

【調 査】

実現ボラティリティ (Realized Volatility)

—— サーベイと日経 225 株価指数への応用* ——

渡部敏明・中島上智

近年の計量ファイナンスの研究では、資産価格のボラティリティの推定量として、日中の高頻度データから計算される実現ボラティリティ (Realized Volatility; RV) を用いることが多い。これまで RV を利用したボラティリティの予測モデルには様々な定式化が提案されてきた。先行研究の中で、予測精度が比較的優れていると論じられているモデルとしては、Heterogeneous Autoregressive (HAR) モデルがあり、日次 RV の変動が過去の日次、週次、月次といった周期の異なる RV の関数として定式化されている。また、日次の資産価格収益率からボラティリティを未知の変数として推定するモデルとして、GARCH モデルや Stochastic Volatility (SV) モデルがあるが、これらに RV を加えた、Realized GARCH (RGARCH) モデルや Realized SV (RSV) モデルによるボラティリティ予測も注目されている。本稿では、まず、RV を用いたモデルについて先行研究のサーベイを行う。次に、日経 225 株価指数の日次リターンと日次 RV を用いて、先行研究で扱われている主要なモデルについて、ボラティリティの予測精度の比較を行う。予測期間を新型コロナウイルス感染症の拡大前後で分け、優れているモデルが時期によって異なるかどうかを検証する。分析の結果、コロナの発生直後を含む期間では、HAR モデルや RGARCH モデルの拡張である REGARCH モデルの予測精度が比較的高く、それ以外の期間についてはコロナ前もコロナ継続期間も RSV モデルの予測精度が高いことが明らかになった。JEL Classification Codes: C22, C53, C58, G17

1. 序論

資産価格のボラティリティは、資産価格変化率(リターン)の分散もしくは標準偏差を表し、ファイナンスの研究でも実務でも重要な変数である。例えば、オプション価格の評価、Value-at-Risk (VaR) や期待ショートフォール (Expected shortfall; ES) の計算で、ボラティリティは重要な役割を果たす。しかし、ボラティリティの値は観測されないで、過去のリターンから推定する必要がある。また、ボラティリティもリターン同様、確率的に変動するとの考えが主流になっており、時間を通じて変動するボラティリティをどのように推定・予測すべきかは、計量ファイナンスの重要な課題である。市場が効率的であれば、リターンに自己相関はなく、過去のリターンから将来のリターンを予測できないが、ボラティリティは、上昇(低下)するとしばらくボラティリティが高い(低い)日が続くことが知られており、このボラティリティの高

い自己相関はボラティリティ・クラスタリングと呼ばれる。そこで、ボラティリティは過去の値からある程度予測可能なので、より正確にボラティリティを予測できる時系列モデルの開発が進められている。

古くは、ボラティリティの変動を定式化する時系列モデルを構築し、日次リターンを用いてそのモデルを推定することにより、ボラティリティを推定・予測していた。そうした時系列モデルには大きく分けて 2 種類ある。一つは、Engle(1982) の Autoregressive Conditional Heteroskedasticity (ARCH) モデル、Bollerslev(1986) の Generalized ARCH (GARCH) モデル、およびそれらを拡張したモデルで、本稿ではそうしたモデルを総称して ARCH 型モデルと呼ぶ。ARCH 型モデルは、パラメータの値とボラティリティの初期値が与えられると、 t 期のボラティリティが $t-1$ 期に既知になる。そのため、誤差項の分布が与えられれば、尤度を簡単に計算できるので、パラメータを最尤推

定できる。そこで、さまざまな拡張が行われてきた。株式市場では株価が上がった日の翌日より下がった日の翌日の方がボラティリティが上昇する傾向があることが知られており、ARCH, GARCH モデルはボラティリティ・クラスタリングは捉えられるものの、そうしたボラティリティ変動の非対称性は捉えられないので、ボラティリティ変動の非対称性を捉えるため、Nelson(1991)の Exponential GARCH(EGARCH)モデル、Glosten et al.(1993)の GJR-GARCH モデル、Ding et al.(1993)の Asymmetric Power GARCH(APGARCH)モデルなどが提案された。これらのモデルについて詳しくは、渡部(2000, 2007)を参照されたい。

もう一つは、Stochastic Volatility(SV)モデルである。SV モデルはボラティリティの変動を表す式に確率変数である誤差項が加わるので、 t 期のボラティリティが $t-1$ 期に既知にならず、観測されない状態変数になる。このモデルは非線形状態空間モデルなので、尤度の計算が難しいため、モデルの発展よりも推定法の開発が行われてきた。現在よく用いられるのは、マルコフ連鎖モンテカルロ法(Markov chain Monte Carlo; MCMC)を用いたベイズ推定である。SV モデルもボラティリティ変動の非対称性を考慮したモデルに拡張されている。SV モデルやその MCMC を用いたベイズ推定について詳しくは、Watanabe and Omori(2004)、渡部(2000, 2005)、Omori et al.(2007)、Omori and Watanabe(2008)を参照されたい。

その後、2000 年代頃から、資産価格の日中の高頻度データが普及し始めたことを背景として、そうしたデータから計算される実現ボラティリティ(Realized Volatility; RV)がボラティリティの推定量として注目され始めた。ARCH 型モデルや SV モデルなどでは、ボラティリティを時系列モデルにおける未知の変数として推定しなければならないが、日次 RV は、単純なものであれば、例えば、日中における 5 分ごとの資産価格のリターンの 2 乗を一日分合計すればよいだけであるため、簡便であり取り扱いやすい。一方、資産価格にはマイクロストラクチャ

ャー・ノイズと呼ばれるノイズが含まれており、そのため RV にはバイアスが生じ、RV の計算に用いる日中リターンの時間間隔を短くすればするほど、バイアスが大きくなる。また、ボラティリティの予測という目的のためには、RV を時系列データとしてフィットの良い時系列モデルの開発が必要になる。こうした RV の性質や問題点を背景として、RV に関する研究は 2000 年代と 2010 年代を中心に急速に拡大した。RV は自己相関係数の減衰が遅く、長期記憶性を持つ可能性があることがわかっており、そのような時系列にふさわしいモデルの開発が進んできた。まず、長期記憶の時系列モデルとして、古くから研究が進んでいた、Autoregressive Fractionally Integrated Moving Average(AR-FIMA)モデルが RV に適していると主張する研究がみられた。その後、Corsi(2009)によって提案された Heterogeneous Autoregressive(HAR)モデルが、RV に良好なフィットを示すことから、広く使われるようになった。HAR モデルは、日次 RV を過去の日次、週次、月次といった 3 つの異なる周期の RV の関数として表す。一見、アドホックな定式化であり、長期記憶の時系列モデルではないが、様々な実証分析でボラティリティの予測精度が高いことが示されており、HAR モデルを基礎とした拡張が多く研究されてきた。HAR モデルも、ボラティリティ変動の非対称性を考慮したモデルに拡張されている。HAR モデルについて詳しくは、渡部(2020)を参照されたい。

渡部(2007)において、こうした RV に関する研究をサーベイしたが、その後もこの分野は発展が著しい。主な発展は以下の 3 つである。

(1) 資産価格にジャンプがある場合、RV は連続的に変化するボラティリティだけでなく、ジャンプも含むので、RV を連続成分(Continuous Component)とジャンプ成分(Jump Component)に分割する統計的手法が Barndorff-Nielsen and Shephard(2004, 2006)、Corsi et al.(2010)、Andersen et al.(2012)らによって提案された。また、Andersen et al.(2007)は、過去の日次、週次、月次の連続成分とジャンプ成分

を別々に HAR モデルの説明変数とする HAR-CJ モデルを提案した。(2) GARCH モデルや SV モデルに RV を組み込んだモデルが提案された。まず, Takahashi et al.(2009)は, SV モデルの中で推定される日次のボラティリティがその日の日中のリターンから計算された RV と整合的になるような仮定を置き, ボラティリティを推計することを提案した。すなわち, 日次のリターンの情報からボラティリティを推定する通常の SV モデルに, RV の情報も加えてボラティリティの推定の精度を上げるという方法である。このモデルは, その後, Realized SV (RSV)モデルとして, 様々な拡張が行われている。RSV モデルとその拡張については, 高橋他(2020)を参照されたい。同様に, ARCH 型モデルに RV を組み合わせた Realized GARCH (RGARCH) モデル(Hansen et al.(2012), Watanabe(2012)), Realized EGARCH(REGARCH) モデル(Hansen and Huang(2016))も提案され, これについても拡張が行われている。(3) ボラティリティは真の値がわからないので, 予測精度を比較するためには代理変数が必要である。真のボラティリティの代理変数として, 古くは日次リターンの 2 乗が使われていたが(渡部(2000)), Andersen and Bollerslev(1998)や Hansen and Lunde(2006a)は, それでは予測精度の正確な比較ができないので, RV を使うことを提案している。その後, Patton(2011)が, 代理変数が真のボラティリティの条件付き不偏推定量であれば, Mean Square Error(MSE)や Quasi-likelihood(QLIKE)などの損失関数は, 真のボラティリティを用いた場合と同じ順位になることを理論的に示したため, 近年ではこの方法が用いられることが多い。その場合, 高頻度の日中リターンを使って計算した RV にはマイクロストラクチャー・ノイズによるバイアスがあるので, 例えば, 30 分間隔などの比較的時間間隔の長い日中リターンを用いて計算した RV や Barndorff-Nielsen et al.(2008, 2009)によって提案されたマイクロストラクチャー・ノイズを考慮して計算する Realized Kernel(RK)などを代理変数として用いる。

本稿では, こうした新たな発展も含めて, RV の研究のサーベイを行う。また, 日経 225 株価指数の日次リターンと日次 RV を用いて, HAR, HAR-CJ, EGARCH, REGARCH, SV, RSV モデルの推定と 1 期先のボラティリティの予測を行い, Patton(2011)の方法により予測精度の比較を行う。すべてのモデルで株式市場のボラティリティ変動の非対称性を考慮した。HAR, REGARCH, RSV モデルでは, 5 分ごとのリターンの 2 乗を足し合わせた RV を用いる。REGARCH, RSV モデルで RV のバイアスが除去できるかどうかを分析するために, RK を用いた推定と予測も行う。リターンの分布については, 正規分布と t 分布を仮定する。また, 予測期間を新型コロナウイルス感染症の拡大前後で分け, 予測精度の高いモデルが時期によって異なるかどうかを検証する。分析の結果, コロナの発生直後を含む期間では, HAR モデルや REGARCH モデルの予測精度が比較的高く, それ以外の期間についてはコロナ前もコロナ継続期間も RSV モデルの予測精度が高いことが明らかになった。Takahashi et al.(2021)でも同様の分析を行っている。そこでは, 日経 225 株価指数だけでなく, DJIA, FTSE, STOXX といった海外の株価指数も分析しているが, コロナ前の期間しか分析していない。また RV しか用いておらず, HAR-CJ モデルも用いていない。リターンの分布も正規分布しか用いていない。

本稿の構成は以下の通りである。続く第 2 節では, RV の定義と基本的な性質について説明する。第 3 節では, RV に関する主要なモデルについて説明する。第 4 節では, 日経 225 株価指数のデータを用いた実証分析を行う。第 5 節では結論を述べる。

2. Realized Volatility

資産価格の対数値 $p(s)$ が以下の拡散過程に従っているとする。

$$dp(s) = \mu(s)ds + \sigma(s)dW(s). \quad (1)$$

ここで, s は時間, $\mu(s)$ はドリフト, $\sigma(s)$ は

瞬時的ボラティリティ, $W(s)$ はウィナー過程である. ファイナンス理論では $\sigma(s)$ を瞬時的ボラティリティと呼ぶが, ここでは以下 $\sigma^2(s)$ を瞬時的ボラティリティと呼ぶ. オプション価格を導出するための Black and Scholes (1973) モデルでは $\mu(s), \sigma(s)$ を時間を通じて一定と仮定するが, ここでは時間を通じて変化することを許容する.

このとき, $t-1$ を第 $t-1$ 日の最終時点, t を第 t 日の最終時点とすると, その間の第 t 日の真のボラティリティは次のように定義される.

$$IV_t = \int_{t-1}^t \sigma^2(s) ds. \quad (2)$$

これは瞬時的ボラティリティ $\sigma^2(s)$ を $t-1$ 時点から t 時点まで積分したものである, 積分ボラティリティ (Integrated Volatility; IV) と呼ばれる.

最も単純な第 t 日の RV は, $t-1$ 時点から t 時点までの日中リターン $\{r_{t-1+1/n}, r_{t-1+2/n}, \dots, r_t\}$ を使って次のように計算される.

$$RV_t = \sum_{i=1}^n r_{t-1+i/n}^2. \quad (3)$$

ここで, 資産価格にノイズがなければ, $n \rightarrow \infty$ とすると, RV_t は IV_t に確率収束する.

$$\text{plim}_{n \rightarrow \infty} RV_t = IV_t. \quad (4)$$

そこで, n を大きくすればするほど (日中リターンの時間間隔を短くすればするほど), RV は IV のより精度の高い推定量になる.

しかし, 実際の資産価格には bid-ask bounce, 非同期取引などによるマイクロストラクチャー・ノイズと呼ばれるノイズが加わる. 市場の板には指値注文の売値 (ask) と買値 (bid) が並んでおり, そこに成り行きの買い注文が入ると最も安い売値 (best ask) で, 成り行きの売り注文が入ると最も高い買値 (best bid) で約定する. そこで, 真の価格が変わらなくても, 約定価格は売値で約定するか買値で約定するかで変動する. これが bid-ask bounce によるノイズである. また, 流動性の低い株式は高い株式より情報が反映するのが遅れるので, それらの株価を平均した株価指数の変化率には自己相関が生じ

る. これが, 非同期取引によるノイズである. こうしたマイクロストラクチャー・ノイズについて詳しくは, Campbell et al. (1997) の Chapter 3 を参照されたい.

以下, 約定価格や株価指数などの観測される価格の対数値 $\tilde{p}(s)$ と (1) 式に従う真の価格の対数値 $p(s)$ の間に以下の関係が成り立つものとする.

$$\tilde{p}(s) = p(s) + \eta(s), \eta(s) \sim \text{WN}(0, \sigma_\eta^2). \quad (5)$$

ここで, $\eta(s)$ がマイクロストラクチャー・ノイズを表し, 簡単化のため, 平均 0, 分散 σ_η^2 のホワイト・ノイズであり, 任意の t_0 時点から t_1 時点までの真の価格の変化率 $p(t_1) - p(t_0)$ と無相関であると仮定する. この仮定の下で, t_0 時点から t_1 時点までの観測される価格の変化率 $\tilde{p}(t_1) - \tilde{p}(t_0)$ の分散は以下のように表せる.

$$\begin{aligned} \text{Var}(\tilde{p}(t_1) - \tilde{p}(t_0)) &= \text{Var}(p(t_1) - p(t_0)) \\ &\quad + \text{Var}(\eta(t_1) - \eta(t_0)) \\ &= \int_{t_0}^{t_1} \sigma^2(s) ds + 2\sigma_\eta^2. \end{aligned} \quad (6)$$

ここで, t_1 と t_0 の時間間隔を短くすると, (6) 式右辺第 1 項の真の価格変化率の分散が小さくなるのに対して, 第 2 項のマイクロストラクチャー・ノイズの変分の分散は $2\sigma_\eta^2$ で一定である. そこで, マイクロストラクチャー・ノイズの変分の変動が相対的に大きくなり, 観測される価格の変化率の 2 乗和として計算した RV のマイクロストラクチャー・ノイズによるバイアスが大きくなる.

