



Midastar モデルについて閾値自己回帰モデルと混合 頻度データ

茂木, 快治

(Citation)

国民経済雑誌, 228(3):107-117

(Issue Date)

2024-09-10

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/0100491624>

(URL)

<https://hdl.handle.net/20.500.14094/0100491624>



国民経済雑誌

THE
KOKUMIN-KEIZAI ZASSHI

(JOURNAL OF ECONOMICS & BUSINESS ADMINISTRATION)

Midastar モデルについて

—— 閾値自己回帰モデルと混合頻度データ

茂 木 快 治

国民経済雑誌 第228巻 第3号 抜刷

2024年9月

神戸大学経済経営学会

Midastar モデルについて

—— 閾値自己回帰モデルと混合頻度データ

茂 木 快 治^a

経済・金融系の時系列変数の中で、閾値効果を有するものは多い。閾値効果とは、ある閾値変数 x が何らかの閾値 μ よりも小さな値をとっているときと大きな値をとっているときとで、目的変数 y の時系列的性質が異なるという非線形効果の一種である。Tong (1978) の閾値自己回帰モデル (threshold autoregression; TAR; タール) は、目的変数 y と閾値変数 x の観測頻度が同一であることを前提として定式化されている。一方、実証分析においては、目的変数と閾値変数の観測頻度が互いに異なる状況はしばしば生じる。この課題を解決すべく、Motegi and Dennis (2024) はマイダスター (Midastar) という新たなモデルを提案した。Midastar では、目的変数 y は低い頻度、閾値変数 x^* は高い頻度で観測されると仮定される。Midastar は、閾値変数 x^* の時制集約を避け、TAR よりも高い精度で閾値効果を捕捉できるという特長をもつ。本稿では、Midastar モデルの理論と応用を解説し、今後の研究の展望を述べる。

キーワード 閾値効果, 時系列分析, Mixed Data Sampling (MIDAS),
ワイルド・ブートストラップ法

1 はじめに

経済・金融系の時系列変数の中で、閾値効果 (threshold effect) を有するものは多い。閾値効果とは、ある閾値変数 x が何らかの閾値 μ よりも小さな値をとっているときと大きな値をとっているときとで、目的変数 y の時系列的性質が異なるという非線形効果 (nonlinear effect) の一種である。Tong (1978) の閾値自己回帰モデル (threshold autoregression; TAR; タール) は、閾値効果を捕捉することのできる最も古典的かつ有名なモデルのひとつである。TAR モデルは、目的変数 y と閾値変数 x の観測頻度が同一であることを前提として定式化されている。一方、実証分析においては、目的変数と閾値変数の観測頻度が互いに異なる状況はしばしば生じる。このように複数の観測頻度が混在している状況は Mixed Data Sampling (MIDAS; マイダス) と呼ばれ、そのようなデータセットは混合頻度データ (mixed fre-

a 神戸大学大学院経済学研究科, motegi@econ.kobe-u.ac.jp

quency data) と呼ばれる。混合頻度データの計量分析は, Ghysels, Santa-Clara, and Valkanov¹⁾ (2004) を嚆矢として, 約20年間にわたって盛んに研究されている。

Motegi and Dennis (2024) [以下 MD2024] は, MIDAS の手法を TAR モデルに導入することにより, Midastar (マイダスター) という新たなモデルを提案した。Midastar では, 目的変数 y は低い頻度, 閾値変数 x^* は高い頻度で観測されると仮定される。ここで, アスタリスクは観測頻度が高いことを強調するために付してある。Midastar は, 高頻度で観測される閾値変数 x^* をそのまま余すところなく使うという点で, 閾値効果を最大限の精度で捕捉することができる。従来の TAR モデルは, 閾値変数 x^* を集約して目的変数 y と同じ観測頻度に変換するという手続きをへる分, 統計的精度は低下する。特に, もともと存在している x^* から y への閾値効果が, x^* の時制集約 (temporal aggregation) により観察できなくなってしまうという現象は, 「見せかけの非閾値効果」 (spurious non-threshold effect) と呼ばれる。一般に, 時制集約はもともと存在している非線形効果を覆い隠す方向に作用することが多く, 閾値効果についても同様のことがいえる。²⁾

