.2　MSE の集計

次に，MSE の中央値，平均，標準偏差を，予測する統計量ごとに表 5，表 6，表 7 に示す．なお，N_a および N_p ごとにそれぞれの統計量の最小値および最大値の背景を

表６　分散の予測における MSE の中央値，平均，標準偏差

$N_p \backslash N_a$	中央値 (10^{-3})					平均 (10^{-3})					標準偏差 (10^{-3})				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.56086	.37221	.26655	.22238	.22450	.57699	.39981	.33568	.28902	.26146	.31526	.22005	.20415	.18565	.13316
M	.32128	.29258	.24982	.22246	.24206	.41143	.38301	.34963	.31365	.28770	.28234	.27246	.25459	.22463	.16571
MS	.32893	.29007	.24304	.21576	.24019	.40764	.37449	.33927	.30343	.27786	.26797	.26015	.24374	.21567	.15801
B	.31054	.28364	.22458	.19677	.21587	.38107	.34494	.30810	.27883	.25617	.26487	.24439	.22086	.19753	.14103
BS	.31542	.29252	.22608	.20454	.22044	.38288	.34677	.30976	.28014	.25716	.25905	.23862	.21740	.19458	.13905
V	.34641	.30575	.23903	.20586	.21976	.39653	.35383	.31360	.28171	.25788	.24457	.22563	.20963	.19150	.13742
VS	.36346	.31063	.24480	.21187	.21976	.41181	.35919	.31643	.28294	.25822	.24502	.22297	.20804	.18962	.13618
T	.34001	.30149	.23646	.20597	.21877	.39230	.35185	.31243	.28116	.25758	.24274	.22653	.21027	.19189	.13773
TS	.35897	.30897	.24403	.20966	.22011	.40652	.35730	.31551	.28264	.25790	.24309	.22363	.20850	.18973	.13639
36															
H	.34979	.20818	.16177	.13051	.11747	.39498	.21256	.16412	.13426	.12017	.21186	.07648	.05470	.03761	.02346
M	.19667	.16544	.15815	.15269	.13546	.21289	.18342	.16996	.15229	.14451	.11194	.09565	.08318	.05559	.04191
MS	.21011	.16401	.14955	.14148	.13028	.21303	.17635	.16086	.14254	.13422	.10673	.08766	.07548	.05069	.03763
B	.18068	.13546	.12557	.12055	.11355	.19490	.15431	.13705	.12319	.11499	.09480	.07264	.06096	.04352	.02787
BS	.18451	.13931	.13243	.12188	.11388	.19819	.15694	.13870	.12482	.11560	.09310	.07044	.05898	.04219	.02676
V	.20043	.15687	.13594	.12238	.11440	.21571	.16533	.14275	.12631	.11673	.10366	.06680	.05563	.04015	.02603
VS	.20521	.16356	.13795	.12471	.11404	.23145	.17118	.14534	.12769	.11656	.11099	.06739	.05494	.03931	.02544
T	.20061	.15391	.13511	.12167	.11409	.21228	.16340	.14161	.12584	.11644	.10072	.06642	.05569	.04038	.02623
TS	.20367	.16048	.13679	.12427	.11401	.22678	.16922	.14444	.12742	.11652	.10790	.06674	.05498	.03938	.02550
60															
H	.27067	.16095	.12386	.09777	.08153	.34910	.17013	.12549	.09914	.08447	.21240	.05905	.03245	.01417	.01285
M	.14891	.12665	.13001	.11197	.10770	.16001	.13470	.12193	.11392	.10940	.08557	.05784	.04435	.02348	.02022
MS	.15105	.12205	.12037	.10255	.09599	.16221	.12898	.11421	.10435	.09874	.08400	.05188	.03870	.02102	.01871
B	.13019	.10720	.09787	.08372	.07735	.14660	.11266	.09672	.08828	.08002	.05984	.03860	.02918	.01835	.01491
BS	.13773	.11248	.10124	.08548	.07700	.15065	.11601	.09925	.08968	.08011	.06119	.03812	.02857	.01761	.01461
V	.14286	.11904	.10767	.08710	.07884	.17113	.12416	.10340	.09124	.08147	.09212	.04098	.02773	.01590	.01414
VS	.15754	.12372	.10994	.08957	.07869	.18726	.13017	.10668	.09244	.08130	.10045	.04330	.02849	.01536	.01385
T	.14335	.11797	.10599	.08663	.07855	.16825	.12222	.10239	.09080	.08129	.08952	.03980	.02752	.01608	.01420
TS	.15263	.12235	.10906	.08958	.07851	.18314	.12836	.10576	.09219	.08121	.09735	.04202	.02820	.01546	.01390

表７　ベータ値の予測における MSE の中央値，平均，標準偏差

$N_p \backslash N_a$	中央値 (10^{0})					平均 (10^{0})					標準偏差 (10^{0})				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.99051	.58026	.55769	.48830	.53704	1.11276	.68737	.64377	.59053	.59310	.59225	.37866	.37209	.36155	.30608
M	.63082	.56452	.58924	.59237	.58795	.74695	.67018	.67230	.66024	.66521	.42600	.37040	.37646	.38097	.30689
MS	.67085	.53379	.56365	.55009	.56664	.75679	.64559	.63913	.61690	.62269	.42768	.36158	.36939	.37113	.30689
B	.57622	.48500	.52076	.50215	.53299	.70913	.60684	.60236	.58006	.58723	.42754	.35619	.36148	.36464	.30542
BS	.61470	.49138	.50917	.49569	.53413	.74263	.61189	.60237	.57733	.58458	.43476	.35668	.36290	.36368	.30529
ML	.72379	.48446	.49442	.49092	.52993	.81942	.60512	.58881	.57884	.59240	.45455	.35853	.35648	.36472	.31229
36															
H	.62433	.35875	.27958	.23710	.21739	.73929	.37607	.31261	.25337	.23789	.41184	.16018	.14570	.06101	.05617
M	.32683	.31629	.31500	.30204	.28268	.35376	.33190	.32705	.30939	.30236	.17959	.12189	.11451	.07782	.06145
MS	.33862	.28661	.28011	.26359	.23869	.37246	.31278	.29673	.26817	.25985	.19480	.13297	.12516	.06911	.06340
B	.30159	.25709	.24649	.23044	.20770	.33027	.28289	.26684	.23950	.23153	.17596	.12359	.12111	.06584	.06021
BS	.33031	.25457	.23837	.22577	.20719	.36715	.28860	.26672	.23653	.22821	.19861	.13453	.12884	.06537	.06035
ML	.39358	.27387	.24191	.22023	.20501	.43628	.27309	.24928	.23562	.23437	.20848	.09350	.08547	.06581	.06405
60															
H	.59464	.29979	.21051	.18751	.15869	.72838	.34386	.26303	.19994	.17353	.43113	.17649	.13289	.04149	.03679
M	.26181	.25736	.24243	.23799	.23361	.31072	.28322	.26502	.24597	.23582	.18396	.10959	.09072	.04372	.03566
MS	.25862	.21221	.19875	.19839	.16838	.33608	.26707	.23639	.20651	.19197	.21731	.14156	.11142	.04711	.04335
B	.23652	.19304	.17737	.16399	.14758	.29702	.24316	.21264	.18388	.16791	.19237	.13094	.10732	.04561	.04046
BS	.25452	.18975	.17192	.15869	.14392	.33660	.24864	.21267	.18094	.16371	.22358	.14733	.11695	.04655	.04093
ML	.34734	.21115	.18052	.16416	.15003	.39836	.22745	.19624	.18025	.17046	.19674	.08273	.06699	.04474	.04333

グレーで塗り潰しており，より薄いグレーで塗り潰している方が最小値である．

まず，平均の予測の結果（表５）をみると，MSE の中央値や平均は N_a や N_p によらず，Historical モデルが最も大きく，Mean（all）モデルあるいは Vasicek（all）モデルが最も小さい結果となっている．とくに，Mean（all）モデルは推定値を平均化するだけの単純なモデルであるから，その予測精度が最も高いならば，平均の予

測には Mean（all）モデルが最適だと容易に結論づけられる．なお，Mean モデルと Vasicek モデルの結果の差が非常に小さいのは，多くの，あるいはすべての時点で $\mathrm{Var}(\theta) \to +0$ と計算されるためである．これは母平均のばらつきに対して個別銘柄の標本平均のばらつき，すなわち標準誤差が極めて大きく推定されることを意味する．ゆえに，平均の予測の場合，Mean モデルを用いても Vasicek モデルを用いても大差がなく，より計算が簡単な Mean モデルのほうが適しているといえる．