こうしたマイクロストラクチャー・ノイズによる RV のバイアスを軽減する方法がいくつか提案されている. Ait-Sahalia et al. (2005), Bandi and Russell (2006, 2008) らはマイクロストラクチャー・ノイズを考慮した RV の計算に用いる日中リターンも最適な時間間隔の選択法, Zhang et al. (2005), Zhang (2006) らは Two (multi) scale estimator と呼ばれる異なる時間間隔の日中リターンを用いて計算した RV を組

み合わせる方法, Barndorff-Nielsen et al.(2008, 2009)はマイクロストラクチャー・ノイズによって生じる日中リターンの自己相関を考慮して計算する Realized Kernel(RK), Podolskij and Vetter(2009), Jacod et al.(2009)はプレアベレジ法を提案している¹⁾. Liu et al.(2015)がそうした方法を用いても5分ごとのリターンの2乗和として計算した単純なRVにボラティリティのサンプル内でのフィットや予測精度で上回らないとの結果を様々な資産について示しているが, 本稿の実証分析では, 5分ごとのリターン2乗和として計算した単純なRVに加え, RKも用いる.

RVを計算する上でもう1つ問題になるのは, 市場が閉まっていて取引がなく, 高頻度の資産価格が得られない時間帯をどう扱うかである. 東京証券取引所の現物株式の取引は9:30-11:30の前場と12:30-15:00の後場だけで, 15:00から次の日の9:30まで(以下, 夜間と呼ぶ)と11:30から12:30まで(昼休みと呼ぶ)は取引がない. 後場の終了から次の日の後場の終了までの24時間のボラティリティを推定したい場合, 夜間と昼休みを無視してRVを計算すると, 24時間のボラティリティを過小評価してしまう. 夜間のリターンの2乗と昼休みのリターンの2乗を加えると, それらは時間間隔が長いので, 時間の離散化によるRVの誤差が大きくなる. そこで, Hansen and Lunde(2005a)は, 以下のように, まず夜間と昼休みを無視してRVを計算し, それに定数 c を掛けて24時間のボラティリティの推定値に調整する方法を提案した.

$$RV_t^* = cRV_t, \quad c = \frac{\sum_{t=1}^T (y_t - \bar{y})^2}{\sum_{t=1}^T RV_t}. \quad (7)$$

ここで, RV_t は夜間と昼休みを無視して計算したRV, $y_t (t=1, \dots, T)$ は日次リターン, $\bar{y} = (1/T) \sum_{t=1}^T y_t$, RV_t^* は調整後のRV, c は調整パラメータである. これは, サンプル期間で日次リターンの分散とRVの平均が等しくなるように調整している. Hansen and Lunde(2005b)は別の調整方法を提案している.

3. Realized Volatility に関するモデル

3.1 HAR モデル

RVは自己相関係数の減衰が遅く, 長期記憶性を持つ可能性が指摘されている. 詳しくは, Andersen et al.(2001a, b, 2003), 渡部(2007)等を参照されたい. そのため, 長期記憶モデルである Autoregressive Fractionally Integrated Moving Average(ARFIMA)モデルが用いられることが多かった. 長期記憶性やARFIMAモデルについて詳しくは, Beran(1994), 矢島(2003), 田中(2006第7章)等を参照されたい. しかし, 近年よく用いられるのは, Corsi(2009)によって提案された Heterogeneous Autoregressive(HAR)モデルである. このモデルは以下のように日次RVを過去の日次, 週次, 月次といった周期の異なるRVの関数として表す²⁾.

$$\log RV_t = \alpha + \beta_d \log RV_{t-1} + \beta_w \log RV_{t-5:t-1} + \beta_m \ln RV_{t-22:t-1} + \nu_t. \quad (8)$$

ここで,

$$RV_{t-5:t-1} = \frac{1}{5} \sum_{i=1}^5 RV_{t-i}$$

$$RV_{t-22:t-1} = \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}$$

であり, それぞれ過去の週次, 月次RVを表す³⁾. HARモデルは長期記憶モデルではないが, 長期記憶過程をうまく近似できること, ボラティリティの予測精度が高いこと, パラメータを最小二乗法で簡単に推定できることからよく用いられる. 本稿では, 日次ボラティリティの予測しか行わないが, 週次や月次のボラティリティを予測したい時には, 被説明変数を週次や月次のRVに置き換えるだけでよい. 詳しくは, Andersen et al.(2007)や渡部(2020)を参照されたい. なお, HARモデルの拡張として, 状態空間モデルを用いたHARモデルも提案されている(Asai et al.(2012), Buccheria and Corsi(2021)).

株式市場では価格が下がった日の翌日の方がボラティリティが上昇する傾向があることが知られている. (8)式はこうしたボラティリティ

変動の非対称性を考慮していないので、本稿では、Bekaert and Hoerova(2014)に従い、以下の HAR モデルを用いる。

$$\log RV_t = \alpha + \beta_d \log RV_{t-1} + \beta_w \log RV_{t-5:t-1} + \beta_m \log RV_{t-22:t-1} + \delta y_{t-1}^- + \nu_t. \quad (9)$$

ここで、 $y_{t-1}^- = \min\{y_{t-1}, 0\}$ である。このモデルでは、 $\delta < 0$ であれば、価格が上がった日の翌日より下がった日の翌日の方がボラティリティがより上昇する。

RV の変動を表すモデルには ARFIMA モデルや HAR モデルの他に Unobserved Components(UC) モデルがある。このモデルについては、Barndorff-Nielsen and Shephard(2001, 2002), Barndorff-Nielsen et al.(2004), Nagakura and Watanabe(2015)を参照されたい。

3.2 HAR-CJ モデル

資産価格の対数値 $p(s)$ が拡散過程(1)式ではなく、以下のジャンプ拡散過程に従っているとする。

$$dp(s) = \mu(s)ds + \sigma(s)dW(s) + \kappa(s)dN(s). \quad (10)$$

ここで、 $\kappa(s)$ はジャンプの大きさ、 $N(s)$ は数え上げ測度(Counting Measure)である。

このとき、価格にノイズが無ければ、RV は IV ではなく、以下のように IV にジャンプを加えたものに収束する。

$$\text{plim}_{n \rightarrow \infty} RV_t = \int_{t-1}^t \sigma^2(s)ds + \sum_{t-1 < s \leq t} \kappa^2(s)ds. \quad (11)$$

ここで、右辺の第 1 項と第 2 項はそれぞれ連続成分(Continuous Component)、ジャンプ成分(Jump Component)と呼ばれ、以下、それらの推定値をそれぞれ C_t, J_t と表す。

連続成分はジャンプ成分よりも持続性が高いので、Andersen et al.(2007)は HAR モデルの説明変数を $RV_{t-1}, RV_{t-5:t-1}, RV_{t-22:t-1}$ ではなく、 $C_{t-1}, C_{t-5:t-1}, C_{t-22:t-1}, J_{t-1}, J_{t-5:t-1}, J_{t-22:t-1}$ にした HAR-CJ モデルを提案している。本稿では、株式市場におけるボラティリティ変動の非対称性

を考慮した以下のモデルを用いる。

$$\begin{aligned} \log RV_t = & \alpha + \beta_d \log C_{t-1} + \beta_w \log C_{t-5:t-1} \\ & + \beta_m \log C_{t-22:t-1} + \gamma_d \log(1 + J_{t-1}), \\ & + \gamma_w \ln(1 + J_{t-5:t-1}) \\ & + \gamma_m \ln(1 + J_{t-22:t-1}) \\ & + \delta y_{t-1}^- + \nu_t. \end{aligned} \quad (12)$$

ここで、

$$\begin{aligned} C_{t-5:t-1} &= \frac{1}{5} \sum_{i=1}^5 C_{t-i}, \quad C_{t-22:t-1} = \frac{1}{22} \sum_{i=1}^{22} C_{t-i}, \\ J_{t-5:t-1} &= \frac{1}{5} \sum_{i=1}^5 J_{t-i}, \quad J_{t-22:t-1} = \frac{1}{22} \sum_{i=1}^{22} J_{t-i}. \end{aligned}$$

Barndorff-Nielsen and Shephard(2004, 2006), Andersen et al.(2012), Corsi et al.(2010)らが RV を C と J に分割する方法を提案している。本稿では、Barndorff-Nielsen and Shephard(2004, 2006)の方法を用いる。Barndorff-Nielsen and Shephard(2004, 2006)は、価格にノイズが無い場合、以下の Bipower Variation(BV)が連続成分に収束することを示している。

$$BV_t = \mu_1^{-2} \frac{n}{n-1} \sum_{i=2}^n |r_{t-1+i/n}| |r_{t-1+(i-1)/n}|. \quad (13)$$

ここで、 $\mu_1 = \sqrt{2/\pi}$ 。

通常、この BV_t をそのまま C_t として用いるのではなく、 $RV_t - BV_t$ を用いてジャンプの検定を行い、統計的に有意であれば $J_t = RV_t - BV_t$ 、有意でなければ $J_t = 0$ とする。その上で、 $C_t = RV_t - J_t$ とする。ジャンプの検定にはいくつかの方法が提案されているが、本稿では、Barndorff-Nielsen and Shephard(2006), Huang and Tauchen(2005)の検定を用いる。この検定では以下の統計量を用いる。

$$Z_t^* = \frac{\sqrt{n} (RV_t - BV_t) RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5) \text{Max}(1, TQ_t BV_t^{-2})}}. \quad (14)$$

ここで、

$$TQ_t = n\mu_{4/3}^{-3} \frac{n}{n-4} \sum_{i=3}^n |r_{t-1+i/n}|^{4/3} |r_{t-1+(i-1)/n}|^{4/3} |r_{t-1+(i-2)/n}|^{4/3}, \quad (15)$$

$$\mu_{4/3} = 2^{2/3} \Gamma(7/6) \Gamma(1/2)^{-1}.$$

TQ_t は Tri-power Quarticity と呼ばれる積分 4 次モーメント (Integrated Quarticity; IQ)

$$IQ_t = \int_{t-1}^t \sigma^4(s) ds, \quad (16)$$

の推定量である。 Z_t^* はジャンプがないという帰無仮説の下で漸近的に標準正規分布に従う。以下では、有意水準を 5% として検定を行った。

3.3 EGARCH モデルと REGARCH モデル

近年では資産価格の高頻度データが利用可能となったため、ボラティリティの推定に RV を使うことが多くなったが、その前は、日次リターンを用いてボラティリティの変動を定式化する時系列モデルを推定することによりボラティリティの推定を行うのが主流であった。そうした時系列モデルは、大きく 2 つに分けられる。一つは、Engle(1982) の Autoregressive Conditional Heteroskedasticity (ARCH) モデル、Bollerslev(1986) の Generalized ARCH (GARCH) モデル、およびそれらを拡張した ARCH 型モデルで、もう一つは Stochastic Volatility (SV) モデルである。Nelson(1991) によって提案された Exponential GARCH (EGARCH) モデルは ARCH 型モデルの一つであり、Hansen and Huang(2016) は、RV のバイアスを考慮してこのモデルに RV を加えた Realized EGARCH (REGARCH) モデルを提案している。本節では、EGARCH モデルと REGARCH モデルについて説明する。

本稿では、EGARCH モデルとして、以下のモデルを用いる。

$$y_t = E(y_t | \mathbf{I}_{t-1}) + e_t, \\ e_t = \exp(h_t/2) \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad (17)$$

$$h_t = \omega + \beta(h_{t-1} - \omega) + \theta \epsilon_{t-1} + \gamma(\epsilon_{t-1}^2 - 1). \quad (18)$$

ここで、 y_t は日次リターン、 h_t は日次ボラティ

リティの対数値、すなわち、 $\exp(h_t)$ が日次ボラティリティ、 $\mathbf{I}_{t-1}$ は $t-1$ 日における情報集合を表す。以下の実証分析で用いる日経 225 株価指数の変化率(リターン)には有意な自己相関がなく、平均も有意に 0 から乖離しないので、 $E(y_t | \mathbf{I}_{t-1}) = 0$ とする。Nelson(1991) は、(18) 式の代わりに、

$$h_t = \omega + \beta(h_{t-1} - \omega) + \theta \epsilon_{t-1} + \gamma[|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)],$$

と定式化しているが、この定式化だと、 $E(|\epsilon_{t-1}|)$ が ϵ_{t-1} の分布に依存するので、以下の REGARCH モデルに倣って、(18) 式のように定式化した。

このモデルは ARCH モデルや GARCH モデルと比べていくつかの長所がある。GARCH モデルでは、ボラティリティ $\sigma_t^2 = \exp(h_t)$ の変動を以下のように定式化する。

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha e_{t-1}^2, \quad \omega > 0, \beta, \alpha \geq 0. \quad (19)$$

このモデルでは被説明変数を σ_t^2 としているので、それが必ず正になるよう、説明変数を正の値しか取らない変数とし、パラメータに非負制約を課さなければならない。それに対して、EGARCH モデルの(18)式では、 $h_t = \log \sigma_t^2$ を被説明変数とするので、パラメータに非負制約が必要なく、負の値を取る変数も説明変数に加えられる。そのため、前日の基準化した誤差項 ϵ_{t-1} を説明変数とすることにより、ボラティリティ変動の非対称性を考慮している。(18)式は、 $\theta < 0$ であれば、株式市場のボラティリティ変動の非対称性と整合的である。また、(19)式の e_{t-1}^2 は $\sigma_{t-1}^2 \epsilon_{t-1}^2$ なので、ボラティリティのショックの持続性は $\beta + \alpha$ で表されるが、(18)式は e_{t-1}^2 を説明変数とするので、 β の値だけで計ることができる。

ARCH 型モデルの特徴は、 t 期のボラティリティの説明変数がすべて $t-1$ 期に値のわかる変数だけで、 $t-1$ 期に値のわからない確率変数が含まれていないことである。そこで、EGARCH モデルの場合、 $h_0 = \omega$ とし、パラメ

ータの値が与えられれば、 $t-1$ 期に t 期のボラティリティ $\exp(h_t)$ の値が既知になる。そのため、 ϵ_t の分布が与えられれば、尤度を簡単に計算できるので、最尤法によってパラメータを推定できる。本稿では、 ϵ_t の分布に標準正規分布を用いる。 ϵ_t の分布は通常正規分布ではないが、それを正規分布と仮定して尤度を計算し最大化する方法を疑似最尤法と呼ぶ。本稿では疑似最尤法を用いて EGARCH モデルのパラメータを推定し、Bollerslev and Wooldridge(1992) の疑似最尤法の標準誤差を計算している。ARCH 型モデルの疑似最尤推定と標準誤差の計算について詳しくは、渡部(2000)を参照されたい。また、 ϵ_t の分布には、分散を 1 に標準化した t 分布も用いる。その他の ARCH 型モデルについては、渡部(2000, 2007)を参照されたい。ARCH 型モデルの中にもボラティリティの長期記憶性を考慮しているモデルがあり、そうしたモデルについては、Baillie et al.(1996), Bollerslev and Mikkelsen(1996), Tse(1998), 竹内(野木森)・渡部(2008)を参照されたい。