MD2024 は, Midastar モデルがいくつかの標準的な仮定の下で望ましい漸近的性質を有することを示した。また, モンテカルロ実験を通じて, Midastar が現実的な設定の下で TAR よりも優れた性能を発揮することを示した。さらに, 月次の原油市場のボラティリティを目的変数として Midastar と TAR の両方を当てはめ, 標本内予測および標本外予測の精度を比較した。MD2024 の用いた閾値変数は, Midastar では日次のボラティリティ指数 (Volatility Index; VIX), TAR ではそれを月次に集計したものである。分析の結果, 標本内予測と標本外予測の両方について, Midastar は TAR よりも優れたパフォーマンスを示した。このように, 時制集約を回避する Midastar モデルの利点は大きく, 今後様々な経済・金融変数のモデル化や予測に Midastar が用いられる可能性は十分にある。

本稿では, MD2024 に則って Midastar モデルの理論と応用を解説し, 今後の研究の展望を述べる。第2節では TAR モデルを概説する。第3節では Midastar モデルの定式化と特長を説明する。第4節では Midastar に基づく統計的推測について説明する。第5節では原油市場のボラティリティに対する Midastar の応用例を紹介する。第6節では結論と今後の展望を述べる。本稿では一貫して次の記号法を用いる。実数の集合を $\mathbb{R}$, 自然数の集合を $\mathbb{N}$ と書く。実数 a を超えない整数の中で最大のものを $\lfloor a \rfloor$ と書く。事象 A が起こるとき 1, 起こらないとき 0 をとる指示関数を $1(A)$ と書く。集合 A の要素の数を $\#A$ と書く。

2 TAR モデル

目的変数 y , 閾値変数 x が整数時点 $\mathbb{L} = \{1, 2, \dots, n\}$ において観測されると仮定する。ここで, n はサンプルサイズを表す。時点 $t \in \mathbb{L}$ における目的変数と閾値変数をそれぞれ y_t ,

x_t と書く。Tong (1978) の TAR モデルは、次式のとおり定式化される。

$$y_t = \begin{cases} \alpha_1 + \sum_{k=1}^p \phi_{1k} y_{t-k} + u_t, & (x_{t-d} < \mu \text{ のとき}), \\ \alpha_2 + \sum_{k=1}^p \phi_{2k} y_{t-k} + u_t, & (x_{t-d} \geq \mu \text{ のとき}), \end{cases} \quad t \in \mathbb{L}. \quad (1)$$

ただし、 $\beta_r = (\alpha_r, \phi_{r1}, \dots, \phi_{rp})^T$ は第 r レジームにおける回帰係数ベクトル、 $p \in \mathbb{N}$ はラグの長さ、 $d \in \mathbb{N}$ は遅延パラメータ、 $\mu \in \mathbb{R}$ は閾値パラメータ、レジームの番号は $r \in \{1, 2\}$ である。誤差項 u_t に課される仮定の詳細については本稿では省略するが、代表的なものとしてはマルティンゲール差分性（すなわち $E(u_t | u_{t-1}, u_{t-2}, \dots) = 0$ ）や有限 2 次モーメント（すなわち $E(u_t^2) < \infty$ ）などがある。

TAR モデルの特徴は、閾値変数 x が閾値 μ を下回っているときと上回っているときとで、目的変数 y の自己回帰係数が互いに異なるという可能性を許容している点にある。このような閾値効果は、様々な経済・金融変数に存在すると考えられる。例えば MD2024 の実証分析にならって、目的変数 y を原油市場のボラティリティ、閾値変数 x を VIX 指数としよう。VIX 指数を金融市場全体の不安定性の尺度と見なすと、第 1 レジームは不安定性の程度がある水準 μ よりも小さな「安定期」、第 2 レジームはその逆の「不安定期」と解釈できる。金融市場の安定期と不安定期とでは、投資家の心理、企業の経営戦略、政策当局の方針など、様々な状況が異なると考えられる。したがって、原油市場のボラティリティの水準や粘性性 (persistence) もレジームごとに異なると考えられ、TAR モデルの定式化は説得力をもつ。

3 Midastar モデル

いま、閾値変数 x^* が目的変数 y よりも高い頻度で観測されるとしよう。観測頻度の比率を $m \in \mathbb{N}$ とし、ひとつの整数時点の中で x^* の観測値は m 個、 y の観測値は 1 個手に入るとする。このような混合頻度の状況を記述するため、任意の整数時点 $t \in \mathbb{L}$ を m 等分して、 $\mathbb{H}_t = \{t-1+1/m, t-1+2/m, \dots, t\}$ を定義する。さらに、 $\mathbb{H} = \bigcup_{t \in \mathbb{L}} \mathbb{H}_t$ を定義する。観測可能な変数は $\{y_t\}_{t \in \mathbb{L}}$ と $\{x_t^*\}_{t \in \mathbb{H}}$ である。