平均の予測において Mean（all）モデルが最適であるという結論は，株価変化率の平均の予測が極めて難しいことを示唆する．後述するが分散やベータ値の予測では Mean（all）モデルが最適という結論に至らない．これらの予測に比べて平均，すなわち期待収益率の予測が困難であるという事実は，Merton（1980）でも指摘されている．

MSE の中央値や平均に関して，all モデルと sector モデルとを比較すると，Mean モデルも Vasicek モデルも all モデルのほうが優れていることがわかる．ゆえに，平均の推定値の分布は業種によって特徴づけられるが，その特徴が経時的に安定しないと考えられる．

MSE の標準偏差をみると，Mean モデルも Vasicek モデルも Historical モデルよりは小さい．また，all モデルと sector モデルとを比較すると，Mean モデルでも Vasicek モデルでも $N_a = 180$ かつ $N_p \geq 36$ の場合を除いて all モデルのほうが小さい．つまり，このとき Mean（all）モデルや Vasicek（all）モデルは MSE の水準とばらつきがともに Historical モデルよりも小さく，いっそう有用なモデルであるといえる．

N_a と N_p ごとにみると，いずれのモデルも MSE の中央値や平均，標準偏差は N_a や N_p が大きいほど小さくなることがわかる．ゆえに，平均の予測に際してはなるべく多くのデータを用いたほうがよいといえる．ただし，Historical モデル以外のモデルに関して，$N_p = 12$ のとき，MSE の平均は $N_a = 180$ の場合よりも $N_a = 120$ の場合のほうが小さく，MSE の標準偏差は $N_a = 60$ の場合よりも $N_a = 36$ の場合のほうが小さい．この点を考慮すると，N_p が小さい場合，すなわち近い将来の値を予測する場合は，あまり多すぎるデータを用いるべきではないといえる．MSE の平均で比較するならば，過去 120 か月間程度の月次データで充分である．

次に，分散の予測結果（表 6）をみると，MSE の中央値や平均は，N_a が小さいときは Historical モデルが，N_a が大きいときは Mean（all）モデルが最大となることがわかる．Mean モデル以外のモデルは，MSE の中央値や平均が Historical モデルのそれを上回ることはなく，Mean モデル以外のモデルではベイズ修正の効果が多少なりともあるといえる．

MSE の中央値や平均が最も小さいモデルはすべて Vasicek モデルであり，Vasicek モデルは Variable モデルや Theoretical モデルよりも計算が簡単なことから，分散の予測に対しては Vasicek モデルが適していると結論付けられる．

また，all モデルと sector モデルとを比較すると，MSE の中央値や平均は，N_a や N_p を問わずどのモデルでも all モデルのほうが小さいことがほとんどである．つまり，平均の予測と同じように，業種ごとに分けて推定する価値がないといえる．あるいは，分散の分布に業種ごとの大きな特徴があるとは考え難いともいえる．

N_a と N_p ごとにみると，やはりそれらが大きいほど MSE の中央値や平均，標準偏差は小さくなる傾向にある．ただし，$N_p = 12$ のとき，MSE の中央値は $N_a = 180$ の場合よりも $N_a = 120$ の場合のほうが小さい．ゆえに，平均の予測と同じように，分散の予測でも基本的にはなるべく多くのデータを用いるべきであるが，近い将来の分散を予測する場合は過去 120 か月分程度のデータで充分であり，それ以上古いものを用いるとかえって予測精度が悪化する恐れがある．このように，より最近のデータに重点を置くほうが予測精度が向上する点は，修正に週次のデータを用いることでより明瞭に観察できることが予想される．

N_a が大きいとき，Variable モデルと Theoretical モデルとの差は非常に小さいが，これは Theoretical モデルで仮定する分布が標本数 N_a の増大にともなって正規分布に近似されるためである．

最後に，ベータ値の予測結果（表 7）をみると，分散の予測結果と同じように Historical モデルと（N_a が小さくないときの）Mean モデルの予測能力がとりわけ低いが，最も優れたモデルを判別するのは難しくなっている．N_a や集計の仕方（つまり，中央値で比べるか平均で比べるか）によってモデルの優劣が若干異なっているためである．

N_a ごとにみると，$N_a = 12$ の場合は Vasicek（all）モデルがとくに優れている．MSE の中央値でみても平均でみても他と大きく離れて最小であるので，この場合は Vasicek（all）モデルが適しているといえる．なお，$N_a = 12$ のとき，Mean モデルは Vasicek モデルにこそ劣るものの，Historical モデルに対しては大きく優れる結果である．一般的に，推定値を得るための標本数 N_a が小さいほど標準誤差が大きくなり，推定値の予測精度が落ちるため，個々を区別せずにすべての修正値に同じ値を用いる Mean モデルが優れる結果になることは不思

議でない．上述の通り，Mean モデルは $N_a \geq 36$ のときHistorical モデルに匹敵するほどの予測能力の低さである．Vasicek（sector）モデルよりも Vasicek（all）モデルのほうが優れる理由も同様であろうと考えられる．

$36 \leq N_a \leq 60$ の場合は ML モデルあるいは Vasicek（sector）モデルが優れることが多い．MSE の平均で見ると，Vasicek（sector）モデルよりも ML モデルのほうが優れるようである．ML モデルが MSE の平均に関してとくに優れるのは，MSE の水準が全体的に高かった1980 年代に ML モデルの MSE だけが相対的に著しく低水準であったことに起因すると考えられる（図 4 参照）．ある一時期だけ顕著に優れる場合，平均はその影響を強く受けるものの，中央値はその影響を受けづらい．また，MSE の標準偏差をみると，このとき ML モデルがとくに小さい．これらの結果は，MSE の水準が著しく高かった 1980 年代に，ML モデルだけは相対的に極めて低い水準であったことに起因する（図 5 参照）．このことを踏まえれば，どちらかといえば ML モデルを用いるべきであるといえる．

$120 \leq N_a \leq 180$ の場合は MSE の中央値をみても平均をみても，ML モデルが最も優れる場合と Vasicek（sector）モデルが最も優れる場合とが混在しており判断が難しい．ただ，この場合の ML モデルの標準偏差はやや大きいため，どちらかといえば Vasicek（sector）モデルのほうが優れているといえるかもしれない．しかし，その差がわずかであることを考えれば，より計算が簡単な ML モデルのほうが適しているともいえる．

ベータ値の予測において all モデルと sector モデルをと比較すると，平均や分散の予測の場合と異なり，Mean モデルも Vasicek モデルも基本的に sector モデルのほうが優れる結果となっている．ただし，$N_a = 12$ の場合は all モデルのほうが優れるようである．上でも述べたように，短期間のデータを用いた推定では業種ごとの特徴がうまくとらえられないためだと考えられる．

N_a と N_p ごとに予測精度を比較すると，やはりそれらが大きいほど MSE の中央値や平均，標準偏差が小さくなる．しかし，ベータ値の予測の場合は，平均や分散の予測の場合よりも，その例外が顕著であり，$N_p = 12$ のとき，$N_a = 180$ の場合の MSE の中央値と平均は，$N_a = 120$ の場合に大きく劣り，とくに中央値では Historical モデル以外のすべてのモデルで $N_a = 36$ の場合よりも大きくなっている．中央値だけをみれば，$N_p = 12$ の場合は $N_a = 36$ のときにそれらの値が最小になるモデルが多いようである．ただし，MSE の標準偏差は $N_a = 36$ のとき，Historical モデルを除いて，$N_a = 60$ や $N_a = 120$ の場合よりも小さくなるモデルが多いが，$N_a = 180$ の場合よりは大きい．これらのことから，ベータ値の予測でも基本的にはなるべく多くのデータを用いて予測するべきだが，近い将来の値を予測したい場合は，あまり遠くない過去，具体的には過去 36 か月分程度のデータが最適であるといえる．

ところで，平均の予測では Mean（all）モデル（および Vasicek（all））が圧倒的に優れていた．その要因を考えると，推定値と事後の値との共分散が負であった可能性が挙げられる．Mean（all）モデルでは修正値の標準偏差が 0 になるため，MSE のうちの相関の誤差も 0 となり，MSE は平均の誤差と事後の値の標本分散の和に一致する．つまり Mean モデル以外のモデルの MSE は Mean モデルの MSE よりも修正値の標本分散だけ大きく，事後の値との標本共分散の 2 倍だけ小さな値となる．ゆえに，共分散が負であるならば，必ず Mean（all）モデルのほうが優れることになる．