$x_t = \log(RV_t)$ とすると Hansen and Huang (2016) の提案した Realized EGARCH(RE-GARCH) モデルは(17)式と以下の式で表される。

$$h_t = \omega + \beta(h_{t-1} - \omega) + \tau_1 \epsilon_{t-1} + \tau_2(\epsilon_{t-1}^2 - 1) + \gamma u_{t-1}, \quad (20)$$

$$x_t = \xi + \phi h_t + \delta_1 \epsilon_t + \delta_2(\epsilon_t^2 - 1) + u_t, \quad u_t \sim N(0, \sigma_u^2). \quad (21)$$

(20)式は真のボラティリティの変動を表しており、(21)式は RV と真のボラティリティの関係を表している。 ϕ は通常 1 から有意に乖離しないので、以下では、 $\phi=1$ と仮定した。そこで、 $\xi=0$ であれば、 x_t は h_t の不偏推定量になる。Hansen and Lunde (2006b) はマイクロストラクチャー・ノイズによる RV のバイアスが正にも負にもなり得ることを示している。以下の実証分析では、 y_t を終値の対数階差として計算しているので、真のボラティリティ $\exp(h_t)$ は 24 時間のボラティリティである。それに対して、 RV_t には夜間と昼休みを無視して計算した RV

を用いるので、RV は真のボラティリティ $\exp(h_t)$ を過小評価する。そこで、RV のマイクロストラクチャー・ノイズによるバイアスが正で、夜間と昼休みを無視したことによる負のバイアスを上回らない限り、 $\xi < 0$ になる。

RV のバイアスを考慮して ARCH 型モデルを同様に拡張したモデルには、他にも Hansen et al.(2012) の Realized GARCH モデル、Huang et al.(2016) の Realized HAR GARCH モデル、Chen and Watanabe(2019) の Realized Threshold GARCH モデルなどがある。Chen et al.(2021) は、これらのモデルからベイズ推定法によって 1 期先のボラティリティと VaR, ES の予測を行い、予測精度を比較している。

3.4 SV モデルと RSV モデル

RSV モデルを説明するために、まず、SV モデルを導入する。標準的な SV モデルは次式で表される。

$$y_t = \sqrt{\lambda_t} \epsilon_t \exp(h_t/2), \quad (22)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad (23)$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N(\mathbf{0}, \mathbf{S}), \mathbf{S} = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}. \quad (24)$$

(22)式は、(17)式において、 $E(y_t | \mathbf{I}_{t-1}) = 0$ とし、 $\sqrt{\lambda_t}$ を加えたものである。SV モデルにおけるボラティリティも $\exp(h_t)$ であり、 h_t はボラティリティの対数値である。この対数ボラティリティは、(23)式で AR(1)過程に従うと仮定する。なお、その過程は定常であると仮定し、 $|\phi| < 1$ とする。対数ボラティリティの初期値は定常分布に従っていると仮定し、 $h_1 \sim N(\mu, \sigma^2/(1-\phi^2))$ とする。(24)式で、誤差項である (ϵ_t, η_t) は、2 変量正規分布に従うと仮定している。その相関係数である ρ が負であれば、 y_t が正の場合よりも負の場合の方が翌日の対数ボラティリティ h_{t+1} が上昇しやすいという株式市場で観測されるボラティリティ変動の非対称性と整合的である。 λ_t は観測方程式(22)におけるリターンの分布を規定する変数であり、全ての t において $\lambda_t = 1$ の場合は、標準正規分布となる。一方、 $(\nu-2)/\lambda_t \sim \chi^2(\nu)$ と

仮定すると、 $\sqrt{\lambda_t}\epsilon_t$ は分散を 1 に基準化した自由度 ν の t 分布に従う。本稿では、この t 分布と正規分布の両方のケースを推定する。

このモデルは、ARCH 型モデルと異なり、ボラティリティの変動を表す (23) 式に確率変数 η_t があるため、 $t-1$ 期に t 期のボラティリティの値は未知である。そこで、 h_t を観測されない状態変数 (潜在変数) とする非線形状態空間モデルになるので、尤度を解析的に求めるのが難しい。そのため、最尤法に代わる推定法がいくつか提案されてきたが (渡部 (2000) 第 3.2 節を参照されたい)、近年ではベイズ推定を用いることが多い。頻度論統計学では、パラメータを定数、サンプルを確率変数と考えるのに対して、ベイズ統計学では、パラメータを確率変数、サンプルを定数と考え、パラメータにデータを観測する前の事前分布を設定し、データを観測した後の事後分布に更新することにより、パラメータの推定を行う。ベイズの定理によると事後分布は事前分布と尤度の積に比例するので、SV モデルのように尤度を解析的に求めるのが難しいモデルの場合、ベイズの定理を使って解析的に事後分布を求めるのも難しい。そこで、マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo; MCMC) を用いて事後分布からパラメータと潜在変数をサンプリングし、得られたサンプルに基づいてパラメータを推定する。MCMC は 1 回前にサンプリングされた値に依存させて次のサンプリングを行う方法の総称である (MCMC 法について詳しくは、中妻 (2007, 2013)、伊庭他 (2005)、和合 (2005) 等を参照されたい)。

さて、SV モデルに RV を組み込んだ RSV モデルは、次式で表される。

$$y_t = \sqrt{\lambda_t \epsilon_t} \exp(h_t/2), \quad (25)$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad (26)$$

$$x_t = \xi + h_t + u_t, \quad (27)$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \\ u_t \end{pmatrix} \sim N(\mathbf{0}, \Sigma), \Sigma = \begin{pmatrix} 1 & \rho\sigma & 0 \\ \rho\sigma & \sigma^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{pmatrix}. \quad (28)$$

(25) 式と (26) 式は、SV モデルの (22) 式と (23) 式と同一である。RSV モデルでは、(27) 式が追加されており、 x_t である RV の対数が SV モデルにおける対数ボラティリティ (つまり、真のボラティリティ IV_t の推定量) と整合的であるようにモデル化している。 ξ は RV に含まれるバイアスを補正するためのパラメータであり、 $\xi=0$ の場合に、RV と SV のボラティリティのそれぞれの対数値が平均的に一致するような定式化となっている。3.3 節で説明した RE-GARCH モデル同様、RV のマイクロストラクチャー・ノイズによるバイアスが正で、夜間と昼休みを無視したことによる負のバイアスを上回らない限り、 $\xi < 0$ になる。(27) 式の右辺の h_t を βh_t というように係数パラメータを追加するとより柔軟にバイアスを考慮するモデルとなるが、先行研究では β を 1 としても予測精度に大きな違いはないため、本稿では、 $\beta=1$ としている。この式の誤差項 u_t と他の誤差項の相関は 0 と仮定する。

RSV モデルも最尤推定が難しいので、MCMC を用いてベイズ推定する。まず、 $\mathbf{Y} = \{(y_t, x_t)\}_{t=1}^T$, $\mathbf{h} = \{h_t\}_{t=1}^T$, $\boldsymbol{\lambda} = \{\lambda_t\}_{t=1}^T$, $\boldsymbol{\omega} = (\phi, \sigma, \mu, \rho, \nu, \xi, \sigma_u)$ と定義する。データ $\mathbf{Y}$ が得られたときの RSV モデルにおける同時事後確率密度関数 $\pi(\mathbf{h}, \boldsymbol{\lambda}, \boldsymbol{\omega} | \mathbf{Y})$ からの確率標本は、MCMC 法を用いたサンプリングによって得ることができる。サンプリングの詳細については、本稿の補論、または、Takahashi et al. (2009, 2016, 2021), Omori and Watanabe (2015), 大森・渡部 (2013), 高橋他 (2020) を参照されたい。Shirota et al. (2014) は RSV モデルの (26) 式を長期記憶モデルである ARFIMA モデルにしている。

4. 日経 225 株価指数を用いた実証分析

4.1 データと予測方法

本節では、日経 225 株価指数を用いた実証分析を行う。データとしては、日次リターン (%) について、終値を P_t として対数階差を $y_t = 100(\log P_t - \log P_{t-1})$ と計算したものを用いる。

RV について、日中の 5 分ごとのリターンの 2 乗和を計算した値を用いる。また、Barndorff-Nielsen et al.(2011)に従い、RK は日中の 1 分ごとのリターンを用いて、漸近的に最適なバンド幅にした Non flat-top Parzen kernel 推定量を用いて計算した。詳しくは、Ubukata and Watanabe(2014a)を参照されたい。ただし、この RV と RK の計算には昼休みと夜間休みのリターンを含めていない。

比較を行うモデルは、HAR モデル、EGARCH モデル、REGARCH モデル、SV モデル、RSV モデルである。RV を用いているモデルについては、5 分ごとの日中リターンから計算された RV を使うモデルを「RV」、RK を使うモデルを「RK」とする。例えば、HAR モデルであれば、HAR-RV と HAR-RK の両方を推定する。また、HAR モデルについては、5% 有意水準で抽出された連続成分を用いた HAR-C モデル、さらにジャンプ成分も追加した HAR-CJ モデルも推定する。さらに、HAR モデル以外は、誤差項について正規分布と t 分布をそれぞれ仮定したモデルを試す。 t 分布を仮定した場合には EGARCH t といったように「 t 」を付ける。

モデルの推定については、HAR モデルは最小二乗法、EGARCH、REGARCH モデルは疑似最尤法による推定を行う。SV モデルと RSV モデルは、前述の MCMC 法を用いて推定する。その際、事前分布は、先行研究に倣って、 $(\phi+1)/2 \sim B(20, 1.5)$, $\sigma^2 \sim IG(2.5, 0.025)$, $(\rho+1)/2 \sim B(1, 1)$, $\mu \sim N(0, 1)$, $\nu \sim G(16, 0.8)$, $\xi \sim N(0, 1)$, $\sigma_\xi^2 \sim IG(2.5, 0.025)$ とおく。ただし、 B, G, IG はそれぞれ、ベータ(Beta)分布、ガンマ(Gamma)分布、逆ガンマ(Inverse Gamma)分布である。MCMC 法は、稼動検査期間(burn-in period)として最初の 1,000 個を捨てた後、10,000 個のサンプルを発生させた。MCMC 法によって得られた標本自己相関関数はいずれのパラメータについても十分に減衰していて定常分布への収束が早く、また標本経路も状態空間を万遍なく十分に訪れていることが確認される。

モデル比較には、1 日先のボラティリティの予測値を用いる。

HAR モデルでは、(9), (12)式のパラメータに最小二乗推定値を代入することで $E(\log RV_{t+1} | \mathbf{I}_t)$ を計算し、誤差項 ϵ_t の分布を正規分布と仮定し、残差分散 $\hat{\sigma}^2$ を用いて、 RV_{t+1} の予測値を以下のように計算した。

$$E(RV_{t+1} | \mathbf{I}_t) = \exp \left[E(\log RV_{t+1} | \mathbf{I}_t) + \frac{1}{2} \hat{\sigma}^2 \right]. \quad (29)$$

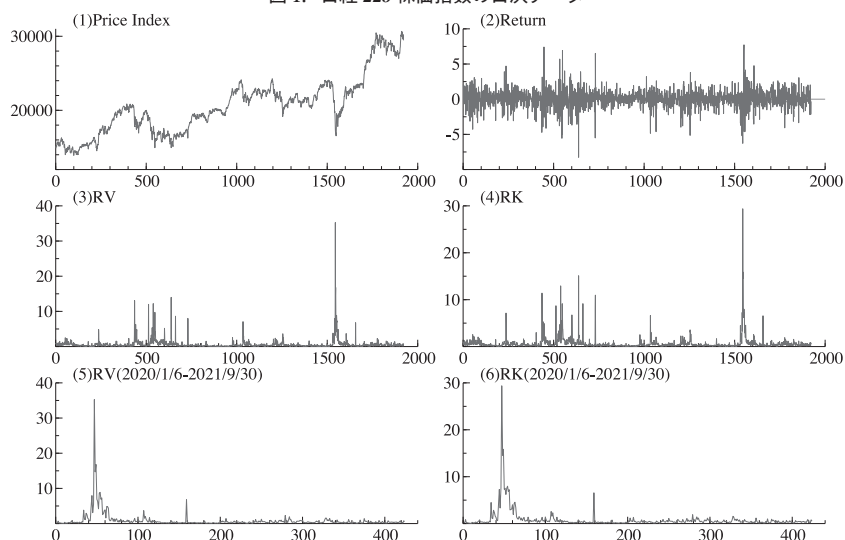
これは夜間と昼休みを除いたボラティリティの予測値なので、第 2 節で説明した Hansen and Lunde(2005a)の方法で、終値から終値までの 24 時間のボラティリティの予測値に調整した。REGARCH モデルと RSV モデルでは、この調整は(21), (27)式のパラメータ ξ によって行われるので、調整は必要ない。EGARCH モデルと REGARCH モデルでは、それぞれ(18)式、(20)-(21)式のパラメータに疑似最尤推定値を代入して h_{t+1} を計算し、 $\exp(h_{t+1})$ をボラティリティの予測値とした。

SV, RSV モデルは MCMC の各回のサンプルで、 h_{T+1} を発生させ、 $\exp(h_{T+1})$ の事後平均をボラティリティの予測値とした。

データの期間は、2013 年 11 月 14 日から 2021 年 9 月 30 日までの 1,922 日分を用意した。図 1 は、データの全期間の日経 225 株価指数の終値、日次リターン、RK, RV の時系列プロットである。RV, RK が最も高くなったのは、2020 年 3 月 13 日であり、得体のしれない新型コロナウイルスに関するニュースがボラティリティを大きく押し上げたものと考えられる。なお、この日の前後の数日間は金融市場でボラティリティの高い展開が続き、日中の取引で価格が大きく変化する状況がみられた。しかし、それ以降は感染者数が増加しても株価指数は上昇傾向にあり、RK, RV も低い日が続いている。

予測の計算については、まず、2013 年 11 月 14 日から 2017 年 12 月 13 日までの期間(観測個数は 1,000 個)を用いてモデルを推定し、翌日の 2017 年 12 月 14 日のボラティリティを予

図 1. 日経 225 株価指数の日次データ



注) (1) 指数, (2) リターン, (3) RV, (4) RK, (5) 新型コロナウイルス禍の RV, (6) 新型コロナウイルス禍の RK. 期間について, (1)-(4) は 2013/11/14 日から 2021/9/30 日まで, (5)-(6) は 2020/1/6 日から 2021/9/30 日までである.

測する. 次に, サンプル期間を 1 日後ろにずらして, 2013 年 11 月 15 日から 2017 年 12 月 14 日までの期間(観測個数は 1,000 個)を用いてモデルを推定し, 2017 年 12 月 15 日のボラティリティを予測する. これを繰り返し, 2021 年 9 月 30 日のボラティリティ予測が得られるまで, すなわち, 922 日分の予測が得られるまで推定を行う. このようにして, モデルのパラメータを毎回, 新たに推定しながら, ボラティリティの予測値を計算する.