閾値変数 x^* の時制集約は、それぞれの整数時点内における m 個の観測値の算術平均をとって、次式で表される。

$$x_t = \frac{1}{m} \sum_{j=1}^m x_{t-1+\frac{j}{m}}^*, \quad t \in \mathbb{L}. \quad (2)$$

従来の TAR モデルは、高頻度の $\{x_t^*\}_{t \in \mathbb{H}}$ が観測可能であったとしても、時制集約後の $\{x_t\}_{t \in \mathbb{L}}$ を用いて (1) 式を当てはめる。これは x^* に関する情報の損失を意味するため、統計的に望ましくないアプローチである。

MD2024 は、Midastar モデルを次のとおり定式化した。

$$y_t = \begin{cases} \alpha_1 + \sum_{k=1}^p \phi_{1k} y_{t-k} + u_t, & (x_{t-\frac{d}{m}}^* < \mu \text{ のとき}), \\ \alpha_2 + \sum_{k=1}^p \phi_{2k} y_{t-k} + u_t, & (x_{t-\frac{d}{m}}^* \geq \mu \text{ のとき}), \end{cases} \quad t \in \mathbb{L}. \quad (3)$$

ただし、 $d \in \mathbb{N}$ は遅延パラメータである。(1)式と(3)式の唯一の違いは、第 t 期のレジームを決定づける条件が(1)式では x_{t-d} と μ の大小であるのに対して、(3)式では $x_{t-\frac{d}{m}}^*$ と μ の大小であるという点である。つまり、レジームを規定する遅延構造が、(1)式では低頻度ベース、(3)式では高頻度ベースで定義されているのである。

(3)式において閾値効果が存在しないとき、すなわち $\beta_1 = \beta_2$ のとき、TAR モデルと Midastar モデルはともに正しく定式化されている。両モデルとも不要な2つのレジーム構造という冗長性 (redundancy) を有するが、これは定式化の誤りを意味せず、 (β_1, β_2) の推定量にバイアスは生じない。この意味で、閾値効果が存在しないときは、TAR と Midastar との間に本質的な差はないといえる。

一方、(3)式において閾値効果が存在するとき、すなわち $\beta_1 \neq \beta_2$ のとき、Midastar は正しく定式化されており、推定にバイアスがかからない。それに対して、TAR は真の遅延構造を捉えることができず、定式化の誤りを伴う。これにより、TAR に基づく (β_1, β_2) の推定量にはバイアスが生じる。しかも、 (β_1, β_2) の擬似真値 (pseudo-true value) は互いに近い値をとることが多く、その場合は本来存在するはずの閾値効果を観察できなくなる (見せかけの非閾値効果)。言い換えると、 β_1 にかかるバイアスと β_2 にかかるバイアスは、それぞれ閾値効果を覆い隠すような符号と大きさをとりやすいのである。

Midastar のもうひとつの特長は、パラメータの数が観測頻度の比率 m に依存しない点である。通常の MIDAS 回帰モデルでは、被説明変数 y を説明変数 x^* の高頻度ラグに直接回帰するため、 m が大きくなるにつれてパラメータの数が増加し、高次元 (high dimensionality) あるいは同じことだがパラメータ過多 (parameter proliferation) の問題が発生する。それに対して Midastar では、 x^* がレジームを決めるという間接的な仕組みを有しているため、高次元の問題は生じない。これは月次の目的変数 y と日次の閾値変数 x^* の組み合わせなど、幅広い実証分析を可能にする特長である。つまり Midastar は、TAR の混合頻度データへの拡張であると同時に、MIDAS の抱える高次元の問題に対するひとつの解決案でもある。

4 Midastar モデルに基づく統計的推測

Midastar モデルに基づく統計的推測を説明するため、(3)式をベクトル形式で書き換える。まず、パラメータを次のとおりベクトル化する。

$$\underbrace{\beta}_{K \times 1} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \quad \underbrace{\gamma}_{2 \times 1} = \begin{bmatrix} d \\ \mu \end{bmatrix}, \quad \underbrace{\theta}_{(K+2) \times 1} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}.$$