そこで，推定値と事後の値との共分散が負であった割合を表 8 に示す．やはり，平均の予測の場合はその割合が著しく高いことがわかる．また，ベータ値の予測において $N_a = 12$ の場合に Mean（all）モデルが比較的優れていたことと同様に，$N_a = 12$ の場合のこの割合も，（ベータ値の予測におけるほかの場合と比べれば）やや大きめである．一方で，分散の予測の場合はこの割合が非常に小さく，すべて 0 である．

表 8　共分散が負である割合（%）

$N_p \backslash N_a$	12	36	60	120	180
平均					
12	61.52	75.94	81.93	77.39	74.13
36	73.33	80.37	77.44	84.81	89.83
60	76.64	84.18	78.75	89.20	85.62
分散					
12	0.00	0.00	0.00	0.00	0.00
36	0.00	0.00	0.00	0.00	0.00
60	0.00	0.00	0.00	0.00	0.00
ベータ値					
12	7.32	2.61	2.49	1.15	1.00
36	5.51	0.00	0.00	0.00	0.00
60	5.92	0.00	0.00	0.00	0.00

平均の予測において推定値と事後の値の共分散が負であるということは，過去の株価変化率が比較的高かった銘柄は，将来比較的低くなる傾向にあることを示している．つまり，過去の株価変化率を参考にして，それが大きい銘柄に対して重点的に投資を行ういういわゆるモメンタム投資は，この結果をみる限りあまり合理的とはいえないことになる．

修正値と事後の値との共分散が正であったからといって Mean（all）モデルの MSE が他のモデルに劣るわけではない．共分散が正であっても，推定値の分散が共分散の 2 倍よりも大きい場合は Mean（all）モデルのほうが優れることになる．そこで，Historical モデルの推定値の分散が共分散の 2 倍よりも大きくなる割合を表 9 に示す．これは単に Mean（all）モデルが Historical モデルよりも優れる割合に等しい．この結果は MSE の平均や中央値の比較と整合的であることがわかるが，平均の予測において常に 100% とはならないことは厄介である．これは先に述べたように平均の予測では Mean（all）モデルが最も優れるものの，時期によっては（総合的に見て最も予測精度の低い）Historical モデルよりも劣る場合があることを意味する．ただし，それは $N_p \leq 36$ の場合に限られる．

表 9　推定値の分散が共分散の 2 倍よりも大きい割合（%）

$N_p \backslash N_a$	12	36	60	120	180
平均					
12	99.73	98.26	95.95	93.87	93.03
36	100.00	100.00	97.31	100.00	100.00
60	100.00	100.00	100.00	100.00	100.00
分散					
12	86.72	62.61	51.09	30.65	26.87
36	92.75	69.47	44.44	22.78	13.56
60	91.59	69.36	47.99	20.66	0.00
ベータ値					
12	91.33	49.86	33.33	19.16	13.93
36	97.68	61.99	33.00	6.75	0.00
60	98.13	67.00	37.00	4.69	0.00

4.3　MSE の各成分の集計

次に，MSE の各成分，すなわち平均の誤差，標準偏差の誤差，相関の誤差のそれぞれの平均を予測する統計量ごとに表 10，表 11，表 12 に示す．これらの表も表 5，表 6，表 7 と同様に，N_a および N_p ごとに最小値および最大値の背景をグレーで塗り潰している．

3 つの表を見ると，平均の誤差は，どの統計量でも N_a を大きくするほど小さくなっていることがわかる．ただし，ベータ値の予測の場合，$N_p = 12$ ならば $N_a = 36$ のときに MSE の中央値が最小となっていたように，平均の誤差もこのとき $N_a = 60$ の場合よりも小さい．この例外を除けば，この結果は，事後期間 N_p の長さによらず，より長期間のデータを用いるほど事後の値の平均を予測する能力が効果的に高まることを示している．ゆえに，修正値全体の平均が他のモデルの修正値の平均に一致するような調整を行うとすれば，その効果は N_a が小さいほど大きいといえる．この調整によって縮小（あるいは拡大）する MSE の平均は，当然ながらその 2 モデルの MSE における平均の誤差の平均の差に等しい．

たとえば，平均の予測では Mean（all）モデルが MSE の中央値や平均に関して最も優れていたが，平均の誤差の平均は $N_a \leq 36$ のとき，Vasicek（all）モデルが最小となっている（表 10 参照）．そこで，Mean（all）モデルによる修正値全体の平均が Vasicek（all）モデルによる修正値全体の平均に一致するように調整すると，$\{N_a, N_p\} = \{12, 12\}$ の場合，その調整後の MSE の平均は $0.0006327 + 0.0007400 + 0.0000000 = 0.0013727$ となり，Mean（all）モデルの MSE の平均である $0.0006434 + 0.0007400 + 0.0000000 = 0.0013834$ よりは若干小さくなる．しかし，この場合，その効果は非常にわずかであり，さらにこの調整をするためには Vasicek（all）モデルの修正値を計算しなければならないため，到底現実的とはいえないうえに，これが予測精度の向上に貢献することを理論的に示すことが容易でない．計算の容易さの点において，この調整が有用であるのは，修正値の平均を Historical モデルあるいは Mean（all）モデルの修正値の平均に一致させるように調整することで予測精度が向上する場合であるといる．これらのモデルの修正値の平均は，Vasicek モデルや Variable モデルでは必要不可欠なパラメータであり，それ以外のモデルに適用するとしてもこの計算が簡単であるためである．また，Historical モデルや Mean（all）モデルでは，推定値がそれぞれ独立に同一の分布に従うとき，修正値の平均の期待値が事後の値の平均の期待値に一致することが明らかであるから，この調整の根拠も説明しやすい．その点を踏まえると，平均の予測においては Mean（all）モデルが最も優れていたため，この調整を施す価値はないことがわかる．

分散の予測の場合，Vasicek（all）モデルが最も優れていたが，平均の誤差は常に Historical モデルよりも Vasicek（all）モデルのほうが小さく，さらにいえば Vasicek（all）モデルの平均の誤差は $\{N_a, N_p\} = \{12, 12\}$ を除いて常に最小である（表 11 参照）．ゆえに，分散の予測の場合でも修正値の平均を調整する操作は必要ないと思われる．

ベータ値の予測の場合は，$N_a \leq 120$ のときに ML モデルの平均の誤差が著しく小さい点が目立つ（表 12 参照）．ML モデルはその修正値の計算が極めて簡単であるから，ML モデルの平均に合わせて調整する操作は有用であるといえる．ただし，ML モデル自体，ベータ値の予測において比較的優れたモデルであったから，必ずしも適用できるとは限らない．ベータ値の予

表 10　平均の予測における MSE の各成分の集計

$N_p \backslash N_a$	平均の誤差の平均 (10^{-2})					標準偏差の誤差の平均 (10^{-2})					相関の誤差の平均 (10^{-2})				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.06434	.04952	.04951	.03720	.03375	.00451	.01537	.02386	.03436	.04383	.15617	.08181	.05551	.03281	.02554
M	.06434	.04952	.04951	.03720	.03375	.07400	.06787	.06326	.06045	.06463	.00000	.00000	.00000	.00000	.00000
MS	.06450	.04974	.04952	.03724	.03374	.03297	.04366	.04595	.04890	.05595	.05155	.02773	.01934	.01248	.00926
B	.06327	.04928	.04951	.03720	.03375	.07059	.06754	.06326	.06045	.06463	.00444	.00041	.00000	.00000	.00000
BS	.06373	.04960	.04949	.03723	.03374	.03238	.04351	.04592	.04890	.05595	.05293	.02799	.01940	.01248	.00926
36															
H	.05020	.03461	.03022	.01935	.01782	.01957	.00104	.00186	.00521	.00788	.08404	.04346	.02889	.01833	.01326
M	.05020	.03461	.03022	.01935	.01782	.02011	.01823	.01709	.01736	.01754	.00000	.00000	.00000	.00000	.00000
MS	.05045	.03474	.03019	.01936	.01778	.00400	.00739	.00916	.01138	.01335	.02700	.01423	.00975	.00702	.00443
B	.04901	.03428	.03022	.01935	.01782	.01857	.01802	.01709	.01736	.01754	.00262	.00033	.00000	.00000	.00000
BS	.04952	.03451	.03016	.01936	.01778	.00390	.00733	.00914	.01138	.01335	.02795	.01442	.00979	.00702	.00443
60															
H	.04954	.02963	.02157	.01391	.00818	.03213	.00309	.00081	.00152	.00294	.06044	.03103	.02120	.01395	.00957
M	.04954	.02963	.02157	.01391	.00818	.01009	.00929	.00886	.00950	.00942	.00000	.00000	.00000	.00000	.00000
MS	.04967	.02966	.02157	.01393	.00815	.00173	.00254	.00363	.00522	.00653	.01970	.01023	.00724	.00540	.00306
B	.04837	.02937	.02157	.01391	.00818	.00919	.00918	.00886	.00950	.00942	.00193	.00017	.00000	.00000	.00000
BS	.04869	.02945	.02154	.01393	.00815	.00186	.00252	.00363	.00522	.00653	.02034	.01033	.00727	.00540	.00306