モデル比較では, まず, 922 日分の予測の精度を確認する. 次に, この予測期間を新型コロナウイルス感染症の拡大前と後の期間に分けて, モデル比較を行う. 具体的には, 世界保健機構(WHO)が中国武漢での肺炎患者について新型コロナウイルスの可能性が否定できないと発表した 2020 年 1 月 8 日より前までの 500 個(以下, 「コロナ前」と呼ぶ)と, この日以降の 422 個(「コロナ発生直後・継続期間」と呼ぶ)について, それぞれのモデル比較を行う. 後者の期間については, 前述の通り, コロナ発生直後のボラティリティが高い時期と, その後のボラティリティが低い時期がみられるため, このボラティリティが低い時期のみについても, モデル比

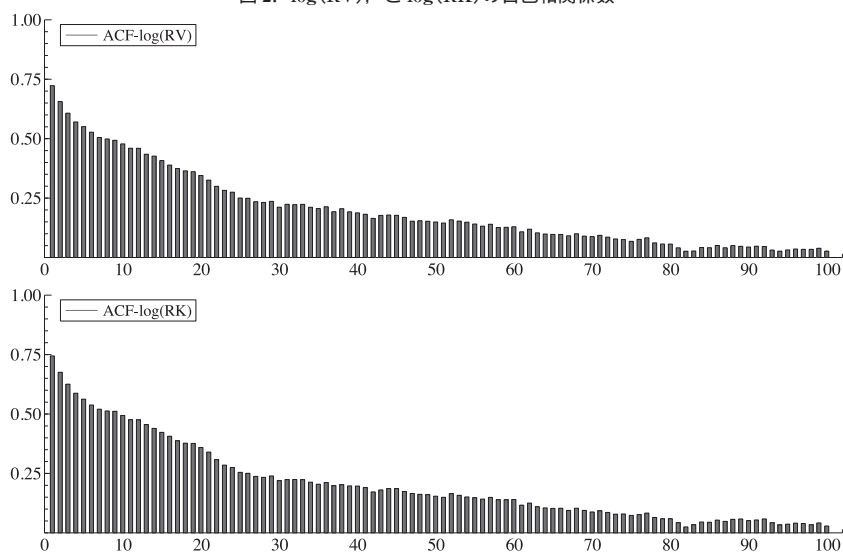
較を行う. 具体的には, 第一回緊急事態宣言が東京都などの 8 都道府県を除いた 39 県で解除された, 2020 年 5 月 14 日から 2021 年 9 月 30 日までの 339 個(「コロナ継続期間」と呼ぶ)である.

表 1 は日次リターン, $\log(RV)$, $\log(RK)$ の要約統計量を, それぞれサンプル期間を(1)全データ期間(2013/11/14 日-2021/9/30 日), (2)予測期間(2017/12/14 日-2021/9/30 日), (3)うちコロナ前(2017/12/14 日-2020/1/7 日), (4)うちコロナ発生直後・継続期間(2020/1/8 日-2021/9/30 日), (5)うちコロナ継続期間(2020/5/14 日-2021/9/30 日)に分けて示している. すべてのサンプル期間において, 日次リターンの平均は有意に 0 から乖離しない. LB(10)は 1 次から 10 次までの自己相関係数がすべて 0 であるという帰無仮説を検定するための Ljung and Box(1978)統計量で, Diebold(1988)に従い, 分散不均一性を調整している. この統計量によると, すべてのサンプル期間で, 日次リターンでは有意水準 10% でも帰無仮説は棄却されない. そこで, (17)式の $E(y_t | \mathbf{I}_{t-1})$ は 0 とする. リターンの標準偏差や RV, RK の平均はコロナ後(コロナの拡大開始時期を含む)の方がコロナ

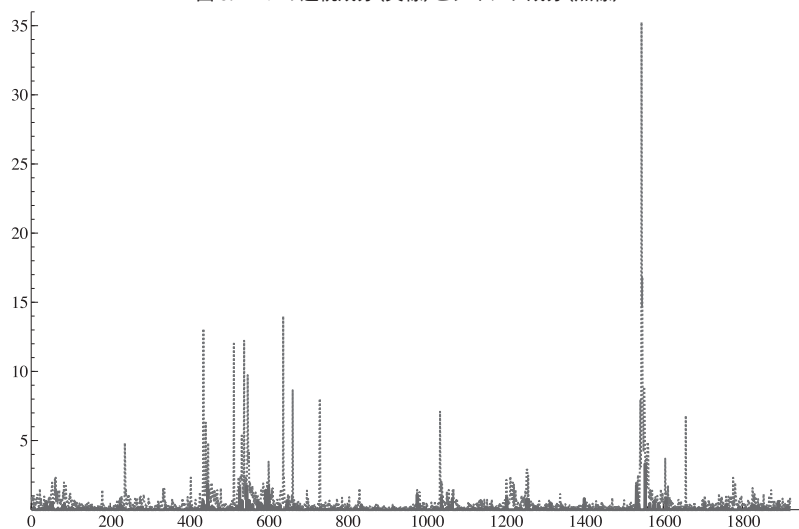
表 1. 日経 225 株価指数の日次データの要約統計量

期間	観測数	平均	標準偏差	最大値	最小値	歪度	尖度	LB(10)
リターン								
(1) 全データ期間	1922	0.037 (0.029)	1.284	7.731	-8.253	-0.164 (0.056)	7.956 (0.112)	6.591
(2) 予測期間	922	0.028 (0.041)	1.251	7.731	-6.274	-0.151 (0.081)	7.683 (0.161)	8.957
(3) うちコロナ前	500	0.007 (0.047)	1.057	3.810	-5.140	-0.733 (0.110)	6.209 (0.219)	6.367
(4) うちコロナ発生直後 ・継続期間	422	0.053 (0.071)	1.447	7.731	-6.274	0.100 (0.119)	7.248 (0.238)	8.076
(5) うちコロナ継続期間	339	0.110 (0.062)	1.141	4.767	-4.067	-0.089 (0.133)	4.266 (0.266)	16.085
log(RV)								
(1) 全データ期間	1922	-0.997 (0.021)	0.914	3.562	-3.894	0.761 (0.056)	4.304 (0.112)	2769.302
(2) 予測期間	922	-1.030 (0.029)	0.888	3.562	-3.894	0.935 (0.081)	5.269 (0.161)	916.443
(3) うちコロナ前	500	-1.228 (0.035)	0.789	1.958	-3.894	0.471 (0.110)	3.945 (0.219)	716.843
(4) うちコロナ発生直後 ・継続期間	422	-0.796 (0.046)	0.940	3.562	-3.025	1.167 (0.119)	5.361 (0.238)	419.452
(5) うちコロナ継続期間	339	-0.976 (0.037)	0.683	1.917	-3.025	0.384 (0.133)	3.708 (0.266)	233.357
log(RK)								
(1) 全データ期間	1922	-0.990 (0.021)	0.900	3.380	-3.923	0.823 (0.056)	4.446 (0.112)	2814.112
(2) 予測期間	922	-1.033 (0.028)	0.859	3.380	-3.923	1.042 (0.081)	5.593 (0.161)	934.414
(3) うちコロナ前	500	-1.225 (0.034)	0.756	1.891	-3.923	0.570 (0.110)	4.166 (0.219)	842.229
(4) うちコロナ発生直後 ・継続期間	422	-0.806 (0.045)	0.918	3.380	-3.067	1.249 (0.119)	5.622 (0.238)	412.977
(5) うちコロナ継続期間	339	-0.982 (0.035)	0.638	1.882	-2.782	0.425 (0.133)	3.784 (0.266)	256.894

注) 期間は、(1)2013/11/14日から2021/9/30日まで、(2)2017/12/14日から2021/9/30日まで、(3)2017/12/14日から2020/1/7日まで、(4)2020/1/8日から2021/9/30日まで、(5)2020/5/14日から2021/9/30日まで。括弧内は標準誤差。LB(10)は1次から10次までの自己相関係数がすべて0であるという帰無仮説を検定するためのLjung and Box(1978)統計量で、Diebold(1988)に従い、分散不均一性を調整した。RV, RKは夜間と昼休みを無視して計算したものである。

図 2. $\log(RV)$ と $\log(RK)$ の自己相関係数

注) 期間は、2013/11/14 日から 2021/9/30 日まで。

図 3. RV の連続成分(実線)とジャンプ成分(点線)

注) Barndorff-Nielsen and Shephard (2004, 2006) の方法によって分割した。ジャンプ成分は有意水準 5% で有意なものだけ抽出し、 RV からジャンプ成分を引いたものを連続成分とした。期間は、2013/11/14 日から 2021/9/30 日まで。

前よりも大きくなっている。

図 2 は $\log(RV)$ と $\log(RK)$ の自己相関係数を示している。横軸はラグ次数であり、どちらもラグ次数が増えても減衰が遅く、長期記憶性を持つ可能性が示唆される。図 3 は第 3.2 節で説明した Barndorff-Nielsen and Shephard (2004, 2006) の方法によって分割した RV の連続成分とジャンプ成分を時系列プロットしたグ

ラフである。ジャンプ成分は有意水準 5% で有意なものだけ抽出し、 RV からジャンプ成分を引いたものを連続成分とした。C と J の分割に Andersen et al. (2012) の方法も用いたが、C と J の値は Barndorff-Nielsen and Shephard (2004, 2006) の方法によって分割した場合とほとんど変わらなかったの、以下の分析では Barndorff-Nielsen and Shephard (2004, 2006) の

表 2. HAR モデルのパラメータの推定結果：最小二乗推定値

	α	β_d	β_w	β_m	γ	γ_w	γ_{dm}	δ	$\bar{R}^2$
HAR-RV モデル									
前期	-0.320 (0.035)	0.363 (0.042)	0.239 (0.060)	0.225 (0.042)				-0.166 (0.034)	0.606
後期	-0.363 (0.038)	0.216 (0.042)	0.455 (0.061)	0.135 (0.042)				-0.202 (0.032)	0.586
HAR-RK モデル									
前期	-0.305 (0.035)	0.405 (0.044)	0.210 (0.061)	0.213 (0.041)				-0.154 (0.037)	0.613
後期	-0.317 (0.037)	0.257 (0.042)	0.440 (0.061)	0.129 (0.042)				-0.171 (0.032)	0.619
HARCJ-RV モデル									
前期	-0.361 (0.058)	0.356 (0.041)	0.218 (0.058)	0.198 (0.045)	0.186 (0.212)	0.586 (0.450)	0.629 (0.631)	-0.164 (0.034)	0.604
後期	-0.375 (0.083)	0.217 (0.042)	0.461 (0.065)	0.087 (0.058)	0.019 (0.152)	-0.133 (0.363)	1.027 (1.101)	-0.206 (0.030)	0.587
HARC-RV モデル									
前期	-0.265 (0.037)	0.360 (0.041)	0.235 (0.061)	0.228 (0.044)				-0.173 (0.034)	0.601
後期	-0.304 (0.039)	0.217 (0.042)	0.461 (0.063)	0.125 (0.043)				-0.207 (0.031)	0.588

注) 括弧内は Newey and West(1987)の標準誤差。 $\bar{R}^2$ は自由度修正済み決定係数。期間は、前期が 2013/11/14 日から 2017/12/13 日まで、後期が 2017/8/22 日から 2021/9/30 日まで。

方法によって分割した C と J だけを用いている。

4.2 各モデルの推定結果

予測精度の検証を行う前に、各モデルのパラメータの推定結果を確認する。データの期間の後半には、新型コロナウイルス感染症の拡大時期が含まれており、期間の前半と後半でモデルのパラメータに変化が生じている可能性がある。そこで、データ期間の最初の 1,000 個(「前期」と呼ぶ)と最後の 1,000 個(「後期」)について、それぞれモデルを推計し、パラメータの推計値を表 2-4 にまとめた。

表 2 は、最小二乗法による 4 つの HAR モデルのパラメータの推定値、Newey and West (1987)の標準誤差、自由度修正済み決定係数を示している。RV, RK のどちらを用いた場合でも、また前期、後期ともに、HAR モデルの過去の日次、週次、月次 RV の係数である $\beta_d, \beta_w, \beta_m$ の推定値はすべて有意水準 1% で統計的に有意な正の値になっている。また、ボラティリティ変動の非対称性を表す δ の推定値

は有意水準 1% で統計的に有意な負の値になっており、株式市場で観測される価格が下がった日の翌日の方がボラティリティがより上昇するという現象と整合的である。RV を連続成分 C とジャンプ成分 J に分割したため、HAR-CJ モデルでは RV を用いた推定だけ行った。前期も後期も連続成分 C の係数は日次、週次、月次すべてで有意水準 1% で統計的に有意な正の値になっているが、ジャンプ成分 J の係数はすべて有意水準 10% でも有意でない。そこで、説明変数からジャンプ成分を削除した HAR-C モデルも RV を用いた推定だけ行った。そこでも前期、後期とも連続成分 C の係数は日次、週次、月次すべてで有意水準 1% で統計的に有意な正の値になっている。また、HAR-CJ モデル、HAR-C モデルでも、前期、後期ともにボラティリティ変動の非対称性を表す δ の推定値は有意水準 1% で統計的に有意な負の値になっている。自由度修正済み決定係数は前期、後期ともすべてのモデルで 0.6 前後と高い値を示している。

表 3 は疑似最尤法によって推定された

表 3. EGARCH モデルと REGARCH モデルのパラメータの推定結果：疑似最尤推定値

	ω	β	θ	γ	ν	対数尤度
EGARCH モデル						
前期	0.354 (0.139)	0.949 (0.018)	-0.158 (0.031)	0.056 (0.015)		-1553.17
後期	0.208 (0.107)	0.894 (0.059)	-0.175 (0.046)	0.081 (0.029)		-1490.21
EGARCHt モデル						
前期	0.443 (0.167)	0.947 (0.023)	-0.193 (0.037)	0.053 (0.018)	4.793 (0.781)	-1518.02
後期	0.251 (0.118)	0.869 (0.065)	-0.213 (0.051)	0.100 (0.031)	5.346 (0.825)	-1464.08