ただし、回帰係数の個数は $K=2(p+1)$ である。Midastar には d と μ という 2 つの局外パラメータ (nuisance parameter) が含まれる。ラグの長さ p は既知であると仮定する。

次に、レジームを規定する二項変数を次のとおり定義する。

$$I_{1t}^*(\mu) = \mathbf{1}(x_t^* < \mu), \quad I_{2t}^*(\mu) = \mathbf{1}(x_t^* \geq \mu), \quad t \in \mathbb{H}. \quad (4)$$

(4) 式を用いて、説明変数を次のとおりベクトル化する。

$$\underbrace{\mathbf{z}_t}_{(p+1) \times 1} = \begin{bmatrix} 1 \\ y_t \\ \vdots \\ y_{t+1-p} \end{bmatrix}, \quad \underbrace{\mathbf{Z}_t(\boldsymbol{\gamma})}_{K \times 1} = \begin{bmatrix} \mathbf{z}_{t-1} I_{1, t-\frac{d}{m}}^*(\mu) \\ \mathbf{z}_{t-1} I_{2, t-\frac{d}{m}}^*(\mu) \end{bmatrix}, \quad t \in \mathbb{L}. \quad (5)$$

ここで、 $\mathbf{Z}_t(\boldsymbol{\gamma})$ という表記には注意が必要である。 $\mathbf{Z}_t(\boldsymbol{\gamma})$ が観察されるのは時点 $t - \min\{1, d/m\}$ であり、これは任意の $d \in \mathbb{N}$ の下で時点 t よりも前である。これ以降、 $\mathbf{Z}_t(\boldsymbol{\gamma})$ は y_t よりも先に観察されることに注意する。以上の準備をへて、(3) 式は次のとおりベクトル形式で表現できる。

$$y_t = \mathbf{Z}_t(\boldsymbol{\gamma})^\top \boldsymbol{\beta} + u_t, \quad t \in \mathbb{L}. \quad (6)$$

第 4.1 節ではパラメータ $\boldsymbol{\theta}$ の推定、第 4.2 節では閾値効果の有無の検定について説明する。

4.1 パラメータの推定

Midastar モデルのパラメータ $\boldsymbol{\theta}$ の推定は、TAR モデルと同様にして、プロファイリング (profiling) と呼ばれる二段階推定法によって実行される。その準備として、パラメータの選択空間 (choice space) を定義する必要がある。まず閾値パラメータ μ の選択空間を定義すべく、(2) 式に基づいて $\{x_t\}_{t \in \mathbb{L}}$ を定義し、それを昇順に並び替えたものを $x_{[1]} \leq \dots \leq x_{[n]}$ と書く。一般に、 μ の選択空間は次式で定義される。

$$\mathcal{X}_{\kappa, n} = \{x_{\lfloor 0.5(1-\kappa)n \rfloor}, \dots, x_{\lfloor (1-0.5(1-\kappa))n \rfloor}\}.$$

ただし、 $\kappa \in [0, 1)$ は $\#\mathcal{X}_{\kappa, n}$ のサンプルサイズ n に対する比率を表す。 κ の値を所与とすると、それぞれのレジームは平均的に少なくとも $\lfloor 0.5(1-\kappa)n \rfloor$ 回実現する。例えば $\kappa=0.9$ のように、 κ の値を過大に設定してしまうと、有限標本においてレジームの識別 (identification) が困難となり得る。Andrews (1993) の提案以降、 $\kappa=0.7$ と設定するのが慣例となっており、これは各レジームが平均的に少なくとも $\lfloor 0.15n \rfloor$ 回実現することを意味する。回帰パラメータ $\boldsymbol{\beta}$ の選択空間を $B \subseteq \mathbb{R}^K$ 、遅延パラメータ d の選択空間を $D \subseteq \mathbb{N}$ 、局外パラメータ $\boldsymbol{\gamma}$ の選択空間を $\Gamma = D \times \mathcal{X}_{\kappa, n}$ 、そして $\boldsymbol{\theta}$ の選択空間を $\Theta = B \times \Gamma$ と定義する。