表 11　分散の予測における MSE の各成分の集計

$N_p \backslash N_a$	平均の誤差の平均 (10^{-3})					標準偏差の誤差の平均 (10^{-3})					相関の誤差の平均 (10^{-3})				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.04371	.03696	.03210	.02317	.01430	.05606	.06425	.06499	.08956	.09049	.47721	.29860	.23859	.17630	.15667
M	.04371	.03696	.03210	.02317	.01430	.36772	.34604	.31752	.29049	.27340	.00000	.00000	.00000	.00000	.00000
MS	.04396	.03698	.03223	.02314	.01430	.20210	.20975	.19366	.18925	.17769	.16158	.12776	.11337	.09104	.08588
B	.04390	.03596	.03107	.02216	.01345	.15906	.13619	.11637	.11654	.10852	.17812	.17279	.16066	.14013	.13420
BS	.04409	.03597	.03109	.02226	.01355	.13968	.12079	.10398	.11028	.10466	.19911	.19001	.17469	.14760	.13896
V	.04644	.03688	.03191	.02313	.01433	.11165	.10678	.09676	.10700	.10376	.23844	.21017	.18493	.15158	.13979
VS	.04559	.03695	.03205	.02303	.01423	.09311	.09693	.08886	.10323	.10154	.27311	.22531	.19553	.15668	.14244
T	.04468	.03659	.03187	.02291	.01440	.11250	.10927	.09877	.10860	.10534	.23512	.20599	.18179	.14965	.13784
TS	.04450	.03671	.03196	.02306	.01418	.09496	.09847	.09011	.10451	.10226	.26707	.22211	.19345	.15506	.14146
36															
H	.03647	.02160	.01751	.01004	.00713	.06255	.01487	.01191	.02045	.02258	.29595	.17609	.13470	.10377	.09045
M	.03647	.02160	.01751	.01004	.00713	.17642	.16182	.15245	.14225	.13738	.00000	.00000	.00000	.00000	.00000
MS	.03653	.02171	.01764	.01004	.00704	.07382	.07269	.07322	.07439	.07236	.10267	.08195	.07000	.05811	.05481
B	.03431	.01975	.01648	.00908	.00622	.05133	.03446	.03092	.03271	.03146	.10926	.10010	.08965	.08140	.07730
BS	.03475	.01979	.01651	.00917	.00630	.04301	.02764	.02556	.02987	.02955	.12043	.10951	.09663	.08578	.07975
V	.03925	.02168	.01739	.00996	.00709	.03259	.02234	.02234	.02831	.02910	.14387	.12131	.10302	.08804	.08054
VS	.03843	.02177	.01755	.00994	.00706	.02768	.01905	.01930	.02668	.02794	.16534	.13036	.10848	.09108	.08156
T	.03785	.02137	.01739	.00985	.00715	.03290	.02329	.02314	.02896	.02982	.14154	.11874	.10108	.08703	.07948
TS	.03758	.02149	.01746	.00994	.00705	.02818	.01952	.01979	.02718	.02821	.16102	.12821	.10718	.09030	.08126
60															
H	.03384	.01826	.01268	.00455	.00419	.07574	.01375	.00747	.01054	.01106	.23953	.13812	.10535	.08404	.06923
M	.03384	.01826	.01268	.00455	.00419	.12617	.11644	.10926	.10937	.10522	.00000	.00000	.00000	.00000	.00000
MS	.03404	.01839	.01276	.00458	.00405	.04221	.04558	.04652	.05211	.05046	.08596	.06501	.05492	.04766	.04422
B	.02923	.01640	.01119	.00375	.00330	.02702	.01829	.01611	.01898	.01746	.09035	.07797	.06942	.06555	.05926
BS	.02981	.01648	.01127	.00382	.00336	.02212	.01427	.01306	.01702	.01609	.09872	.08525	.07492	.06884	.06066
V	.03719	.01845	.01259	.00445	.00414	.01753	.01150	.01105	.01601	.01575	.11641	.09420	.07976	.07078	.06158
VS	.03633	.01855	.01274	.00445	.00413	.01709	.01020	.00977	.01489	.01481	.13384	.10142	.08417	.07310	.06235
T	.03568	.01814	.01256	.00439	.00421	.01767	.01193	.01160	.01640	.01619	.11491	.09214	.07823	.07001	.06090
TS	.03539	.01828	.01263	.00446	.00413	.01713	.01034	.01000	.01519	.01500	.13061	.09974	.08313	.07254	.06208

測では $N_a = 12$ の場合に限って Vasicek（all）モデルが明らかに ML モデルよりも優れていたから，この場合にはこの調整が有効となることが想像できる．実際，$\{N_a, N_p\} = \{12, 12\}$ の場合に Vasicek（all）モデルの修正値の平均を ML モデルの修正値の平均に調整すると，MSE の平均は $0.10591 + 0.29380 + 0.30942 = 0.70913$ から $0.08744 + 0.29380 + 0.30942 = 0.71844$ に 3% 弱改善する．そもそも，ML モデルは推定値全体の平均，ひいてはベータ値の母数の全体の平均が 1 と予め仮定していたモデルであるから，ML モデルがとくに予測精度が高かったことも考えれば，ベータ値の予測においてはそのように調整することが，すなわち修正値全体の平均が 1

表 12　ベータ値の予測における MSE の各成分の集計

$N_p \backslash N_a$	平均の誤差の平均 (10^0)					標準偏差の誤差の平均 (10^0)					相関の誤差の平均 (10^0)				
	12	36	60	120	180	12	36	60	120	180	12	36	60	120	180
12															
H	.11596	.08392	.09488	.07450	.06048	.06445	.07769	.10983	.16321	.20289	.93235	.52575	.43906	.35283	.32973
M	.11596	.08392	.09488	.07450	.06048	.63100	.58626	.57742	.58575	.60473	.00000	.00000	.00000	.00000	.00000
MS	.11746	.08592	.09722	.07497	.06021	.25936	.27643	.28699	.31485	.34742	.37996	.28325	.25493	.22707	.21506
B	.10591	.08155	.09245	.07410	.05954	.29380	.21034	.20372	.22059	.25007	.30942	.31495	.30618	.28537	.27762
BS	.10878	.08250	.09364	.07411	.05962	.19174	.18821	.19237	.21599	.24800	.44211	.34117	.31636	.28723	.27696
ML	.08744	.06854	.07656	.07266	.06322	.11041	.18608	.21954	.27096	.30937	.62157	.35050	.29271	.23522	.21982
36															
H	.08151	.06798	.06115	.02583	.01151	.10636	.00950	.01049	.03255	.04941	.55141	.29859	.24097	.19499	.17697
M	.08151	.06798	.06115	.02583	.01151	.27225	.26391	.26590	.28356	.29085	.00000	.00000	.00000	.00000	.00000
MS	.08350	.07134	.06416	.02638	.01120	.05862	.07550	.08580	.11259	.12941	.23034	.16594	.14676	.12920	.11924
B	.07198	.06497	.05922	.02546	.01076	.08288	.04373	.04215	.05905	.07345	.17541	.17420	.16547	.15498	.14732
BS	.07513	.06707	.06079	.02559	.01081	.03393	.03411	.03751	.05706	.07259	.25809	.18742	.16842	.15389	.14481
ML	.04791	.04036	.03840	.02196	.01290	.02077	.03367	.05023	.08367	.10349	.36761	.19906	.16065	.12999	.11798
60															
H	.10352	.07761	.05391	.01711	.00536	.15045	.01454	.00724	.02042	.03046	.47441	.25170	.20188	.16241	.13771
M	.10352	.07761	.05391	.01711	.00536	.20720	.20560	.21111	.22886	.23046	.00000	.00000	.00000	.00000	.00000
MS	.10739	.08218	.05703	.01750	.00514	.03737	.04984	.06013	.08568	.09730	.19132	.13506	.11923	.10333	.08953
B	.09233	.07453	.05190	.01669	.00466	.05806	.02524	.02464	.04038	.04965	.14663	.14339	.13610	.12680	.11360
BS	.09722	.07719	.05360	.01683	.00473	.02311	.01877	.02156	.03900	.04909	.21626	.15268	.13751	.12511	.10989
ML	.05367	.04100	.03106	.01355	.00613	.02841	.01865	.03058	.05843	.07252	.31628	.16780	.13459	.10827	.09181