	ω	β	τ_1	τ_2	γ	ξ	δ_1	δ_2	σ_u^2	ν	対数尤度
REGARCH-RV モデル											
前期	0.363 (0.162)	0.940 (0.017)	-0.138 (0.017)	0.051 (0.014)	0.300 (0.042)	-1.208 (0.066)	-0.146 (0.017)	0.103 (0.013)	0.259		-2283.26
後期	0.134 (0.136)	0.928 (0.015)	-0.122 (0.015)	0.070 (0.008)	0.296 (0.027)	-1.186 (0.055)	-0.151 (0.016)	0.116 (0.009)	0.246 (0.015)		-2197.81
REGARCHt-RV モデル											
前期	0.358 (0.164)	0.935 (0.016)	-0.144 (0.017)	0.052 (0.016)	0.321 (0.040)	-1.215 (0.071)	-0.147 (0.017)	0.102 (0.014)	0.258 (0.018)	5.419 (1.035)	-2245.14
後期	0.132 (0.154)	0.935 (0.014)	-0.116 (0.015)	0.073 (0.008)	0.298 (0.026)	-1.190 (0.057)	-0.152 (0.016)	0.115 (0.009)	0.246 (0.015)	5.797 (0.926)	-2172.02
REGARCH-RK モデル											
前期	0.351 (0.157)	0.932 (0.019)	-0.138 (0.018)	0.053 (0.017)	0.329 (0.048)	-1.192 (0.067)	-0.144 (0.018)	0.103 (0.014)	0.249 (0.017)		-2265.72
後期	0.117 (0.142)	0.934 (0.013)	-0.115 (0.014)	0.066 (0.007)	0.325 (0.029)	-1.182 (0.055)	-0.147 (0.016)	0.108 (0.009)	0.209 (0.014)		-2111.38
REGARCHt-RK モデル											
前期	0.340 (0.158)	0.927 (0.018)	-0.144 (0.018)	0.055 (0.018)	0.355 (0.046)	-1.195 (0.071)	-0.145 (0.018)	0.101 (0.015)	0.249 (0.017)	5.459 (1.053)	-2226.21
後期	0.106 (0.165)	0.941 (0.012)	-0.109 (0.014)	0.068 (0.008)	0.324 (0.028)	-1.183 (0.057)	-0.148 (0.016)	0.107 (0.009)	0.208 (0.014)	5.929 (0.954)	-2086.63

注) 括弧内は Bollerslev and Wooldridge(1992)の疑似最尤推定量の標準誤差。期間は、前期が 2013/11/14 日から 2017/12/13 日まで、後期が 2017/8/22 日から 2021/9/30 日まで。

EGARCH モデルと REGARCH モデルのパラメータの推定値と疑似最尤推定量の標準誤差を示している。ボラティリティの持続性を表す β が先行研究と比べて若干低いものの、すべてのモデルでボラティリティ・クラスタリングを示している。EGARCH モデルの θ と REGARCH モデルの τ_1 はいずれも有意水準 1% で有意な負の値を示しており、株式市場におけるボラティリティ変動の非対称性と整合的である。REGARCH モデルにおいて RV のバイアスを表す ξ はすべて有意水準 1% で有意な負の値になっており、これは第 3.3 節で予想した通りである。

帰無仮説を ϵ_t の分布が標準正規分布、対立仮説を分散を 1 に基準化した t 分布として、表 3 に示されている対数尤度を用いて尤度比検定を行うと、すべてのモデルで帰無仮説は棄却される。

表 4 は、ベイズ推定法によって推定された SV モデルと RSV モデルの事後平均と事後標準偏差を示している。 ϕ はボラティリティの持続性を捉えるパラメータであり、推定値は 0.9 程度と、やはり先行研究と比べて若干低いものの、ボラティリティが持続的であることがわかる。 ρ はボラティリティ変動の非対称性を捉え

表 4. パラメータの推定結果(3)：事後平均

	ϕ	σ	ρ	μ	ξ	σ_u	ν
SV モデル							
前期	0.909 (0.026) [0.37; 164.8]	0.421 (0.078) [0.52; 267.4]	-0.609 (0.074) [0.80; 120.2]	0.213 (0.142) [0.43; 9.9]			
後期	0.890 (0.026) [0.32; 97.0]	0.437 (0.067) [0.30; 198.4]	-0.558 (0.071) [0.47; 97.2]	0.067 (0.131) [0.86; 12.7]			
SVt モデル							
前期	0.937 (0.018) [0.24; 87.5]	0.319 (0.050) [0.22; 138.9]	-0.699 (0.066) [0.10; 51.8]	0.273 (0.141) [0.52; 8.2]			16.194 (4.565) [0.79; 175.9]
後期	0.913 (0.021) [0.10; 85.9]	0.362 (0.055) [0.15; 148.6]	-0.608 (0.070) [0.13; 68.1]	0.114 (0.130) [0.08; 13.5]			19.036 (4.731) [0.95; 137.8]
RSV-RV モデル							
前期	0.913 (0.016) [0.14; 25.1]	0.327 (0.024) [0.10; 51.8]	-0.336 (0.059) [0.24; 26.9]	0.155 (0.126) [0.95; 7.2]	-1.024 (0.053) [0.65; 27.7]	0.430 (0.020) [0.15; 36.7]	
後期	0.920 (0.015) [0.91; 15.5]	0.296 (0.023) [0.40; 34.3]	-0.319 (0.058) [0.68; 19.2]	0.032 (0.127) [0.95; 7.0]	-1.022 (0.055) [0.83; 32.5]	0.432 (0.017) [0.71; 19.5]	
RSVt-RV モデル							
前期	0.918 (0.016) [0.55; 26.9]	0.314 (0.025) [0.43; 53.6]	-0.355 (0.062) [0.64; 35.7]	0.174 (0.130) [0.19; 6.0]	-1.035 (0.060) [0.99; 27.7]	0.437 (0.020) [0.83; 34.5]	21.154 (4.450) [0.73; 105.3]
後期	0.925 (0.014) [0.06; 7.7]	0.286 (0.021) [0.15; 17.4]	-0.328 (0.059) [0.89; 8.7]	0.057 (0.133) [0.33; 6.9]	-1.043 (0.059) [0.29; 29.9]	0.437 (0.017) [0.48; 7.7]	20.258 (4.650) [0.29; 59.6]
RSV-RK モデル							
前期	0.907 (0.017) [0.20; 19.4]	0.339 (0.026) [0.63; 42.8]	-0.300 (0.056) [0.73; 22.8]	0.132 (0.126) [0.25; 13.1]	-0.992 (0.056) [0.20; 47.2]	0.411 (0.021) [0.93; 25.5]	
後期	0.927 (0.014) [0.30; 14.6]	0.278 (0.020) [0.22; 37.6]	-0.305 (0.058) [0.09; 16.9]	0.045 (0.130) [0.55; 6.5]	-1.040 (0.056) [0.45; 31.2]	0.394 (0.016) [0.70; 17.9]	
RSVt-RK モデル							
前期	0.912 (0.016) [0.99; 13.7]	0.328 (0.024) [0.75; 31.6]	-0.319 (0.060) [0.87; 19.6]	0.167 (0.131) [0.28; 10.6]	-1.024 (0.060) [0.24; 41.0]	0.418 (0.020) [0.45; 20.7]	21.376 (4.825) [0.11; 52.5]
後期	0.931 (0.013) [0.92; 7.2]	0.271 (0.019) [0.95; 19.3]	-0.310 (0.058) [0.43; 14.2]	0.063 (0.134) [0.16; 3.0]	-1.055 (0.059) [0.33; 18.5]	0.398 (0.015) [0.63; 12.1]	20.335 (4.390) [0.42; 59.1]

注) 丸括弧内は事後標準偏差。角括弧内は左から Geweke(1992)の CD 統計量の p 値, 非効率性因子。期間は, 前期が 2013/11/14 日から 2017/12/13 日まで, 後期が 2017/8/22 日から 2021/9/30 日まで。

るパラメータであり、推定値はどのモデルでも負となっていることから、日経 225 株価指数のボラティリティと前日のリターンには負の相関が存在することが示唆される。ξ は RV の観測方程式の定数項であり、REGARCH モデルと同様負となっている。サンプルの前期と後期でパラメータの推定値を比べると、ほとんど変わらないことがわかる。RSV モデルでは、 σ と μ が多少、前期と後期で異なった値となっているが、それ以外のパラメータはほぼ同じ水準となっている。なお、RSV モデルは SV モデルに比べて、パラメータの推定の際に RV の情報が追加されているため、共通のパラメータの事後標準偏差が小さくなっている。t 分布モデルの自由度 ν の事後平均は、15 から 20 程度の大きさとなっており、SVt モデルの方が RSVt モデルよりもやや小さめとなっている。

表 4 には、MCMC 法によって発生させた事後密度分布からのパラメータのサンプルについて、サンプルの分布が事後分布に収束しているかどうかを検定する、Geweke(1992)の収束診断(Convergence Diagnostic, CD)統計量の p 値を示している。具体的には、10,000 回のサンプリングのうち、最初の 1,000 個のサンプルと後半の 5,000 個のサンプルの平均が等しいという帰無仮説を検定し、その p 値を計算している。結果をみると、どのパラメータについても、0.05 を超えており、5% 有意水準でサンプル系列の最初と後半の平均が等しいという帰無仮説は棄却されない。すなわち、サンプルは事後密度分布に収束していると判断してよい。また、表では、非効率性因子(inefficiency factor)も示している(例えば、大森・渡部(2013)を参照)。非効率性因子は無相関の標本から計算する標本平均と同じ分散を得るために何倍のサンプル数が必要であるかを表す効率性の尺度である。SV モデルの一部のパラメータを除いて、非効率性因子は 100 を下回っており、比較的、効率的なサンプリングが行われていると考えることができる。

4.3 予測精度の結果

予測の精度を評価するためには、1 日先の真のボラティリティが必要となるが、観測できない値であるため、代理変数として、30 分おきの日中リターンの 2 乗和で計算した RV を用いる。また、頑健性のチェックとして、Barndorff-Nielsen et al.(2011)の Realized Kernel(RK)も用いる。ただし、これらの系列には、昼休みや夜間休みのリターンが含まれていないため、1 日のボラティリティを過小評価することになる。そこで、第 2 節で説明した Hansen and Lunde(2005a)の方法を用いて、休憩時間の影響を補正する。

ボラティリティの予測精度の評価基準として、平均二乗誤差(Mean Square Error; MSE)と疑似尤度(Quasi-likelihood; QLIKE)の 2 つを用いる。Patton(2011)は、代理変数が真のボラティリティの条件付き不偏推定量であれば、これら 2 つの評価基準は真のボラティリティを用いた場合と同じ順位になることを理論的に示している。第 $t+1$ 日の真のボラティリティを σ_{t+1}^2 、第 t 日までのデータを用いて計算した第 $t+1$ 日のボラティリティの予測値を $\hat{\sigma}_{t+1|t}^2$ とおく。2 つの評価基準である損失関数は次式で表される。

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^N (\hat{\sigma}_{t+1|t}^2 - \sigma_{t+1}^2)^2,$$

$$\text{QLIKE} = \frac{1}{N} \sum_{t=1}^N \left(\log \hat{\sigma}_{t+1|t}^2 + \frac{\sigma_{t+1}^2}{\hat{\sigma}_{t+1|t}^2} \right).$$

ただし、ここで $t=1, \dots, N$ は予測を行う期間である。なお、これらの式からわかる通り、異常値があった場合、MSE はそれに影響されやすい一方、QLIKE の方が比較的、安定しているといえる。

各モデルから得られた予測値について、MSE と QLIKE の結果をまとめたのが表 5 である。なお、HAR-CJ モデルについては、前述の通りジャンプ成分が有意でなかったため、予測精度の比較から除外している。また、EGARCH, SV, RSV モデルについて、正規分布のモデルより t 分布のモデルの方が全体的に予測精度が良い結果となったため、ここでは t 分

表5. 予測精度のモデル比較

評価基準 代理変数	MSE		QLIKE	
	RV(30分)	RK	RV(30分)	RK
(1) 全期間				
EGARCHt	17.2882	8.9115	1.1079*	1.1584*
SVt	17.4268	8.9323	1.0936	1.1489*
HAR-RV	16.4735	6.7064	1.0345	1.0796
HAR-RK	15.5650	6.2098	1.0319	1.0796
HAR-C	16.8388	6.8973	1.0322	1.0798
REGARCHt-RV	15.0913	6.3513	1.0409	1.0791
REGARCHt-RK	14.7369	6.1854	1.0363	1.0772
RSVt-RV	15.6916	6.6814	1.0349	1.0774
RSVt-RK	15.3874	6.4641	1.0323	1.0776
(2) コロナ前				
EGARCHt	2.1175	1.2035	0.8800*	0.9095*
SVt	1.9474	1.0456	0.8500*	0.8837*
HAR-RV	2.0344	1.1544	0.8247	0.8514*
HAR-RK	2.0101	1.1275	0.8210	0.8483
HAR-C	2.0389	1.1541	0.8234	0.8502
REGARCHt-RV	2.0298	1.0838	0.8326	0.8567*
REGARCHt-RK	1.9898	1.0496	0.8272	0.8519
RSVt-RV	1.9217	0.9785	0.8161	0.8443
RSVt-RK	1.9018	0.9714	0.8103	0.8399
(3) コロナ発生直後・継続期間				
EGARCHt	35.2629	18.0442	1.3778*	1.4533*
SVt	35.7674	18.2768	1.3823*	1.4631*
HAR-RV	33.5814	13.2846	1.2831	1.3501
HAR-RK	31.6253	12.2314	1.2817	1.3537
HAR-C	34.3742	13.7020	1.2797	1.3519
REGARCHt-RV	30.5671	12.5925	1.2878*	1.3426
REGARCHt-RK	29.8401	12.2705	1.2840	1.3440
RSVt-RV	32.0066	13.4383	1.2942	1.3535
RSVt-RK	31.3657	12.9721	1.2953	1.3592
(4) コロナ継続期間				
EGARCHt	1.7606	1.6868	1.1554*	1.2500*
SVt	1.5343	1.4650	1.1494*	1.2591*
HAR-RV	1.4380	1.2976	1.0799	1.1619
HAR-RK	1.4405	1.3013	1.0759	1.1638
HAR-C	1.4596	1.3124	1.0794	1.1673
REGARCHt-RV	1.5436	1.3730	1.0825	1.1553
REGARCHt-RK	1.5074	1.3424	1.0780	1.1569
RSVt-RV	1.4511	1.2848	1.0811	1.1627
RSVt-RK	1.4530	1.2911	1.0798	1.1680

注) 真のボラティリティの代理変数として、RV(30分間隔の日中リターンにより算出)とRKを使用。予測期間については、(1)全期間は2017/12/14日から2021/9/30日まで(1,000日間)、(2)コロナ前は2017/12/14日から2020/1/7日まで(500日間)、(3)コロナ発生直後・継続期間は2020/1/8日から2021/9/30日まで(422日間)、(4)コロナ継続期間は2020/5/14日から2021/9/30日まで(339日間)。*は10%有意水準でMCSに含まれないモデル。

布の結果のみを報告している。まず、全期間をみると、MSEについては、RV(30分)の場合、REGARCHt-RKモデルが、RKの場合、HAR-RKモデルが最も優れていることがわかる。一方、QLIKEについては、RV(30分)の場合、HAR-RKモデルが、RKの場合、REGARCHt-RKモデルが最も優れていることがわかる。

この結果を期間別に仔細にみると、コロナ前の期間については、どの場合もRSVt-RKモデルが最も優れていることがわかる。また、コロナ発生直後・継続期間については、REGARCHt-RKモデル、HAR-RKモデル、HAR-Cモデルが最も優れたモデルとして挙げられる。一方、コロナ継続期間についてみると、HARモデル、REGARCHモデル、RSVtモデルと様々なモデルが最も優れたモデルとなっている。これらの結果から、コロナ発生直後のボラティリティが非常に高い時期を含む期間では、REGARCHtモデルやHARモデルが最も優れたモデルとなった。一方、こうした時期を除くと、コロナ前でもコロナ継続期間でも、一貫してRSVtモデルが最も優れたモデルになることがわかった。この背景としては、RSVtモデルはREGARCHtモデルやHARモデルに比べて潜在変数が多く、自由度の高いモデルとなっており、通常の時期については予測精度が高いが、ボラティリティが非常に高い時期には潜在変数が多いために予測値が振れやすく、REGARCHtモデルやHARモデルの方が予測精度がやや高まる傾向にあるのではないかと考えられる。

さて、こうしたMSEとQLIKEのモデル間の差が統計的に意味のあるものかどうか、2種類の検定方法を用いて、検証する。1つ目は、Hansen et al.(2011)によって提案された、モデル信頼集合(Model Confidence Set; MSC)である。MCSを用いると、真のモデルを仮定することなく、他のモデルより予測精度が優れているモデルの集合を特定することができる。検証の対象となっているモデルの集合を M_0 とする。その中から2つのモデル i と j を取り出し、各時点の損失関数のモデル間の差を $d_{ij,t}$ とおく。

M_0 の部分集合 M_1 について ($M_1 \subset M_0$), 次式の帰無仮説を検定する.

$$H_0: E[d_{ij,t}] = 0, \quad \forall i, j \in M_1, i > j. \quad (30)$$

この帰無仮説が棄却された場合, 予測力が有意に低いモデルを集合 M_1 から除外する. 次々とこの帰無仮説による検定を行い, M_1 に含まれる全ての (i, j) について帰無仮説が棄却されなくなるまで行う. 式(30)に対する検定統計量は, ブートストラップ法を用いて計算される(詳しくは, Hansen et al.(2011)を参照されたい). 本稿の計算では, RのパッケージMCSを用いて計算した.