パラメータ $\boldsymbol{\theta}$ の最小二乗推定量 $\hat{\boldsymbol{\theta}}$ は、プロファイリングによって求められる。局外パラメータ $\boldsymbol{\gamma}$ の値を固定すると、(5) 式における説明変数 $\mathbf{Z}_t(\boldsymbol{\gamma})$ は標本から計算可能である。したがって、 $\boldsymbol{\gamma}$ を所与とする $\boldsymbol{\beta}$ の条件付き最小二乗推定量は次式で与えられる。

$$\hat{\beta}(\gamma) = \left\{ \sum_{i \in \mathbb{L}} \mathbf{Z}_i(\gamma) \mathbf{Z}_i(\gamma)^\top \right\}^{-1} \left\{ \sum_{i \in \mathbb{L}} \mathbf{Z}_i(\gamma) y_i \right\}. \quad (7)$$

また、条件付き残差は $\hat{u}_i(\gamma) = y_i - \mathbf{Z}_i(\gamma)^\top \hat{\beta}(\gamma)$ で与えられる。次式のとおり、 γ の最小二乗推定量は、すべての $\gamma \in \Gamma$ の中で条件付き残差の二乗和を最小化するものとして与えられる。

$$\hat{\gamma} = \arg \min_{\gamma \in \Gamma} \sum_{i \in \mathbb{L}} \hat{u}_i(\gamma)^2. \quad (8)$$

(8)式を(7)式に代入して、 β の最小二乗推定量 $\hat{\beta} = \hat{\beta}(\hat{\gamma})$ を得る。最終的に、 θ の最小二乗推定量は、 $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ で与えられる。

MD2024 は、いくつかの標準的な前提条件の下で、最小二乗推定量 $\hat{\theta}$ の漸近的性質を導出した。第一に、 $\hat{\beta}$ は真の値 β_0 の一致推定量 (consistent estimator) である。第二に、閾値効果が存在しないとき、すなわち $\beta_1 = \beta_2$ のとき、 $\hat{\beta}$ は一般に漸近正規性 (asymptotic normality) を満たさない。第三に、閾値効果が存在するとき、すなわち $\beta_1 \neq \beta_2$ のとき、 $\hat{\beta}$ は漸近正規性を満たす。詳細については MD2024 の Theorem 1 に譲る。

4.2 閾値効果の有無の検定

次の帰無仮説 H_0 および対立仮説 H_1 を考える。

$$H_0 : \beta_1 = \beta_2 \quad H_1 : \beta_1 \neq \beta_2.$$

帰無仮説 H_0 の下で、(6)式は閾値変数 x^* の値によらず通常の $\text{AR}(p)$ モデルへと退化し (degenerate)、局外パラメータ γ は識別されない。一方、対立仮説 H_1 の下では、(6)式は退化せず、 γ は識別される。したがって、帰無仮説 H_0 は閾値効果が存在しないという仮説に対応し、対立仮説 H_1 は閾値効果が存在するという仮説に対応する。

帰無仮説 H_0 の下で γ は識別されないため、通常の χ^2 検定は漸近的な妥当性 (asymptotic validity) を失う。局外パラメータが対立仮説の下でのみ識別される場合、Hansen (1996) のワイルド・ブートストラップ法 (wild bootstrap) を用いた検定が有効であることはよく知られている。MD2024 は、Hansen (1996) のブートストラップ検定を応用して、Midastar モデルにおける閾値効果の有無を検定することを提案した。

Hansen (1996) のブートストラップ検定では、ひとつひとつの $\gamma \in \Gamma$ について帰無仮説 H_0 と対応するワルド統計量もしくはラグランジェ乗数統計量を計算した後、それらを最大値や平均値などの方法でひとつに集約したものを検定統計量として用いる。そして、データから推定された条件付き回帰スコア (conditional regression score) に標準正規乱数を掛け、多数の独立なブートストラップ検定統計量を生成して、データから計算された検定統計量との大小を比較することによって帰無仮説 H_0 を検定する。より具体的な検定の手順は、TAR モデルにおける閾値効果の有無の検定と同様であるため、ここでは省略する。

いくつかの標準的な前提条件の下で、Midastar モデルに基づく閾値効果の有無の検定は、漸近的な妥当性を有する。第一に、帰無仮説 H_0 が正しいとき、 H_0 が有意水準 $100a\%$ で棄却される確率は、漸近的に所与の $a \in (0, 1)$ と等しい。第二に、対立仮説 H_1 が正しいとき、 H_0 が有意水準 $100a\%$ で棄却される確率は、任意の $a \in (0, 1)$ について漸近的に 1 と等しい（一貫性）。詳細については MD2024 の Theorem 2 に譲る。