となるように調整することが有用であるかもしれない．

Elton and Gruber（1978）や Eun and Resnick（1984）では，修正値の平均を調整した場合の結果も検証していたが，本稿での結果をみる限り，とくに N_a が大きい場合は，その効果が非常に小さいことがわかる．

標準偏差の誤差や相関の誤差は，とくに分散の予測の場合に，平均の誤差と比較してその水準が大きく，また，その大きさもモデルによって大きく異なっている．しかし，修正値の標本標準偏差を他のモデルの修正値の標本標準偏差と一致させるような操作は，たとえそれによって標準偏差の誤差が小さくなったとしても，必ずしも MSE 自体を改善するとは限らない．標本標準偏差の調整として，修正値の平均偏差に適当な定数 a を乗じ[26]，そこに元の修正値の平均を加えると，MSE は元の修正値の標本分散の (a^2-1) 倍だけ大きくなり，元の修正値と事後の値の標本共分散の $2(a-1)$ 倍だけ小さくなる．つまり，たとえば，$a>1$ であって標本共分散が正であるならば，標準偏差の調整は有効でない．この点を踏まえると，標準偏差の調整が MSE の平均の縮小に対して効果があるか否かを簡単に示すのは容易でないし，表から読み取ることもできない．ただし，平均の予測の場合に限っては，最も優れるモデルが（そもそも修正値にばらつきのない）Mean（all）モデルであるから修正値の標本標準偏差は調整のしようがないといえる．

なお，表 10 をみると，平均の予測の結果において Vasicek（all）モデルの相関の誤差の平均が極めて 0 に近いことが読み取れる．これは多くの時点で $\mathrm{Var}(\theta) \to +0$ と計算されていることの証左といえる．つまり，$\hat{x}_i$ の標準偏差に対して，その標準誤差の水準が極めて大きいことを意味する．

4.4　MSE の代表値の差に関する検定結果

以下では，MSE の代表値の差が有意であるか否かを検定した結果を示す．本稿では MSE に関してモデル間の対比較を行い，その代表値の差が有意水準 1% で有意であるか否かを検証した．その結果を十分位数の比較とともに表 13，表 14，表 15 にまとめた．いずれの表も MSE の十分位数のうちで表頭のモデルが表側のモデルよりも大きいものの数を示しており，背景をグレーで塗り潰しているものは代表値の差が有意でないモデル間である．背景が塗り潰されておらず，数字が少なくとも 5 以上であれば表頭のモデルよりも表側のモデルのほうが優れるといえ，とくに 8 や 9 であればそのことが明らかにいえる．5 や 6 であれば有意であることは確かだが，どちらが優れると結論付けるにはやや説得力に欠けるだろう．なお，この数字は対角線に対して対称の位置にあるものの和が必ずしも 9 になるとは限らない．値が同じものはどちらにも計上しないためである．

26　一方のモデルの修正値の標準偏差を他方のモデルの標準偏差に一致させるための a の値としては分散比の正の平方根が適当である．

表 13　平均の予測における MSE の差の検定結果

$N_a \backslash N_p$	**H**			**M**			**MS**			**B**			**BS**		
	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60
H															
12				0	0	0	0	0	0	0	0	0	0	0	0
36				0	0	0	0	0	0	0	0	0	0	0	0
60				0	0	0	0	0	0	0	0	0	0	0	0
120				0	0	0	0	0	0	0	0	0	0	0	0
180				0	0	0	0	0	0	0	0	0	0	0	0
M															
12	9	9	9				9	9	9	2	7	5	9	9	9
36	9	9	9				9	9	9	0	0	0	9	9	9
60	9	9	9				7	9	9	0	0	0	7	9	9
120	9	9	9				6	9	9	0	0	0	6	9	9
180	9	9	9				7	6	7	0	0	1	7	6	7
MS															
12	9	9	9	0	0	0				0	0	0	4	6	6
36	9	9	9	0	0	0				0	0	0	5	4	2
60	9	9	9	2	0	0				2	0	0	3	1	1
120	9	9	9	3	0	0				3	0	0	2	1	2
180	9	9	9	2	3	2				2	3	2	0	0	0
B															
12	9	9	9	2	2	2	9	9	9				9	9	9
36	9	9	9	2	0	0	9	9	9				9	9	9
60	9	9	9	0	0	0	7	9	9				7	9	9
120	9	9	9	0	0	0	6	9	9				6	9	9
180	9	9	9	0	0	0	7	6	7				7	6	7
BS															
12	9	9	9	0	0	0	5	3	2	0	0	0			
36	9	9	9	0	0	0	4	3	2	0	0	0			
60	9	9	9	2	0	0	0	2	2	2	0	0			
120	9	9	9	3	0	0	0	0	1	3	0	0			
180	9	9	9	2	3	2	0	0	0	2	3	2			

まず，平均の予測における結果（表 13）をみると，差が有意である場合に優劣が明瞭であることがわかる．つまり，背景が塗り潰されていない部分では 9 あるいは 0 がほとんどである．とくに Historical の列はすべて有意かつ 9 であり，明らかに他のモデルの修正効果があるといえる．

一方で，Mean（all）モデルと Vasicek（all）モデル，Mean（sector）モデルと Vasicek（sector）モデルの比較では N_a や N_p によらずそれぞれ互いに差が有意でない．つまり，どちらを適用しても予測精度が変わるとはいえない．いずれの 2 組でも明らかに前者のほうが計算が簡単であるから，この場合は Mean モデルのほうが実用的である．また，Mean（all）モデルと Mean（sector）モデル，Vasicek（all）モデルと Vasicek（sector）モデルでは，$N_a = 180$ である場合を除いて差が有意である．また，いずれも all モデルのほうが優れている．計算も all モデルのほうが sector モデルより単純であることから，平均の予測の場合は業種を区別することなくすべてを平均化する Mean（all）モデルが適していると結論付けられる．これらの結果は，ここまで述べた検証結果と一切相違しない．

なお，表 13 において，Mean モデルと Vasicek モデルとでは対角線に対して対称の位置にある成分の和が 9 未満であることが多い．これは多くの，あるいはすべての時点で $\mathrm{Var}(\theta) \to +0$ と推定されることで，Vasicek モデルの修正値が Mean モデルの修正値と完全に一致しているためである．ゆえに，これらのモデル間の差が有意でないことは当然でもある．

次に，分散の予測における結果（表 14）をみると，とくに Mean モデルと比較した結果に関して，全体的に 9 や 8 が少ないことが目立つ．つまり，有意差が認められても，いずれを優れると認めるかが明らかでない場合が多い．とくに Variable モデルや Theoretical モデルは N_a が小さいとき，Mean モデルとの差が有意であっても，いずれが優れているとは判断しがたい．

MSE の集計結果において最も優れていた Vasicek（all）モデルは，他のモデルとの差が有意である場合，相対的に小さい MSE の十分位数の数は概ね 7 以上であって，この結果からも分散の予測では Vasicek（all）モデルが最も予測精度が高いといえる．なお，Vasicek（sector）モデルと Theoretical（all）モデルは，Vasicek（all）モデルと比較して MSE の差が常に有意でないが，これらのモデルは Vasicek（all）モデルよりも計算が複雑である．

all モデルと sector モデルとを比較すると，Mean モデルと Vasicek モデルでは差が有意でないが，Variable モデルと Theoretical モデルでは $N_a \leq 60$ のとき，all モデルのほうが有意に優れる．このことは MSE の集計結果（表 6）において all モデルよりも sector モデルのほうが MSE の中央値や平均が大きかったことと整合的であり，分散の分布が業種ごとに区別しづらいことを示している．