検定方法の2つ目は, Giacomini and White (2006)の方法(以下, GW 検定と呼ぶ)である. 候補となっているモデルの全ての2つずつの組み合わせについて, MSCの計算と同様に各時点の損失関数のモデル間の差 $d_{ij,t}$ を考える. GW 検定には2種類あり, 無条件(unconditional)検定と条件付き(conditional)検定がある.

まず, 無条件GW 検定については, MSCと同様に上記の帰無仮説 H_0 を検定する. Giacomini and White(2006)に倣って, 検定統計量の算出にはNewey-WestのHAC(Heteroscedasticity and Auto-correlation Consistent)推定量を用いる. 次に, 条件付きGW 検定については, 帰無仮説を次式に替える.

$$H_0: E[d_{ij,t+1}|I_t] = 0. \quad (31)$$

ここで, I_t は t 時点までの情報を表しており, t までの情報から予測された $t+1$ 時点の損失関数の差について, 期待値が0であるかどうかを検定する. Takahashi et al.(2021)に倣って, $q_t = (1, d_{ij,t})'$ とおき, 次式の帰無仮説を検定する.

$$H_0: E[q_t d_{ij,t+1}] = 0. \quad (32)$$

検定統計量の計算について詳しくは, Takahashi et al.(2021)を参照されたい. なお, Patton and Sheppard(2009)は, MSE よりも QLIKEの方がGW 検定の検出力が高いことを示して

いるため, 本稿でもQLIKE についてのみGW 検定を行う.

全期間とコロナ前後の期間で, MCSの検定を行った結果が表5に記載されている. 10%の有意水準で, MCSに含まれないモデルに「*」を付けている. MSEを基準にすると, どのモデルもMCSに含まれており, 有意な違いがないことがわかる. 一方, QLIKEを基準にすると, SVモデルは, どの推計期間でも, どちらの真のボラティリティの代理変数でも, MCSに含まれない. また, EGARCHモデルも全期間やコロナ発生直後・継続期間などでMCSに含まれない. この2つのモデルはRVを用いていないモデルであり, それ以外のモデルはRVを用いている. つまり, RVを用いた方が予測精度は統計的に有意に優れていることがわかった.

また, この結果は表6に示されたGW 検定の結果からもわかる. ここでは, 推定期間がデータ全期間の結果のみを示す. ボラティリティの代理変数をRV, RKとした場合の両方とも, EGARCHtモデルとSVtモデルは他のモデルよりもQLIKEが高めに出る傾向にあり, その差は5%有意水準で統計的に有意といえる.

表5のMSEとQLIKEの値から, 以下のことがわかる. (1) RVやRKを使わないEGARCH, SVモデルは, RVもしくはRKを使うHAR, HAR-C, REGARCH, RSVと比べて, 統計的に有意にボラティリティの予測精度が低い. (2) 統計的に有意ではないが, コロナの発生直後を含む期間ではHARモデルやREGARCHモデルの予測精度が比較的高く, それ以外の期間についてはコロナ前もコロナ継続期間もRSVモデルの予測精度が高い. (3) 統計的に有意ではないが, ほとんどの場合でHAR, REGARCH, RSVモデルではRVよりもRKを用いた方が予測精度が高い. (4) RVの連続成分を用いたHAR-Cモデルは, RVを用いたHARモデルの予測精度を上回らない.

表 6. QLIKE における GW 検定の結果(全予測期間).

1. 代理変数が RV(30) の場合

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)EGARCHt		1.00**	0.93**	0.94**	0.93**	0.91**	0.92**	0.94**	0.92**
(2)SVt	-1.45		0.95**	0.96**	0.95**	0.91**	0.92**	0.94**	0.93**
(3)HAR-RV	-4.07**	-3.37**		0.86	0.97	0.06	0.25	0.48	0.65
(4)HAR-RK	-4.11**	-3.50**	-0.94		0.36	0.01	0.07	0.00	0.48
(5)HAR-C	-4.04**	-3.39**	-1.11	0.10		0.05	0.17	0.26	0.53
(6)REGARCHt-RV	-3.89**	-3.11**	0.99	1.39	1.30		0.94*	0.87	0.85
(7)REGARCHt-RK	-4.00**	-3.30**	0.25	0.73	0.58	-1.93*		0.83	0.82
(8)RSVt-RV	-3.98**	-3.71**	0.05	0.47	0.36	-0.92	-0.21		0.75
(9)RSVt-RK	-3.87**	-3.64**	-0.26	0.06	0.01	-1.08	-0.57	-0.68	

2. 代理変数が RK の場合

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)EGARCHt		0.99	0.92**	0.93**	0.93**	0.88**	0.90**	0.93**	0.95**
(2)SVt	-1.10		0.92**	0.94**	0.92**	0.92**	0.92**	0.93**	0.95**
(3)HAR-RV	-4.20**	-3.99**		0.56	0.14	0.47	0.73	0.99	0.86
(4)HAR-RK	-3.99**	-3.84**	-0.00		0.05	0.50	0.72	0.89	0.98
(5)HAR-C	-3.98**	-3.80**	0.09	0.08		0.56	0.80	0.90	1.00
(6)REGARCHt-RV	-4.61**	-4.17**	-0.09	-0.07	-0.10		1.00	0.98	0.95
(7)REGARCHt-RK	-4.55**	-4.22**	-0.50	-0.40	-0.44	-0.69		0.00	0.39
(8)RSVt-RV	-4.52**	-4.46**	-0.54	-0.41	-0.44	-0.35	0.05		0.25
(9)RSVt-RK	-4.18**	-4.17**	-0.45	-0.54	-0.48	-0.19	0.07	0.05	

注) 左下の三角行列の値は、無条件 GW 検定の検定統計量(t 値)。値が正の場合、行のモデルの損失関数が列のモデルのそれよりも大きいことを示す。右上の三角行列の値は、条件付き GW 検定における予測損失関数の差(行のモデルから列のモデルの損失を差し引いたもの)が正となる割合。*, * はそれぞれ 1%, 5% の水準で統計的に有意。

5. 結論

本稿では、まず、Realized Volatility(RV)に関する先行研究のサーベイを行った。次に、日経 225 株価指数の RV を用いて、RV を用いた主要なボラティリティモデルについて、それぞれ推定を行ったうえで、ボラティリティの予測精度についてモデル比較を行った。分析の特徴として、できるだけ多くのモデルについて実際の日本のデータを用いて比較した。また、予測期間を新型コロナウイルス感染症の拡大前後で分け、優れているモデルが時期によって異なるかどうかを検証した。分析の結果、コロナの発生直後を含む期間では、HAR モデルや REGARCH モデルの予測精度が比較的高く、それ以外の期間についてはコロナ前もコロナ継続期間も RSV モデルの予測精度が高いことが明らかになった。

本稿には以下のようにさらなる拡張の余地がある。

- (1) 本稿ではリターンの分布に正規分布と t 分布を仮定した。これらは左右対称な分布であり、非対称な分布も含めて包括的にモデル比較を行うことは重要である。非対称分布については、Aas and Haff(2006)を参照されたい。中島・大森(2011), Nakajima and Omori(2012), Takahashi et al.(2016)は SV モデルや RSV モデルにおいて一般化双曲型非対称 t 分布, Abanto-Valle et al.(2015)は SV モデルにおいて Azzalini and Capitanio(2003)の非対称 t 分布, Watanabe(2012)は RGARCH モデルにおいて Fernández and Steel(1998)の非対称 t 分布を用いている。こうしたリターンの分布は、VaR や ES などに応用する場合、特に重要である。
- (2) 本稿では 1 日先のボラティリティの予測だけを行ったが、週次、月次など将来の多期間のボラティリティの予測の比較も重要である。多期間のボラティリティを予測する場合、EGARCH モデルや REGARCH モデルでも

シミュレーションが必要になるのに対して、HAR モデルでは被説明変数を多期間のボラティリティにするだけで良い。詳しくは、渡部(2020)を参照されたい。

- (3) 本稿では Realized Measure(RM)に5分ごとのリターンを使って計算したRVとRKを用いたが、その他のRMを用いた場合やREGARCHモデルやRSVモデルで異なる複数のRMを同時に用いた場合との比較は興味深い。これについては、Hansen and Huang(2016)や石原(2015)を参照されたい。
- (4) 本稿ではEGARCHモデルとREGARCHモデルでボラティリティの予測を行う際にパラメータを疑似最尤推定値として予測を行ったが、MCMCによってパラメータとボラティリティの予測値を同時に事後分布からサンプリングした場合、予測精度が高まるかどうかにも興味深い。EGARCHモデルやREGARCHモデルのMCMCを用いたベイズ推定については、三井・渡部(2003)、Asai(2006)、Chen et al.(2021)を参照されたい。
- (5) オプション価格は原資産価格のボラティリティに依存するので、オプション価格への応用も重要である。オプション価格を用いた比較については、Bandi et al.(2008)やUbukata and Watanabe(2014a)を参照されたい。
- (6) ボラティリティにはオプション価格から計算されるインプライドボラティリティがあり、近年ではBlack and Scholes(1973)モデルを仮定せず、ボラティリティの変動を許容するジャンプ拡散過程(10)式の下で計算されるモデルフリー・インプライドボラティリティが用いられる。これは危険中立測度の下でのボラティリティであり、そこから過去のリターンから予測される実測度の下でのボラティリティを引いたものは分散リスクプレミアムと呼ばれ、将来のリターンや景気の予測に役立つ変数として注目を集めている。モデルフリー・インプライドボラティリティについてはJiang and Tian(2005)、杉原(2010)、分散リスクプレミアムとその予測力についてはBollerslev et al.(2009, 2014)、Bekaert and

Hoerova(2014)、Ubukata and Watanabe(2014b)を参照されたい。実測度の下でのボラティリティは本稿で説明したモデルで予測する必要があるので、分散プレミアムの予測力の比較も興味深い。

- (7) 株価のボラティリティに関する別の論点として、先行研究では日本銀行が行っているETF買い入れが株価市場のリスクプレミアムを押し下げる効果があるという結果が提示されている(安達他(2021)を参照されたい)。本稿の分析の枠組みを用いてこうしたETF買い入れの効果を分析することも有用であると考えられる。

これらは今後の課題としたい。

(一橋大学経済研究所
・一橋大学経済研究所)

補論. SVモデル, RSVモデルの事後分布からのサンプリング法

本稿では、SVモデルとRSVモデルをベイズ推定の枠組みでMCMC法により推定した。この補論では、これらのモデルについて事後分布からのサンプリング法を説明する。基本的にRSVモデルの事後分布からのサンプリング方法について説明し、SVモデルはその特殊ケースとして考える。

ω の事前分布を $\pi(\omega)$ とおく。RSVモデルのパラメータ ω と対数ボラティリティ h の同時事後分布は、次式で表される。

$$\begin{aligned} \pi(\omega, h, \lambda | Y) &\propto \pi(\omega) \prod_{t=1}^T \frac{1}{\sqrt{\lambda_t} e^{h_t/2}} \exp\left(-\frac{y_t^2}{2\lambda_t e^{h_t}}\right) \\ &\times \frac{\sqrt{1-\phi^2}}{\sigma} \exp\left\{-\frac{(1-\phi^2)\bar{h}_1^2}{2\sigma^2}\right\} \\ &\times \prod_{t=1}^{T-1} \frac{1}{\sigma\sqrt{1-\rho^2}} \\ &\exp\left\{-\frac{(\bar{h}_{t+1}-\phi\bar{h}_t-\bar{y}_t)^2}{2\sigma^2(1-\rho^2)}\right\} \\ &\times \prod_{t=1}^T \frac{((\nu-2)/2)^{\nu/2}}{\Gamma(\nu/2)} \lambda_t^{-\frac{\nu+1}{2}-1} \\ &\exp\left(-\frac{\nu-2}{2\lambda_t}\right) \\ &\times \prod_{t=1}^T \frac{1}{\sigma_u} \exp\left\{-\frac{(x_t-\xi-h_t)^2}{2\sigma_u^2}\right\}, \end{aligned}$$

ただし, $\bar{h}_t = h_t - \mu$, $\bar{y}_t = \rho \sigma y_t e^{-h_t/2} / \sqrt{\lambda_t}$ である.

A1. ϕ のサンプリング

ϕ の事前分布を $\pi(\phi)$ とおく. ϕ の条件付事後分布 $\pi(\phi|\cdot)$ は, 次式のように書ける.

$$\begin{aligned} \pi(\phi|\cdot) &\propto \pi(\phi) \sqrt{1-\phi^2} \exp \left\{ -\frac{(1-\phi^2) \bar{h}_1^2}{2\sigma^2} \right. \\ &\quad \left. - \sum_{t=1}^{T-1} \frac{(\bar{h}_{t+1} - \phi \bar{h}_t - \bar{y}_t)^2}{2\sigma^2(1-\rho^2)} \right\} \\ &\propto \pi(\phi) \sqrt{1-\phi^2} \exp \left\{ -\frac{(\phi - \mu_\phi)^2}{2\sigma_\phi^2} \right\}, \quad (33) \end{aligned}$$

ただし,

$$\mu_\phi = \frac{\sum_{t=1}^{T-1} (\bar{h}_{t+1} - \bar{y}_t) \bar{h}_t}{s_h^2}, \quad \sigma_\phi^2 = \frac{\sigma^2(1-\rho^2)}{s_h^2},$$

および, $s_h^2 = \rho^2 \bar{h}_1^2 + \sum_{t=2}^{T-1} \bar{h}_t^2$ である. この条件付事後分布はシンプルな分布の密度関数の形をしていないため, 直接サンプリングすることはできないが, Metropolis Hasting(MH) アルゴリズムを用いて次のようにサンプリングを行うことができる. まず, (33)式の一部が正規分布の密度関数の核の形をしていることに注目して, 提案分布を $TN(-1, 1)(\mu_\phi, \sigma_\phi^2)$ において候補点 ϕ^* を発生させる. ただし, $TN(a, b)(\mu, \sigma^2)$ は平均 μ , 分散 σ^2 の定義域 (a, b) に限定された切断正規分布を表す. 現在の点 ϕ からこの候補点 ϕ^* を次の確率で受容する.

$$\min \left\{ \frac{\pi(\phi^*) \sqrt{1-\phi^{*2}}}{\pi(\phi) \sqrt{1-\phi^2}}, 1 \right\}.$$

棄却した場合には, 現在の点に留まる.