5 Midastar モデルによる実証分析の例

本節では、MD2024 の行った実証分析の要点を紹介する。原油は世界の商品市場（commodity market）で最も盛んに取引されている商品のひとつであり、世界の主要なエネルギー供給源のひとつである。原油価格のボラティリティのモデル化と予測は、市場参加者の金融リスク管理や中央銀行のインフレ目標政策と深く関わる研究課題である。Andersen and Bollerslev（1998）が実現ボラティリティ（realized volatility; RV）という概念を提案して以来、RV は金融資産のボラティリティの尺度としてしばしば用いられる。

Chen, Qiao, and Zhang（2022）[以下 CQZ2022] は、ウェスト・テキサス・インターミディエイト（West Texas Intermediate; WTI）の月次 RV に対して TAR モデルを当てはめた。CQZ2022 の用いた閾値変数は、スタンダード・アンド・プアーズ500種指数（S&P 500）の月次 RV である。閾値効果の有無を TAR に基づくワイルド・ブートストラップ法で検定した結果、CQZ2022 は統計的に有意な閾値効果を検出した。つまり、S&P 500 の RV の水準が閾値を下回っているときと上回っているときとでは、WTI の RV の挙動は有意に異なることが示唆された。さらに、標本外予測において、TAR モデルはベンチマークの AR モデルよりも高い精度で WTI の RV を予測できるとのことである。

CQZ2022 は、S&P 500 の RV を金融市場全体の不確実性の尺度として用いている。確かに S&P 500 の RV はひとつの妥当な選択肢と考えられるが、シカゴ・オプション取引所（Chicago Board Options Exchange; CBOE）の公表するボラティリティ指数（Volatility Index; VIX）も、金融市場全体の不確実性の尺度として広く認知されている。データの入手しやすさという観点でも、日次の RV を計算するためには 5 分刻みなどの日中価格データが必要であるが、VIX 指数の日次データは Federal Reserve Economic Data（FRED）等で公開されている。閾値変数の観測頻度が高いほど、強い閾値効果が存在する可能性が高い。このことは、閾値変数の時制集約に伴う見せかけの非閾値効果の裏返しといえる。以上の根拠に基づき、MD2024 は目的変数については CQZ2022 と同じく WTI の月次 RV を用いる一方で、閾値変数については CQZ2022 と異なり日次の VIX 指数を用いた。このような混合頻度データを伴う分析は、Midastar モデルによってのみ実現可能である。

MD2024 の標本期間は、1990年1月から2023年12月の $n=408$ か月（34年間）である。厳

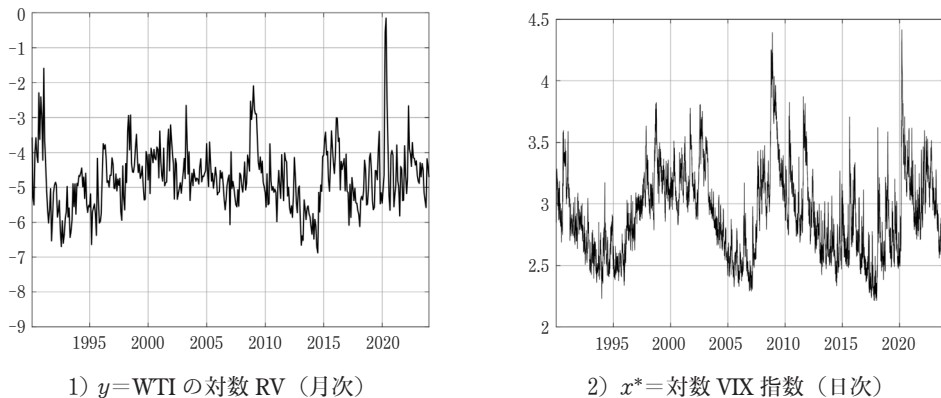
密には各月に含まれる営業日の数は月ごとに異なるため、観測頻度の比率は $\{m_t\}_{t \in \mathbb{L}}$ という時変 (time-varying) の値をとる。時変の観測頻度の比率をそのまま取り扱うのは数式表記やプログラミングを煩雑化させるため、MD2024 は $m=18$ という固定値を採用した。つまり、各月の営業日の数を18とし、実際の営業日の数がそれを超える月については、第18営業日以降の算術平均値をとる。 $m=18$ という大きな観測頻度の比率を取り扱うことができるのは、Midastar モデルのパラメータの数が m の値に依存していないからである。