分散の予測では Variable モデルと Theoretical モデルが定義でき，検証したが，これらの結果はあまりよくない．これらのモデルはいずれも Vasicek モデルよりも仮定が現実的であるにもかかわらず，Vasicek モデルに対して有意に優れることはなく，むしろ Theoretical（all）モデルを除いては明らかに Vasicek（all）モデルに劣る．ただし，Historical モデルに対しては Variable モデルも Theoretical モデルも有意に優れているため，修正の効果が全くないわけではない．Variable モデルと Theoretical モデルは互いに差が有意でないが，これは標本数 N_a の増大にともなって Theoretical モデルで仮定していた分布（確率密度関数は g_T）が正規分布に近づくためである．この結果をみると，$N_a = 12$ 程度でも正規分布を仮定する場合と比較して有意な差が得られないことがわかる．なお，検定結果を無視して十分位数の大小関係だけを比較すると，仮定がより現実的な Theoretical モデルのほうが若干優れるようである．

最後に，ベータ値の予測における結果（表 15）をみると，$N_a \geq 36$ のときに同等の性能であった Vasicek

表 14　分散の予測における MSE の差の検定結果

$N_a \backslash N_p$	H			M			MS			B			BS			V			VS			T			TS		
	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60
H																											
12				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36				2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60				4	4	5	4	3	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
120				5	7	9	4	6	7	2	1	0	2	0	0	2	0	0	2	0	0	2	0	0	2	0	0
180				8	8	9	7	8	9	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0	2	1	0
M																											
12	9	9	9				5	4	6	0	3	3	0	3	3	4	5	4	7	8	9	4	5	4	6	6	9
36	7	8	9				2	2	2	0	3	2	0	3	2	3	4	3	4	4	3	3	4	2	3	4	3
60	5	5	4				3	1	1	0	1	1	0	2	1	3	3	2	4	3	2	3	3	2	4	3	2
120	4	2	0				1	1	0	0	0	0	0	1	0	2	2	0	3	2	0	0	2	0	3	2	0
180	1	1	0				0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
MS																											
12	9	9	9	4	5	3				0	3	3	0	3	3	4	5	5	7	8	9	3	4	5	6	7	9
36	8	9	9	7	7	7				0	2	1	1	2	2	4	4	4	4	4	5	3	3	3	4	4	5
60	5	6	9	6	8	8				0	1	1	0	2	1	4	3	2	5	3	2	3	3	2	5	3	2
120	5	3	2	8	8	9				0	0	0	1	0	0	2	2	0	3	2	0	2	2	0	3	2	0
180	2	1	0	9	9	9				0	0	0	0	0	0	1	0	0	2	0	0	1	0	0	1	0	0
B																											
12	9	9	9	9	6	6	9	6	6				7	8	7	8	9	8	8	9	9	8	9	9	8	9	9
36	9	9	9	9	6	7	9	7	8				5	8	9	8	8	9	8	9	9	8	8	9	8	8	9
60	8	9	9	9	8	8	9	8	8				8	7	9	7	7	9	7	8	9	7	7	9	7	8	9
120	7	8	9	9	9	9	9	9	9				7	7	7	6	8	7	7	8	7	6	8	7	7	8	7
180	7	8	9	9	9	9	9	9	9				7	6	5	7	6	9	6	5	8	6	6	8	6	5	8
BS																											
12	9	9	9	9	6	6	9	6	6	2	1	2				8	7	7	8	9	9	8	8	7	8	9	9
36	9	9	9	9	6	7	8	7	7	4	1	0				8	8	9	8	9	9	8	7	9	8	9	9
60	8	9	9	9	7	8	9	7	8	1	2	0				7	7	9	7	9	9	7	6	9	7	8	9
120	7	9	9	9	8	9	8	9	9	2	2	2				8	7	6	7	8	7	7	5	6	7	8	7
180	7	8	9	9	9	9	9	9	9	2	3	4				5	8	9	5	5	8	5	7	9	5	6	8
V																											
12	9	9	9	5	4	5	5	4	4	1	0	1	1	2	2				9	9	9	0	2	3	8	9	9
36	9	9	9	6	5	6	5	5	5	1	1	0	1	1	0				8	9	9	1	0	0	7	8	9
60	9	9	9	6	6	7	5	6	7	2	2	0	2	2	0				8	8	9	2	1	0	8	8	8
120	7	9	9	7	7	9	7	7	9	3	1	2	1	2	3				7	8	7	2	1	0	7	8	7
180	7	8	9	9	9	9	8	9	9	2	3	0	4	1	0				5	3	4	2	2	1	6	3	2
VS																											
12	9	9	9	2	1	0	2	1	0	1	0	0	1	0	0	0	0	0				0	0	0	0	0	0
36	9	9	9	5	5	6	5	5	4	1	0	0	1	0	0	1	0	0				1	0	0	1	1	0
60	9	9	9	5	6	7	4	6	7	2	1	0	2	0	0	1	1	0				1	0	0	2	1	0
120	7	9	9	6	7	9	6	7	9	2	1	2	2	1	2	2	1	2				2	1	1	4	2	2
180	7	8	9	8	9	9	7	9	9	3	4	1	4	4	1	4	6	5				1	6	4	4	5	5
T																											
12	9	9	9	5	4	5	6	5	4	1	0	0	1	1	2	9	7	6	9	9	9				9	9	9
36	9	9	9	6	5	7	6	6	6	1	1	0	1	2	0	8	9	9	8	9	9				9	9	9
60	9	9	9	6	6	7	6	6	7	2	2	0	2	3	0	7	8	9	8	9	9				8	9	9
120	7	9	9	9	7	9	7	7	9	3	1	2	2	4	3	7	8	9	7	8	8				7	8	8
180	7	8	9	9	9	9	8	9	9	3	3	1	4	2	0	7	7	8	8	3	5				7	3	4
TS																											
12	9	9	9	3	3	0	3	2	0	1	0	0	1	0	0	1	0	0	9	9	9	0	0	0			
36	9	9	9	6	5	6	5	5	4	1	1	0	1	0	0	2	1	0	8	8	9	0	0	0			
60	9	9	9	5	6	7	4	6	7	2	1	0	2	1	0	1	1	1	7	8	9	1	0	0			
120	7	9	9	6	7	9	6	7	9	2	1	2	2	1	2	2	1	2	5	7	7	2	1	1			
180	7	8	9	8	9	9	8	9	9	3	4	1	4	3	1	3	6	7	5	4	4	2	6	5			

(sector）モデルと ML モデルとの差が $N_a \leq 120$（かつ $N_a \geq 36$）であるときに有意でないことが読み取れる．さらに，Vasicek（all）モデルも Vasicek（sector）モデルや ML モデルと互いに差が有意でない場合が多い．$N_a \geq 36$ であるならば Vasicek（all）モデルと ML モデルは常に有意でない．しかし，そもそも，ML モデルの計算は簡単であり，MSE の集計結果（表 7）をみても Vascek（all）モデルよりも ML モデルのほうが優れていたから，

表 15　ベータ値の予測における MSE の差の検定結果

$N_a \backslash N_p$	**H**			**M**			**MS**			**B**			**BS**			**ML**		
	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60	12	36	60
H																		
12				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36				2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
60				8	8	7	4	1	0	0	0	0	0	0	0	0	0	0
120				9	9	9	8	7	7	1	1	0	1	0	0	2	1	0
180				9	9	9	9	9	9	2	2	2	1	0	0	3	3	4
M																		
12	9	9	9				7	8	6	0	0	1	2	6	6	9	9	9
36	7	9	8				0	2	1	0	0	1	0	0	1	0	0	0
60	1	1	2				0	0	1	0	0	1	0	0	1	0	0	0
120	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0
180	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0
MS																		
12	9	9	9	2	1	3				0	0	0	1	3	2	8	9	9
36	9	9	9	9	7	8				0	0	0	0	0	0	0	0	0
60	5	8	9	9	9	8				0	0	0	0	0	0	0	0	0
120	1	2	2	9	9	9				0	0	0	0	0	0	0	0	0
180	0	0	0	9	9	9				0	0	0	0	0	0	0	0	0
B																		
12	9	9	9	9	9	8	9	9	9				9	9	9	9	9	9
36	9	9	9	9	9	8	9	9	9				7	6	3	3	6	5
60	9	9	9	9	9	8	9	9	9				3	1	1	1	3	4
120	8	8	9	9	9	9	9	9	9				1	0	1	1	2	1
180	7	7	7	9	9	9	9	9	9				3	0	0	7	6	7
BS																		
12	9	9	9	7	3	3	8	6	7	0	0	0				9	9	9
36	9	9	9	9	9	8	9	9	9	2	3	6				3	6	5
60	9	9	9	9	9	8	9	9	9	6	8	8				1	5	7
120	8	9	9	9	9	9	9	9	9	8	9	8				3	4	4
180	8	9	9	9	9	9	9	9	9	6	9	9				8	8	8
ML																		
12	9	9	9	0	0	0	1	0	0	0	0	0	0	0	0			
36	9	9	9	9	9	9	9	9	9	6	3	4	6	3	4			
60	9	9	9	9	9	9	9	9	9	8	6	5	8	4	2			
120	7	8	9	9	9	9	9	9	9	8	7	8	6	5	5			
180	6	6	5	9	9	9	9	9	9	2	3	2	1	1	1			