A2. (σ, ρ) のサンプリング

σ と ρ の同時条件付事後分布を考える. $\mathbf{g} = (\sigma, \rho)'$ と定義して, 事前分布を $\pi(\mathbf{g})$ と設定すると, $\mathbf{g}$ の条件付事後分布は次式で表される.

$$\begin{aligned} \pi(\mathbf{g}|\cdot) &\propto \pi(\mathbf{g}) \times \sigma^T (1-\rho^2)^{\frac{T-1}{2}} \exp \\ &\quad \left\{ -\frac{(1-\phi^2) \bar{h}_1^2}{2\sigma^2} - \sum_{t=1}^{T-1} \frac{(\bar{h}_{t+1} - \phi \bar{h}_t - \bar{y}_t)^2}{2\sigma^2(1-\rho^2)} \right\}. \quad (34) \end{aligned}$$

この条件付事後確率密度関数からも確率標本を直接サンプリングすることはできないため, MH アルゴリズムを用いる. パラメータ (σ, ρ) には制約 $R = \{\mathbf{g} : \sigma > 0, |\rho| < 1\}$ がある. このまま提案分布を考えてもよいが, ここでは $\mathbf{g}$ から $\mathbf{w} \equiv (w_1, w_2)'$, $w_1 = \log \sigma$, $w_2 = \log(1+\rho) - \log(1-\rho)$ への変数変換を

行う. この変換により, $\mathbf{w}$ の定義域は $R' = \{-\infty < w_i < \infty; i=1, 2\}$ となる. 次に条件付事後確率密度関数(34)のモード(またはモードの近似値)を数値的に求め, その値を $\hat{\mathbf{g}}$ とおき, また, 上の変数変換を施した値を $\hat{\mathbf{w}}$ とする. 提案分布として $N(\mathbf{w}_*, \mathbf{Q}_*)$ を用いて, 候補点 $\mathbf{w}^*$ を発生させる. ただし,

$$\begin{aligned} \mathbf{w}_* &= \hat{\mathbf{w}} + \mathbf{Q}_* \frac{\partial \log \pi(\mathbf{w}|\cdot)}{\partial \mathbf{w}} \Big|_{\mathbf{w}=\hat{\mathbf{w}}}, \\ \mathbf{Q}_*^{-1} &= -\frac{\partial^2 \log \pi(\mathbf{w}|\cdot)}{\partial \mathbf{w} \partial \mathbf{w}'} \Big|_{\mathbf{w}=\hat{\mathbf{w}}}. \end{aligned}$$

この候補点 $\mathbf{w}^*$ を次の確率で受容する.

$$\min \left\{ \frac{\pi(\mathbf{g}^*|\cdot) f_N(\mathbf{w}|\mathbf{w}_*, \mathbf{Q}_*) |J(\mathbf{g})|}{\pi(\mathbf{g}|\cdot) f_N(\mathbf{w}^*|\mathbf{w}_*, \mathbf{Q}_*) |J(\mathbf{g}^*)|}, 1 \right\}.$$

ただし, $f_N(\mathbf{x}|\mathbf{m}, \mathbf{\Omega})$ は平均ベクトル $\mathbf{m}$, 分散共分散行列 $\mathbf{\Omega}$ の二変量正規分布の密度関数を点 $\mathbf{x}$ で評価したものであり, $J(\cdot)$ は上の変数変換のヤコビアンを表す. $\mathbf{g}^*$ は $\mathbf{w}^*$ を逆変換した点である.

A3. μ のサンプリング

μ の事前分布を $N(\mu_0, v_0^2)$ とおく. μ の条件付事後分布は次式で表される.

$$\begin{aligned} \pi(\mu|\cdot) &\propto \exp \left\{ -\frac{(\mu - \mu_0)^2}{2v_0^2} - \frac{(1-\phi^2) \bar{h}_1^2}{2\sigma^2} \right. \\ &\quad \left. - \sum_{t=1}^{T-1} \frac{(\bar{h}_{t+1} - \mu) - \phi(h_t - \mu) - \bar{y}_t}{2\sigma^2(1-\rho^2)} \right\} \\ &\propto \exp \left\{ -\frac{(\mu - \hat{\mu})^2}{2\sigma_\mu^2} \right\}. \end{aligned}$$

ただし,

$$\begin{aligned} \sigma_\mu^2 &= \left\{ \frac{1}{v_0^2} + \frac{(1-\rho^2)(1-\phi^2) + (T-1)(1-\phi)^2}{\sigma^2(1-\rho^2)} \right\}^{-1}, \\ \hat{\mu} &= \sigma_\mu^2 \left\{ \frac{\mu_0}{v_0^2} + \frac{(1-\rho^2)(1-\phi^2)h_1 + (1-\phi)\sum_{t=1}^{T-1}(h_{t+1} - \phi h_t - \bar{y}_t)}{\sigma^2(1-\rho^2)} \right\}. \end{aligned}$$

この条件付事後確率密度関数は正規分布の密度関数の核の形をしているので, サンプリングは, $\mu|\cdot \sim N(\hat{\mu}, \sigma_\mu^2)$ となる.

A4. $\mathbf{h}$ のサンプリング

対数ボラティリティ $(h_1, \dots, h_T)$ を効率的にサンプリングするために, multi-move sampler と呼ばれるサンプリング法を用いる. $\mathbf{h}$ の条件付事後分布 $\pi(\mathbf{h}|\omega, \mathbf{Y})$ から, 一度に $\mathbf{h}$ を直接サンプリングす

ることは難しい。そこで、 $(h_1, \dots, h_T)$ をいくつかのブロックに分けて、1つのブロック内の潜在変数のサンプリングを、他のブロックの h_t を条件付けた上で行うという考え方である。

$t=0, \dots, T$ に対して $\alpha_t = h_t - \mu$, および $\tilde{y}_t = y_t \exp(-\mu/2)/\sqrt{\lambda_t}$ と定義し、 $\{\alpha_t\}_{t=1}^T$ を潜在変数とする次の状態空間モデルを考える。

$$\tilde{y}_t = \varepsilon_t \exp(\alpha_t/2), \quad t = 1, \dots, T, \quad (35)$$

$$x_t = \xi + \mu + \alpha_t + u_t, \quad t = 1, \dots, T, \quad (36)$$

$$\alpha_{t+1} = \phi \alpha_t + \eta_t, \quad t = 0, \dots, T-1. \quad (37)$$

$(\alpha_1, \dots, \alpha_T)$ 全体を $K+1$ 個のブロックに分割し、任意のブロックを $(\alpha_{k_{i-1}+1}, \dots, \alpha_{k_i})$ と表す。ただし、 $i=1, \dots, K+1, k_0=0, k_{K+1}=n$ である。ブロックの分割方法は、例えば、Shephard and Pitt(1997)が提案している stochastic knots という方法があり、 $i=1, \dots, K$ に対して $k_i = \text{int}[n(i+U_i)/(K+2)]$ と定める。ただし、 U_i は区間 $[0, 1]$ で定義される一様分布からの乱数を表し、 $\text{int}[\cdot]$ は整数部分を表す関数である。MCMC 法のアプローチにおいて、ブロックの節目である $(k_1, \dots, k_K)$ を毎回ランダムに決めることにより、サンプリングの効率性を高めることができる。

任意のブロック $(\alpha_{k_{i-1}+1}, \dots, \alpha_{k_i})$ の条件付同時事後分布を導出し、サンプリングを行う。具体的な方法は、Takahashi et al.(2016)を参照されたい。

A5. λ のサンプリング

λ_t の条件付事後分布 $\pi(\lambda_t|\cdot)$ は、次式のように書ける。

$$\pi(\lambda_t|\cdot) \propto \lambda_t^{\frac{\nu+1}{2}-1} \exp\left(-\frac{\nu-2}{2\lambda_t}\right) \times f(\lambda_t),$$

ただし、

$$f(\lambda_t) = \begin{cases} \exp\left(-\frac{y_t^2}{2\lambda_t e^{h_t}}\right) \exp\left\{-\frac{(\bar{h}_{t+1} - \phi \bar{h}_t - \bar{y}_t)^2}{2\sigma^2(1-\rho^2)}\right\} & (t = 1, \dots, T-1 \text{ のとき}) \\ \exp\left(-\frac{y_t^2}{2\lambda_t e^{h_t}}\right) & (t = T \text{ のとき}), \end{cases}$$

λ_t も MH アルゴリズムを用いてサンプリングを行う。まず、提案分布を $\lambda_t^* \sim IG((\nu+1)/2, (\nu-2)/2)$ とし、候補点を発生させ、現在の点 λ_t から次の候補点 λ_t^* を次の確率で受容する。

$$\min\left\{\frac{f(\lambda_t^*)}{f(\lambda_t)}, 1\right\}.$$

これを、 $t=1, \dots, T$ に対して行う。

A6. ν のサンプリング

ν の事前分布を $\pi(\nu)$ とおく。 ν の条件付事後分布 $\pi(\nu|\cdot)$ は、次式のように書ける。

$$\pi(\nu|\cdot) \propto \pi(\nu) \prod_{t=1}^T \frac{((\nu-2)/2)^{\nu/2}}{\Gamma(\nu/2)} \lambda_t^{-\nu/2} \exp\left(-\frac{\nu-2}{2\lambda_t}\right).$$

ν についても MH アルゴリズムを用いてサンプリングする。具体的には、ステップ A2 と同様に正規分布による条件付事後分布の近似を用いて、候補点をサンプリングする。

A7. ξ のサンプリング

ξ の事前分布を、 $\xi \sim N(\xi_0, z_0^2)$ とおく。 ξ の条件付事後分布は $\xi|\cdot \sim N(\tilde{\xi}, \tilde{z}_2)$ となり、 $\tilde{\xi}$ と $\tilde{z}_2$ はそれぞれ次式で表される。

$$\tilde{z}_2 = \left(\frac{1}{z_0^2} + \frac{T}{\sigma_u^2}\right)^{-1}, \quad \tilde{\xi} = \tilde{z}_2 \left(\frac{\xi_0}{z_0^2} + \frac{\sum_{t=1}^T (x_t - h_t)}{\sigma_u^2}\right).$$

A8. σ_u のサンプリング

σ_u の事前分布を、 $\sigma_u^2 \sim IG(n_0/2, V_0/2)$ とおく。ただし、 IG は逆ガンマ (Inverse Gamma) 分布である。 σ_u の条件付事後分布は $\sigma_u^2|\cdot \sim IG(\hat{n}/2, \hat{V}/2)$ となり、 $\hat{n}$ と $\hat{V}$ はそれぞれ次式で表される。

$$\hat{n} = n_0 + T, \quad \hat{V} = V_0 + \sum_{t=1}^T (x_t - \xi - h_t)^2.$$

A9. SV モデルの事後分布からのサンプリング

以上の A1 から A8 までのステップが RSV モデルの事後分布からのサンプリングとなる。SV モデルについては、A1 から A6 までのステップがサンプリングの方法となる。ただし、A4 については、式(36)を除いて、multi-move sampler を構築する。

注

* 本研究は JSPS 科研費 JP19H00588, JP20H00073 と一橋大学社会科学高等研究院から助成を受けている。一橋大学経済研究所定例研究会参加者

および討論者の石原庸博氏から多くの貴重なコメントを頂いた。ここに記して感謝の意を表したい。

1) その他の方法に、Xiu(2010), Kunitomo and Sato(2013)らの疑似尤度に基づく方法がある。

2) 本稿では用いないが、RV の対数を取らない以下の定式化も用いられる。

$$\begin{aligned} RV_t &= \alpha + \beta_d RV_{t-1} + \beta_w RV_{t-5:t-1} + \beta_m RV_{t-22:t-1} + v_t, \\ \sqrt{RV_t} &= \alpha + \beta_d \sqrt{RV_{t-1}} + \beta_w \sqrt{RV_{t-5:t-1}} \\ &\quad + \beta_m \sqrt{RV_{t-22:t-1}} + v_t. \end{aligned}$$

これらの定式化については、Andersen et al.(2007), Corsi et al.(2008)を参照されたい。

3) 週次、月次 RV を以下のように定義している論文もある。

$$\begin{aligned} \log RV_{t-5:t-1} &= \frac{1}{5} \sum_{i=1}^5 \log RV_{t-i}, \\ \log RV_{t-22:t-1} &= \frac{1}{22} \sum_{i=1}^{22} \log RV_{t-i}. \end{aligned}$$