図1は、目的変数 (WTI の対数 RV の月次データ) と閾値変数 (対数 VIX 指数の日次データ) をプロットしたものである。1990-1991年の湾岸戦争, 2007-2009年の世界金融危機, 2015年の中国株暴落, 2020年の新型コロナウイルス流行拡大などのたび、原油市場のボラティリティが増大しているのが読み取れる (図1.1)。その中でもコロナショックの影響は突出して大きい。一方, VIX は世界金融危機とコロナショックから同程度の影響を受けている (図1.2)。このことから、コロナショックの原油市場に対する影響は、他の金融市場に対する影響よりも甚大であったことがわかる。

MD2024 は、図1のデータに対して Midastar モデルを当てはめ、閾値効果の有無をワイルド・ブートストラップ法により検定した。その結果、閾値効果が存在しないという帰無仮説 H_0 は、有意水準1%で棄却された。日次の対数 VIX を月次に集計し、通常の TAR モデルに基づいて検定を再実行すると、帰無仮説 H_0 は5%水準では棄却できるものの、1%水準では棄却できなかった。この結果は、閾値変数 x^* の時制集約により閾値効果が弱まることを示唆しており、先行研究と整合的な結果である。

MD2024 はローリング・ウィンドウ型の標本外予測も実行し、Midastar と TAR の予測精度を比較した。その結果、日次の対数 VIX 指数を用いた Midastar の二乗平均平方根誤差

図1 WTI の対数 RV の月次データと対数 VIX 指数の日次データ



WTI の価格データは米エネルギー情報局 (EIA), VIX のデータは FRED より入手。標本期間: 1990年1月-2023年12月 ($n=408$ か月)。

(root mean squared error; RMSE) は0.800, 月次の対数 VIX 指数を用いた TAR の RMSE は0.831となった。つまり, RMSE という尺度に基づけば, Midastar は TAR よりも高い予測精度を達成する。さらに MD2024 は, これら 2 つの MSE の差が統計的に有意であるか否か, Diebold and Mariano (1995) の手法を用いて検定した。片側検定の結果, Midastar と TAR の MSE に差がないという帰無仮説は有意水準10%で棄却され, Midastar の MSE の方が TAR の MSE よりも小さいという対立仮説が支持された。この意味で, Midastar モデルは TAR モデルよりも優れた標本外予測パフォーマンスを発揮するといえる。

MD2024 の実証分析から得られた知見は, 原油市場のボラティリティの予測精度向上, ひいては投資家の効率的な金融リスク管理に資するものである。また, 原油価格の急激な変動はしばしば景気循環や物価水準に大きな影響を与えることから, MD2024 の実証分析は政策当局にとっても注目に値するものである。

6 結論と今後の展望

閾値効果は, 経済・金融系の時系列変数に含まれる非線形効果の中で代表的なものである。閾値効果を考慮した時系列モデルのひとつに, Tong (1978) の TAR モデルがある。TAR では, 目的変数 y と閾値変数 x の観測頻度が同一であるという前提条件がおかれている。一方, 実証分析においては, 目的変数と閾値変数の観測頻度が互いに異なることがよくある。このような混合頻度データに対応すべく, Motegi and Dennis (2024) は Midastar モデルを提案した。

Midastar では, 目的変数 y は低い頻度, 閾値変数 x^* は高い頻度で観測されると仮定される。Midastar は, 高頻度で観測される閾値変数 x^* をそのまま余すところなく使うという点で, 閾値効果を最大限の精度で捉えることができる。それに対して TAR モデルは, 閾値変数 x^* を集約して目的変数 y と同じ観測頻度に変換するという手続きをへる分, 統計的精度は低下する。特に, もともと存在している x^* から y への閾値効果が, x^* の時制集約により覆い隠されてしまう現象は, 「見せかけの非閾値効果」と呼ばれる。

Midastar のもうひとつの特長は, パラメータの数が観測頻度の比率 m に依存しない点である。通常の MIDAS 回帰モデルでは, 被説明変数 y を説明変数 x^* の高頻度ラグに直接回帰するため, 高次元の問題が生じやすい。一方 Midastar では, x^* がレジームを決めるという間接的な仕組みを有しているため, 高次元の問題は生じない。これは月次の目的変数 y と日次の閾値変数 x^* の組み合わせなど, 幅広い実証分析を可能にする利点である。つまり Midastar は, TAR の混合頻度データへの拡張であると同時に, MIDAS の抱える高次元の問題に対するひとつの解決案でもある。