$N_a \geq 36$ であるときは Vasicek（all）モデルよりは ML モデルのほうが適している．また，Vasicek（sector）モデルと比較しても，ML モデルのほうが計算が簡単であるから，これらの間に有意な差が認められないならば，ML モデルがもっとも実用的である．つまり，利用できるデータが過去 36 ヶ月分から 120 ヶ月分程度であるならば，計算の容易さの点において ML モデルが最適である．なお，ML モデルと Vasicek モデルとの差が有意でないことから，ML モデルの仮定，すなわち $\overline{x} = 1$ かつすべての i について $\mathrm{Var}(\theta) : \mathrm{Var}(\tilde{X}_i | \theta_i = \hat{x}_i) = 2 : 1$ という仮定が非現実的ではないことがうかがえる．

一方で，$N_a = 180$ の場合は Vasicek（sector）モデルと ML モデルとの差が有意であって Vasicek（sector）モデルのほうが明らかに優れている．ゆえに，過去の多くのデータが利用できる場合は予測精度の点において Vasicek（sector）モデルが適するといえる．

MSE の集計結果でもみたように，$N_a = 12$ の場合は Vasicek（all）モデルが明らかに優れており，ML モデルや Vasicek（sector）モデルとの差も有意である．$N_a \geq 36$ であるならば ML モデルでも同等の予測精度が得られたが，利用できるデータが極めて少なく過去 12 ヶ月分程度である場合は，Vasicek（all）モデルを用いることでより予測精度を高めることができるといえる．

Vasicek モデルや ML モデルは Historical モデルに対

して多くの場合で有意に優れているが，$N_a = 180$ の場合に Vasicek（all）モデルと ML モデルは Historical モデルと有意な差がない．つまり，利用できる過去のデータが極めて多いならば，Vasicek（all）モデルや ML モデルはベータ値の予測における修正モデルとして有効であるとはいえなくなる．ただし，$N_a = 180$ のときに適していると結論づけた Vasicek（sector）モデルだけは，$N_a = 180$ の場合でも Historical モデルに対して有意に優れる．なお，Mean モデルはおおむね $N_a \leq 36$ 程度のときに Historical モデルに対して有意に優れ，修正の効果が認められるが，それ以外の場合は Historical モデルに劣り，修正が有効とはいえない．

4.5　業種ごとの MSE の集計

最後に，業種ごとに最適なモデルが異なる可能性を考え，MSE を業種ごとに集計した結果を表 16 に示す．とくに all モデルと sector モデルとを比較したときにどちらが優れているかは業種によって異なることが想像に難くない．ただし，表が膨大になってしまうため，モデルは Historical モデル，Mean モデル，Vasicek モデルに限り，期間の長さも $\{N_a, N_p\} = \{60, 60\}$ の場合に限る．また，集計する業種はすべての基準時点においてそこに属する銘柄数が 25 を下回らない業種に限り，それには建設業，食料品，繊維製品，化学，機械，電気機器，輸送用機器，卸売業，銀行業の 9 業種が該当する．

表をみると，MSE の中央値や平均は，平均の予測の場合では建設業や繊維製品，機械が，分散の予測の場合では建設業や機械，卸売業が，ベータ値の予測の場合では機械や電気機器，輸送用機器がとくに大きいことがわかる．一方で，食料品や化学における MSE の中央値や平均はいずれの統計量の予測についても他の業種に比べて著しく小さく，分散の予測の場合に至っては食料品の MSE の中央値が機械の値の 3 分の 1 程度になっている．つまり，MSE の水準自体が業種によって明らかに異なる．ゆえに，すべての業種を総合して MSE を計算すると，（表中の業種だけで述べれば）建設業や機械，あるいは電気機器等の影響を強く受け，食料品や科学における修正効果の有無は MSE の値に表れ難くなっていると考えられる．

MSE の水準の他にも，all モデルと sector モデルのいずれが優れるかという点も業種によって異なっている．たとえば，食料品や化学等は all モデルよりも sector モデルのほうが優れているようであるが，一方で建設業や卸売業等は all モデルのほうが優れるようである．上の検証で平均の予測ならば Mean（all）モデルが，分散の予測ならば Vasicek（all）モデルが，ベータ値の予測ならば Vasicek（sector）モデル（あるいは ML モデル）が総合的にみて優れていると結論づけたが，それらのうちの all モデルのほうが優れるのか sector モデルのほうが優れるのかは，その業種に依存するところが大きいようである．たとえば，$\{N_a, N_p\} = \{60, 60\}$ の場合，電気機

表 16　$\{N_a, N_p\} = \{60, 60\}$ の場合の業種ごとの MSE

	平均 (10^{-2})		分散 (10^{-3})		ベータ値 (10^{0})	
	中央値	平均	中央値	平均	中央値	平均
建設業						
H	.05730	.05219	.13862	.12872	.26764	.20601
M	.03230	.02903	.14721	.15648	.24027	.19012
M-S	.04582	.03748	.14729	.15443	.24891	.21457
B	.03230	.02903	.11158	.10076	.21724	.16367
B-S	.04582	.03748	.11446	.10006	.21985	.17163
食料品						
H	.02369	.01673	.04724	.04751	.18320	.14499
M	.02119	.01284	.07841	.07478	.28817	.27873
M-S	.01761	.01026	.03826	.04081	.16705	.13584
B	.02119	.01284	.03893	.03781	.16124	.12591
B-S	.01761	.01026	.03385	.03287	.14404	.12517
繊維製品						
H	.05128	.04625	.10992	.10231	.25530	.19759
M	.03238	.02511	.10785	.10215	.21332	.18545
M-S	.04106	.03428	.11346	.10719	.23064	.19227
B	.03238	.02511	.09018	.08326	.19769	.15902
B-S	.04106	.03428	.09242	.08258	.20174	.15350
化学						
H	.03038	.02409	.05594	.05375	.21208	.14862
M	.02582	.01779	.05829	.05480	.18889	.13518
M-S	.02219	.01569	.04891	.05017	.18948	.13896
B	.02582	.01779	.04866	.04547	.17762	.11368
B-S	.02219	.01569	.04462	.04255	.17101	.10859
機械						
H	.04920	.04310	.19400	.18025	.33192	.26363
M	.03364	.02845	.15403	.11437	.28223	.24359
M-S	.03652	.02837	.15708	.11787	.27866	.19050
B	.03364	.02845	.13974	.10941	.26633	.19371
B-S	.03652	.02837	.14656	.12870	.26941	.17595
電気機器						
H	.03837	.03439	.11190	.11327	.29719	.24262
M	.03060	.02661	.09403	.08805	.23451	.23925
M-S	.02446	.01895	.09434	.10190	.25491	.18510
B	.03060	.02661	.09117	.09759	.22789	.18481
B-S	.02446	.01895	.08896	.09456	.24319	.17755
輸送用機器						
H	.04630	.04240	.09027	.08922	.28563	.22557
M	.03191	.02807	.07437	.06967	.20975	.19011
M-S	.03388	.02680	.08007	.07241	.24383	.18860
B	.03191	.02807	.07524	.07013	.23204	.19243
B-S	.03388	.02680	.07523	.07029	.23712	.18440
卸売業						
H	.04153	.03701	.13093	.10226	.26096	.22128
M	.02879	.02272	.13010	.11063	.23827	.21545
M-S	.03111	.02566	.12970	.12453	.24295	.22004
B	.02879	.02272	.10860	.09151	.20438	.15806
B-S	.03111	.02566	.10865	.09158	.20352	.15895
銀行業						
H	.02460	.01333	.02101	.01870	.13449	.11565
M	.02413	.01088	.08464	.07772	.27476	.20305
M-S	.02083	.00911	.01525	.01426	.11707	.09647
B	.02413	.01088	.02013	.01855	.11256	.09190
B-S	.02083	.00911	.01550	.01498	.11266	.09141