参考文献

- 安達孔・北村富行・平木一浩(2021)「日本銀行のETF買入れが株式市場のリスク・プレミアムに及ぼす影響」日本銀行ワーキングペーパーシリーズ, 21-J-6.
- 石原庸博(2015)「一般化した Realized Stochastic Volatility モデルの推定-Nikkei225 収益率・実現ボラティリティの暦効果への応用」『経済研究』66(1), 1-18.
- 伊庭幸人・種村正美・大森裕浩・和合肇・佐藤整尚・高橋明彦(2005)「統計科学のフロンティア 12 計算統計 II マルコフ連鎖モンテカルロ法とその周辺」東洋経済新報社.
- 大森裕浩・渡部敏明(2013)「Realized Stochastic Volatility モデル—マルコフ連鎖モンテカルロ法を用いたベイズ分析」『日本統計学会誌』42(2), 273-303.
- 杉原慶彦(2010)「わが国株式市場のモデル・フリー・インプライド・ボラティリティ」『金融研究』29(2), 73-120.
- 高橋慎・大森裕浩・渡部敏明(2020)「Realized Stochastic Volatility モデル—拡張と日本の株価指数への応用—」『統計数理』68(1), 65-85.
- 竹内(野木森)明香・渡部敏明(2008)「日本の株式市場におけるボラティリティの長期記憶性とオプション価格」『現代ファイナンス』24, 45-74.
- 田中勝人(2006)『現代時系列分析』岩波書店.
- 中島上智・大森裕浩(2011)「一般化双曲型非対称 t 分布を用いた確率的ボラティリティ変動モデルの推定と株価収益率データへの応用」『日本統計学会誌』40(2), 1-28.
- 中妻照雄(2007)『入門ベイズ統計学』朝倉書店.
- 中妻照雄(2013)『実践ベイズ統計学』朝倉書店.
- 三井秀俊・渡部敏明(2003)「ベイズ推定法による GARCH オプション価格付けモデルの分析」『日本統計学会誌』33(3), 307-324.
- 矢島美寛(2003)「長期記憶をもつ時系列モデル」刈屋武昭・矢島美寛・田中勝人・竹内啓著『経済時系列の統計 その数理的基礎』第 II 部, 103-202, 岩波書店.
- 和合肇編(2005)『ベイズ計量経済分析 マルコフ連鎖モンテカルロ法とその応用』東洋経済新報社.
- 渡部敏明(2000)『ボラティリティ変動モデル』朝倉書店.
- 渡部敏明(2005)「マルチ・ムーブ・サンプラーを用いた確率的ボラティリティ変動モデルのベイズ推定」和合肇編著『ベイズ計量経済分析』第 9 章, pp. 259-294, 東洋経済新報社.
- 渡部敏明(2007)「Realized Volatility—サーベイと日本の株式市場への応用—」『経済研究』58(4), 352-373.
- 渡部敏明(2020)「Heterogeneous Autoregressive モデル—サーベイと日経 225 株価指数の実現ボラティリティへの応用—」『広島経済大学経済研究論集』42(3), 5-18.
- Aas, K. and I. H. Haff, (2006) “The Generalized Hyperbolic Skew Student’s *t*-distribution,” *Journal of Financial Econometrics*, 4(2), 275-309.
- Abanto-Valle, C. A., V. H. Lachos and D. K. Dey (2015) “Bayesian Estimation of a Skew-Student-t Stochastic Volatility Model,” *Methodology and Computing in Applied Probability*, 17(3), 721-738.
- Ait-Sahalia, Y., P. A. Mykland and L. Zhang (2005) “How Often to Sample a Continuous-time Process in the Presence of Market Microstructure Noise,” *Review of Financial Studies*, 18(2), 351-416.
- Andersen, T. G. and T. Bollerslev (1998) “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts,” *International Economic Review*, 39(4), 885-905.
- Andersen, T. G., T. Bollerslev and F. X. Diebold (2007) “Roughing it up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility,” *Review of Economics and Statistics*, 89(4), 701-720.
- Andersen, T. G., T. Bollerslev, F. X. Diebold and H. Ebens (2001a) “The Distribution of Realized Stock Return Volatility,” *Journal of Financial Economics*, 61(1), 43-76.
- Andersen, T. G., T. Bollerslev, F. X. Diebold and P. Labys (2001b) “The Distribution of Realized Exchange Rate Volatility,” *Journal of the American Statistical Association*, 96(453), 42-55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold and P. Labys (2003) “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71(2), 579-625.
- Andersen, T. G., D. Dobrev and E. Schaumburg (2012) “Jump-robust Volatility Estimation Using Nearest Neighbor Truncation,” *Journal of Econometrics*, 169(1), 75-93.
- Asai, M. (2006) “Comparison of MCMC Methods for Estimating GARCH Models,” *Journal of the Japan Statistical Society*, 36(2), 199-212.
- Asai, M., M. McAleer and M. C. Medeiros (2012) “Asymmetry and Long Memory in Volatility Modeling,” *Journal of Financial Econometrics*, 10(3), 495-512.

- Azzalini, A. and A. Capitanò (2003) "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-distribution," *Journal of the Royal Statistical Society, Series B*, 65 (2), 367-389.
- Baillie, R. T., T. Bollerslev and H. O. Mikkelsen (1996) "Fractionally Integrated Generalized Autoregressive Conditionally Heteroskedasticity," *Journal of Econometrics*, 74(1), 3-30.
- Bandi, F. M. and J. R. Russell (2006) "Separating Microstructure Noise from Volatility," *Journal of Financial Economics*, 79(3), 655-692.
- Bandi, F. M. and J. R. Russell (2008) "Microstructure Noise, Realized Volatility, and Optimal Sampling," *Review of Economic Studies*, 75(2), 339-369.
- Bandi, F. M., J. R. Russell and C. Yang (2008) "Realized Volatility Forecasting and Option Pricing," *Journal of Econometrics*, 147(1), 34-46.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde and N. Shephard (2008) "Designing Realized Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76(6), 1481-1536.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde and N. Shephard (2009) "Realized Kernels in Practice: Trades and Quotes," *Econometrics Journal*, 12(3), C1-C32.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, N. Shephard (2011) "Subsampling Realized Kernels," *Journal of Econometrics*, 160(1), 204-219.
- Barndorff-Nielsen, O. E., B. Nielsen, N. Shephard, and C. Ysusi (2004) "Measuring and Forecasting Financial Variability Using Realized Variance," A. C. Harvey, S. J. Koopman and N. Shephard (eds.), *State Space and Unobserved Components Models: Theory and Applications*, Chapter 10, 205-235, Cambridge University Press.
- Barndorff-Nielsen, O. E. and N. Shephard (2001) "Non-Gaussian OU Based Models and Some of Their Uses in Financial Economics (with discussion)," *Journal of the Royal Statistical Society, Series B*, 63 (2), 167-241.
- Barndorff-Nielsen, O. E. and N. Shephard (2002) "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64 (2), 253-280.
- Barndorff-Nielsen, O. E. and N. Shephard (2004) "Power and Bipower Variation with Stochastic Volatility and Jumps (with discussion)," *Journal of Financial Econometrics*, 2(1), 1-37.
- Barndorff-Nielsen, O. E. and N. Shephard (2006) "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4(1), 1-30.
- Bekaert, G. and M. Hoerova (2014) "The VIX, the Variance Premium and Stock Market Volatility," *Journal of Econometrics*, 183(2), 181-192.
- Beran, J. (1994) *Statistics for long-memory processes*, Chapman & Hall.
- Black, F. and M. Scholes (1973) "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81(3), 637-654.
- Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31(3), 307-327.
- Bollerslev, T., J. Marrone, L. Xu and H. Zhou (2014) "Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence," *Journal of Financial and Quantitative Analysis*, 49(3), 633-611.
- Bollerslev, T. and H. O. Mikkelsen (1996) "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometrics*, 73(1), 151-184.
- Bollerslev, T., G. Tauchen and H. Zhou (2009) "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463-4492.
- Bollerslev, T. and J. M. Wooldridge (1992) "Quasimaximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances," *Econometric Reviews*, 11(2), 143-172.
- Buccheri, G. and F. Corsi (2021) "HARK the SHARK: Realized Volatility Modeling with Measurement Errors and Nonlinear Dependencies," *Journal of Financial Econometrics*, 19(4), 614-649.
- Campbell, J. Y., A. W. Lo and A. C. MacKinlay (1997) *The Econometrics of Financial Markets*, Princeton: Princeton University Press. (祝迫得夫・大橋和彦・中村信弘・本田俊毅・和田賢治訳(2003)『ファイナンスのための計量分析』共立出版.)
- Chen, C. W. S. and T. Watanabe (2019) "Bayesian Modeling and Forecasting of Value-at-Risk via Threshold Realized Volatility," *Applied Stochastic Models for Business and Industry*, 35(3), 747-765.
- Chen, C. W. S., T. Watanabe and E. M. H. Lin (2021) "Bayesian Estimation of Realized GARCH-type Models with Application to Financial Tail Risk Management," *Econometrics and Statistics*, in press.
- Corsi, F. (2009) "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7(2), 174-196.
- Corsi, F., S. Mittnik, C. Pigorsch and U. Pigorsch (2008) "The Volatility of Realized Volatility," *Econometric Reviews*, 27(1-3), 46-78.
- Corsi, F., D. Pirino and R. Renò (2010) "Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting," *Journal of Econometrics*, 159 (2), 276-288.
- Diebold, F. X. (1988) *Empirical Modelling of Exchange Rate Dynamics*, Springer-Verlag.
- Ding, Z., C. W. J. Granger and R. F. Engle (1993) "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1 (1), 83-106.

- Engle, R. F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 987-1008.
- Fernández, C. and M. F. J. Steel (1998) "On Bayesian Modeling of Fat Tail and Skewness," *Journal of the American Statistical Association*, 93(441), 359-371.
- Geweke, J. (1992) "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," J. M. Bernardo, J. O. Berger, A. Dawid and A. Smith (eds.), *Bayesian Statistics 4*, 169-193, Oxford University Press, Oxford.
- Giacomini, F. and H. White (2006) "Tests of Conditional Predictive Ability," *Econometrica*, 74(6), 1545-1578.
- Glosten, L. R., R. Jagannathan and D. E. Runkle (1993) "On the Relation between the Expected Value and the Volatility of Nominal Excess Returns on Stocks," *Journal of Finance*, 48(5), 1779-1801.
- Hansen, P. R. and Z. Huang (2016) "Exponential GARCH Modeling With Realized Measures of Volatility," *Journal of Business & Economic Statistics*, 34(2), 269-287.
- Hansen, P. R., Z. Huang and H. Shek (2012) "Realized GARCH: A Joint Model of Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27(6), 877-906.
- Hansen, P. R. and A. Lunde (2005a) "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?" *Journal of Applied Econometrics*, 20(7), 873-889.
- Hansen, P. R. and A. Lunde (2005b) "A Realized Variance for the Whole Day Based on Intermittent High-frequency Data," *Journal of Financial Econometrics*, 3(4), 525-554.
- Hansen, P. R. and A. Lunde (2006a) "Consistent Ranking of Volatility Models," *Journal of Econometrics*, 131(1-2), 97-121.
- Hansen, P. R. and A. Lunde (2006b) "Realized Variance and Market Microstructure Noise," *Journal of Business & Economic Statistics*, 24(2), 127-161.
- Hansen, P., A. Lunde, and J. Nason (2011) "The Model Confidence Set," *Econometrica*, 79(2), 453-497.
- Huang, Z., H. Liu and T. Wang (2016) "Modeling Long Memory Volatility Using Realized Measures of Volatility: A realized HAR GARCH Model," *Economic Modelling*, 52(B), 812-821.
- Huang, X. and G. Tauchen (2005) "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3(4), 456-499.
- Jacod, J., Li, Y., P. A. Mykland, M. Podolskij and M. Vetter (2009) "Microstructure Noise in the Continuous Case: The Pre-averaging Approach," *Stochastic Processes and their Applications*, 119(7), 2249-2276.
- Jiang, G. J. and Y. S. Tian (2005) "Model-free Implied Volatility and Its Information Content," *Review of Financial Studies*, 18(4), 1305-1342.
- Kunitomo, N. and S. Sato (2013) "Separating Information Maximum Likelihood Estimation of the Integrated Volatility and Covariance with Micro-market Noise," *North American Journal of Economics and Finance*, 26(C), 282-309.
- Liu, L. Y., A. J. Patton and K. Sheppard (2015) "Does Anything Beat 5-minute RV? A Comparison of Realized Measures across Multiple Asset Classes," *Journal of Econometrics*, 187(1), 293-311.
- Ljung G. M. and G. E. P. Box (1978) "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 65(2), 297-303.
- Nagakura, D. and T. Watanabe (2015) "A State Space Approach to Estimating the Integrated Variance under the Existence of Market Microstructure Noise," *Journal of Financial Econometrics*, 13(1), 45-82.
- Nakajima, J. and Y. Omori (2012) "Stochastic Volatility Model with Leverage and Asymmetrically Heavy-tailed Error Using GH Skew Student's *t*-distribution," *Computational Statistics and Data Analysis*, 56(11), 3690-3704.
- Nelson, D. B. (1991) "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59(2), 347-370.
- Newey, W. K. and K. D. West (1987) "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703-708.
- Omori, Y., S. Chib, N. Shephard and J. Nakajima (2007) "Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference," *Journal of Econometrics*, 140(2), 425-449.
- Omori, Y. and T. Watanabe (2008) "Block Sampler and Posterior Mode Estimation for Asymmetric Stochastic Volatility Models," *Computational Statistics and Data Analysis*, 52(6), 2892-2910.
- Omori, Y. and T. Watanabe (2015) "Stochastic Volatility and Realized Stochastic Volatility Models," S. K. Upadhyay, U. Singh, D. K. Dey and A. Loganathan (eds.), *Current Trends in Bayesian Methodology with Applications*, Chapter 21, 435-456, Chapman & Hall/CRC Press.
- Patton, J. A. (2011) "Volatility Forecast Comparison Using Imperfect Volatility Proxies," *Journal of Econometrics*, 160(1), 246-256.
- Patton, J. A. and N. Sheppard (2009) "Evaluating Volatility and Correlation Forecast," Andersen, T. G., Davis, R. A., Kreiß, J. P., Mikosch, T. (eds.), *Handbook of Financial Time Series*, 801-838, Springer-Verlag.
- Podolskij, M. and M. Vetter (2009) "Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps," *Bernoulli*, 15(3), 634-658.
- Shephard, N. and M. Pitt (1997) "Likelihood Analysis

- of Non-Gaussian Measurement Time Series," *Biometrika*, 84, 653-667.
- Shirota, S., T. Hizu and Y. Omori (2014) "Realized Stochastic Volatility with Leverage and Long Memory," *Computational Statistics and Data Analysis*, 76(C), 618-641.
- Takahashi, M., Y. Omori and T. Watanabe (2009) "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously," *Computational Statistics and Data Analysis*, 53 (6), 2404-2426.
- Takahashi, M., T. Watanabe and Y. Omori (2016) "Volatility and Quantile Forecasts by Realized Stochastic Volatility Models with Generalized Hyperbolic Distribution," *International Journal of Forecasting*, 32(2), 437-457.
- Takahashi, M., T. Watanabe and Y. Omori (2021) "Forecasting Daily Volatility of Stock Price Index Using Daily Returns and Realized Volatility," *Econometrics and Statistics*, in press.
- Tse, Y. K. (1998) "The Conditional Heteroskedasticity of the Yen-Dollar Exchange Rate," *Journal of Applied Econometrics*, 13(1), 49-55.
- Ubukata, M. and T. Watanabe (2014a) "Pricing Nikkei 225 Options Using Realized Volatility," *Japanese Economic Review*, 65(4), 431-467.
- Ubukata, M. and T. Watanabe (2014b) "Market Variance Risk Premiums in Japan for Asset Predictability," *Empirical Economics*, 47(1), 169-198.
- Watanabe, T. (2012) "Quantile Forecasts of Financial Returns Using Realized GARCH Models," *Japanese Economic Review*, 63(1), 68-80.
- Watanabe, T. and Y. Omori (2004) "A Multi-move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard and Pitt (1997)," *Biometrika*, 91(1), 246-248.
- Xiu, D. (2010) "Quasi-maximum Likelihood Estimation of Volatility with High Frequency Data," *Journal of Econometrics*, 159(1), 235-250.
- Zhang, L. (2006) "Efficient Estimation of Stochastic Volatility using Noisy Observations: A Multi-Scale Approach," *Bernoulli*, 12(6), 1019-1043.
- Zhang, L., P. A. Mykland and Y. Aït-Sahalia, (2005) "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-frequency Data," *Journal of the American Statistical Association*, 100 (472), 1394-1411.