Motegi and Dennis (2024) は, Midastar モデルに基づくパラメータの推定や仮説検定が,

いくつかの標準的な仮定の下で望ましい漸近的性質を有することを示した。実証分析の事例では、月次の原油市場のボラティリティを目的変数として Midastar と TAR の両方を当てはめ、標本内予測および標本外予測の精度を比較した。その際用いられた閾値変数は、Midastar では日次の VIX 指数、TAR ではそれを月次に集計したものである。分析の結果、標本内予測と標本外予測の両方について、Midastar は TAR よりも優れた性能を発揮した。総括すると、Midastar は経済・金融変数の変動メカニズムの解明や予測に革新をもたらすモデルであるといえる。

最後に、Midastar モデルの今後の拡張可能性について議論する。第一に、目的変数 y^* が高頻度、閾値変数 x が低頻度で観測される場合の「逆 Midastar」(reverse Midastar) を定式化することが強く望まれる。通常の MIDAS 回帰モデルは、被説明変数 y^* が説明変数 x よりも高い観測頻度を有するケースへの対応は難しい。ところが TAR モデルでは、目的変数は基本的に自己回帰過程に従い、閾値変数はレジームを決めるのみであるため、Midastar モデルと同様にして逆 Midastar モデルを定式化できる可能性はある。これが実現すれば、閾値効果の応用研究の幅はより一層広がるであろう。

第二に、ポートフォリオ戦略のパフォーマンス比較など、より実践的なシナリオの下で Midastar の既存モデルに対する優秀性を検証することが求められる。第三に、マクロ経済変数や新型コロナウイルス関連統計など、実証分析の対象を広げていくことも期待される。第四に、Midastar モデルの多変数への一般化も重要な研究課題である。MIDAS と TAR それぞれの研究発展の歴史に鑑みても、以上挙げたような Midastar の課題は、今後の研究の方向性として有望であると考えられる。

注

- 1) 混合頻度データの計量分析の概説としては、例えば Ghysels and Marcellino (2018) の Part IV が詳しい。
- 2) 時制集約により非線形効果が観察しにくくなるという傾向は、例えば Brännäs and Ohlsson (1999), Granger and Lee (1999), Paya and Peel (2006) などで報告されている。

参 考 文 献

- ANDERSEN, T. G., AND T. BOLLERSLEV (1998): "Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies," *The Journal of Finance*, 53, 219-265.
- ANDREWS, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61, 821-856.
- BRÄNNÄS, K., AND H. OHLSSON (1999): "Asymmetric time series and temporal aggregation," *The Review of Economics and Statistics*, 81, 341-344.

- CHEN, Y., G. QIAO, AND F. ZHANG (2022): “Oil price volatility forecasting: Threshold effect from stock market volatility,” *Technological Forecasting & Social Change*, 180, #121704.
- DIEBOLD, F. X., AND R. S. MARIANO (1995): “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, 13, 253–263.
- GHYSELS, E., AND M. MARCELLINO (2018): *Applied Economic Forecasting Using Time Series Methods*. Oxford University Press.
- GHYSELS, E., P. SANTA-CLARA, AND R. VALKANOV (2004): “The MIDAS Touch: Mixed Data Sampling Regression Models,” Working paper at UCLA.
- GRANGER, C. W. J., AND T. -H. LEE (1999): “The effect of aggregation on nonlinearity,” *Econometric Reviews*, 18, 259–269.
- HANSEN, B. E. (1996): “Inference when a nuisance parameter is not identified under the null hypothesis,” *Econometrica*, 64(2), 413–430.
- MOTEGI, K., AND J. W. DENNIS (2024): “Midastar: Threshold autoregression with data sampled at mixed frequencies,” SSRN Working Paper No. 4286939.
- PAYA, I., AND D. A. PEEL (2006): “Temporal aggregation of an ESTAR process: Some implications for purchasing power parity adjustment,” *Journal of Applied Econometrics*, 21, 655–668.
- TONG, H. (1978): “On a threshold model,” in *Pattern Recognition and Signal Processing*, ed. by C. H. Chen. Sijthoff and Noordhoff, Amsterdam.