器などでは平均の予測に Mean（sector）モデルを用いることで，（業種内の）MSE の平均を，Mean（all）モデルを用いる場合の 7 割程度に縮小することができる．業種ごとに all モデルと sector モデルのいずれを用いるかを考慮して予測すれば，総合的な MSE 自体も相当に縮小できることが予想される．

MSE の水準や all モデルと sector モデルの優劣は業種によって異なるものの，（all モデルと sector モデルとを区別しない）モデルの優劣はどの業種でもだいたい同じである．実際，MSE の中央値や平均は，平均の予測では，業種を問わず Mean モデルや Vasicek モデルが Historical モデルに劣ることがなく[27]，分散やベータ値の予測ではほとんどの業種で Vasicek モデルが Historical モデルや Mean モデルよりも優れている．

5.　おわりに

本稿では株価変化率の（期待収益率の推定量としての）平均，分散，ベータ値の予測におけるベイズ修正の効果を，最大で 2282 銘柄を対象にして検証した．その修正モデルとして，推定値をすべて平均化する Mean モデル，推定値とその全体の平均とをそれぞれの分散によって重み付けした加重平均で修正する Vasicek モデル，Vasicek モデルから推定値が従う分布に設けた仮定を緩和した Variable モデルや Theoretical モデル等を用い，その予測能力を比較した[28]．また，各統計量の母数の分布が対象銘柄すべてにおいて同一であると仮定する all モデルと，業種によって異なると仮定する sector モデルとに分けて検証した．さらに，事前期間と事後期間は，それぞれの長さを $\{12, 36, 60, 120, 180\}$ か月間，$\{12, 36, 60\}$ か月間と複数設定し，計 15 のパターンのそれぞれで検証した．

結論として，平均の予測では事前期間や事後期間の長さを問わず，いずれのモデルでも（修正を行わないモデルである）Historical モデルの予測精度を向上させる効果が認められた．中でも，平均の予測には Mean（all）モデルが適していることを示した．株価変化率の平均の予測に Mean（all）モデルを用いるならば，平均・分散モデルに従う最適ポートフォリオが必ず最小分散ポートフォリオに一致することになる．このことは，近年，最小分散ポートフォリオの効率性が注目されている事実と整合的である（石部他（2009）など）．平均の予測において Mean（all）モデルが優れる理由として，推定値と事後の値との共分散が負であることを挙げたが，このことを踏まえれば，推定値と修正値が逆相関するような修正モデルを用いれば，Mean（all）モデルよりも予測能力が高まることが期待できる．

分散の予測では，Vasicek（all）モデルが最も予測精度が高く，Historical モデルによる予測誤差を縮小できることを示した．一方で，事後期間が十分に長いとき，Mean モデルはむしろ予測精度を悪化させることを示した．分散の予測の場合，Variable モデルや Theoretical モデルというモデルも定義して検証したが，これらは修正の効果はあるものの，Vasicek モデルには及ばない結果であった．ベイズ修正によって分散の予測精度が高められることは，相関係数や最小分散ポートフォリオの予測精度向上にも貢献できると考えられる．分散の予測に Mean モデルが利用できるのであれば，最小分散ポートフォリオは等金額ポートフォリオに一致するが，本稿の結果ではすべての銘柄で分散を同一視するのは難しいといえる．

ベータ値の予測では，Mean モデル以外のすべてのモデルが予測精度の面で Historical モデルを改善でき，Mean モデルでも事前期間がある程度短い場合は Historical モデルよりも優れる結果であった．中でも，ベータ値の予測には Vasicek（sector）モデルあるいは ML モデルが適していることを示した．ベータ値の予測に Vasicek モデルが有効であるという事実は，Klemkosky and Martin（1975）や金崎（1987）など，過去にも多くの研究で示されており，本稿の結論もそれらと同様である．

また，事前期間や事後期間ごとに予測精度の水準を比較すると，いずれの統計量の予測においても基本的にはそれらが長いほど予測精度が高まることを確認した．ただし，事後期間が 12 か月程度と極端に短い場合は，あまり遠い過去のデータは用いるべきでなく，おおむね平均や分散の予測ならば 120 か月間程度，ベータ値の予測ならば 36 か月間程度が妥当であるといえた．事後期間の長さが異なるものの，ベータ値の予測に関するこの結論は，安藤・久保田（2003）の研究結果と似ているといえる．

本稿では月次データを用いた検証を行ったが，週次や日次のデータを用いることによっても予測精度が向上することが期待できる．また，業種ごとに予測するモデルを sector モデルとして検証したが，Karolyi（1992）では（ベータ値の予測においては）業種に加えて時価総額による銘柄のグループ化も行うことで予測精度が更に向上する結果となっている．時価総額のデータが得られな

27　上述の通り，平均の予測では Mean モデルと Vasicek モデルとで修正値がほとんど変わらないため，優劣がつけられない．
28　Variable モデルおよび Theoretical モデルの検証は分散の予測の場合に限る．

かったため，このモデルが検証できなかったが，本稿のようなデータでも同様の結果が得られるか，そしてベータ値の予測以外においても時価総額が銘柄をグループ化するための重要な要素になり得るかという点も興味深い．これらについては今後の課題としたい．

参考文献

青木繁伸 (2009) “フリードマン検定（plus 多重比較）”〈http://aoki2.si.gunma-u.ac.jp/R/friedman.html〉（参照日 2015-10-13）.

安達智彦・池田昌幸（2016 近刊）『多期間投資の理論と実践（仮称）』東京大学出版会 第 1，6 章.

安藤由美・久保田敬一 (2003)「マーケットモデルを用いたベータ推定方法についての考察：I」『武蔵大学論集』第 51 巻第 1 号 107–125 頁.

石部真人・角田康夫・坂巻敏史 (2009)「最小分散ポートフォリオのリスクとリターン」『行動経済学』第 2 巻 88–92 頁.

金崎芳輔 (1987)「修正ベータの有効性の検証」『証券アナリストジャーナル』第 25 巻第 11 号 62–75 頁.

吉原正善 (1990)「日本市場における β 値の実証研究」『証券アナリストジャーナル』第 28 巻第 8 号 34–50 頁.

芳野哲也 (2011)「株式のベータ値の推定と修正について」武蔵大学大学院経済学研究科修士論文.

Blume, Marchal, (1971), “On The Assessment of Risk,” *Journal of Finance*, Vol.26, No.1, pp.1–10.

Elton, Edwin J. and Gruber, Martin J., (1978), “Are Betas Best?” *The Journal of Finance*, Vol.33, No.5, pp.1375–1384.

Eubank, Arhur A. Jr. and Zumwalt, J. Kenton, (1979), “An Analysis of the Forecast Error Impact of Alternative Beta Adjustment Techniques and Risk Classes,” *The Journal of Finance*, Vol.34, No.3, pp.761–776.

Eun, Cheol S. and Resnick, Bruce G., (1984), “Estimating the Correlation Structure of International Share Prices,” *The Journal of Finance*, Vol.39, No.5, pp.1311–1324.

James, W. and Stein, Charles, (1961), “Estimation with Quadratic Loss,” *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol.1, pp.361–379.

Karolyi, G. Andrew, (1992), “Predicting Risk: Some New Generalizations,” *Management Science*, Vol.38, No.1, pp.57–74.

Klemkosky, Robert C. and Martin, John D., (1975), “The Adjustment of Beta Forecasts,” *The Journal of Finance*, Vol.30, No.4, pp.1123–1128.

Merton, Robert C., (1980), “On Estimating the Expected Return on the Market: An Exploratory Investigation,” *The Journal of Finance*, Vol.8, No.4, pp.323–361.

Pastor, Lubos, (2001), “A Model Weighting Game in Estimating Expected Returns,” *Financial Times, Mastering Investment*, May 21.

Vasicek, Oldrich A., (1973), “A Note on Using Cross-Sectional Information in Baysian Estimation of Security Betas,” *The Journal of Finance*, Vol.28, No.5, pp.1233–1239.