



## HUB-FS Working Paper Series

FS-2020-J-001

### マクロ経済指標を含んだ INAR モデルによる 倒産リスク伝播構造の視覚化とストレステストへの応用

監物 輝夫

一橋大学大学院経営管理研究科

First version: 2020年7月4日

All the papers in this Discussion Paper Series are presented in the draft form. The papers are not intended to circulate to many and unspecified persons. For that reason any paper can not be reproduced or redistributed without the authors' written consent.

# マクロ経済指標を含んだ INAR モデルによる倒産リスク 伝播構造の視覚化とストレステストへの応用

監物 輝夫 \*

## 概要

Hawkes 過程を用いたイベント発生モデルのモデル化はファイナンス分野でも近年注目されている。本研究では、日本における比較的長期の倒産履歴データを対象に、業種に基づいて複数のイベントタイプに分類し、そのうえで多次元 Hawkes 過程を用いて倒産発生伝播構造をモデル化する。多次元 Hawkes 過程に付随する強度過程を特徴づける手法として、Embrechts and Kirchner [3] による INAR モデルを用いた分析に加えて、マクロ経済指標を INAR モデルに組み込んで業種間での倒産発生伝播構造を比較することで、マクロ経済指標を含んだ INAR モデルによる依存関係のモデル化が、倒産過程の推定でマクロ経済指標を含まないモデルよりも優れていることを示すほか、ストレステストへの応用事例も併せて例示することで、実務の観点からもモデルの有効性を示す。

## 1 はじめに

金融機関における信用リスク管理において、倒産発生に伴う損失の発生に備えるため、デフォルト確率という形で企業のリスクを定量化し、会計上の引当金の算出のほか、バーゼル規制上のリスク・アセットといった指標を開示している。ここで、例えばリーマン期を始めとした金融危機時には、倒産発生が連鎖的に発生する事象や同時に倒産する事象が知られており、こうした事象に対する定量化を行うことは信用リスク管理において非常に重要であり、構造モデルを用いて相関を勘案し、会計や規制上の開示指標以外にも内部管理上の経済資本である VaR の計測等が行われている。

しかし、相関のほか非線形な相関関係も勘案することができるコピュラを用いても、時点ごとの関連性を数値化したものにすぎず、連鎖的に倒産が発生するといった過去の事象が将来の事象に影響を与える事象については表現できず、時点間に影響を与えるようなモデルが求められる。

こうした連鎖的な倒産発生を計量化した先行研究として、Azizpour et al. [1] の研究と監物・中川 [11] の研究が挙げられる。前者の研究では、米国の社債市場のデフォルト・クラスタリングの要因を説明するため、伝播 (contagion) 項だけでなくマクロ経済指標の変数項や潜在変数 (frailty) 項を用いており、contagion 項の説明力が高いことを結論づけている。後者の研究では contagion 項だけでモデル化を試みているが、複数のカテゴリに対して分析を行っており、Kirchner による INAR モデルを用いることで、従来のパラメトリックな関数に対する尤度推定ではなく、線形代数計算に帰着した計算手法により、デフォルト発生をモデル化している。この研究では結論として、従来多く用いられてきた指数減衰型カーネル関数強度モデルによる結果と INAR モデルによる結果が、あくまでも限定されたデータセットに対してであるものの大きな違いはなく、計算負荷の観点から INAR モデルが指数減衰型カーネル関数強度モデルの代替手法として応用可能性がある

---

\* 一橋大学大学院 経営管理研究科 博士課程

している。

倒産発生モデル化については、マクロ経済指標を説明変数に加えることは非常に重要である。倒産は GDP や失業率といった指標との関連性が高く、これらのマクロ経済指標が悪化した場合、倒産が多く発生する傾向にある。こうした過去の経験を踏まえ、信用リスク管理では一定の経済シナリオを仮定したストレステストを実施しており、損失の大きさと経済指標を結びつけた感応度により将来損失額を見積もり、与信ポートフォリオを管理している。

以上を踏まえ、本研究では監物・中川 [11] の研究に対して INAR モデルにマクロ経済指標を加えた拡張モデルにより分析を行う。日本における比較的長期の倒産履歴データを対象に、倒産発生イベントを業種に基づき分類することで複数のイベントタイプを作っており、マクロ経済指標を含まない INAR モデルとの比較分析を通じて、業種間でのデフォルト発生の伝播構造の推計において、マクロ経済指標が有効に機能することを示す。

また、マクロ経済指標を含むことで、前述したストレステストの実施が可能となり、本研究においては推計したモデルを用いて幾つかのテスト用のシナリオを用意し、マクロ経済指標との感応度を確認した上で、実務の観点からもマクロ経済指標を組み込んだモデルの方が有効であることを示す。

## 2 分析データについて

本研究においては、倒産事例データとマクロ経済指標の 2 種類のデータについて使用する。それぞれのデータの内容について整理する。

### 2.1 倒産事例データについて

本研究では、監物・中川 [11] における (株) 東京商工リサーチが集計した倒産事例データを用いて分析を行っている。なお本稿では、倒産時負債総額 10 億円以上という条件でスクリーニングしていることを踏まえ、特に中小企業のデータにバイアスが生じている可能性があると考えられることから、規模別による分析は行わないこととする。

本研究では営業日ベースで倒産発生が起こっているという想定のもとで、1 年間を一律 1/4 単位とする時間幅に分割して、データ起点の 2003 年 1 月 1 日～3 月 31 日を  $t = 0$ 、終点の 2015 年 10 月 1 日～12 月 31 日が  $t = T := 13$  となるように定めた。これは、概ね四半期ごとに公表されているマクロ経済指標の粒度に合わせるもので、倒産発生日月日をそれに対応する  $k/4$  ( $k$  はある整数) と形で表される時点にマッピングした。また、複数の倒産が (同じカテゴリ内で) 発生した場合は同時点での発生件数を「マーク変数」として扱うことにした。

### 2.2 マクロ経済指標について

マクロ経済指標については、本邦において非常に多くの指標があるが、その指標の中で企業の倒産に寄与すると考えられる指標を選定している。データは倒産事例データの対象期間内における期間に合わせて四半期ごとに取得している。月次や日次で取得できる指標については、四半期末の値をその四半期における値としている。四半期のデータを生成後、完全失業率を除いて前期比をとった値を原数値として使用する。また、完全失業率については、前期差分をとった値を原数値として使用する。選定した指標と変数変換法については以下の表の通り。

表1 選定したマクロ経済指標一覧

No.	選定したマクロ経済指標
1	GDP : 日本実質 GDP(季節調整済)
2	UER : 完全失業率 (季節調整済)
3	IPM : 鉱工業生産指数 (季節調整済)
4	CIL : 景気動向指数
5	TPX : TOPIX

表2 変数変換法一覧

No.	変数変換法
1	原数値
2	ラグ四半期
3	ラグ半年
4	ラグ1年
5	移動平均 (半年)
6	移動平均 (1年)

### 3 拡張 Hawkes 過程による倒産発生モデル

当節では、まず一般的な多次元 Hawkes 過程の定義を確認し、倒産発生依存関係をモデル化する際の Hawkes 過程のカーネル関数の定式化についてまとめておく。次に、多次元 Hawkes 過程にマクロ経済指標による説明変数群の項を加えた拡張 Hawkes 過程を定式化する。その上で次節にて、Embrechts and Kirchner [3] が提唱した推定手法から本研究で提示するマクロ経済指標を含んだ INAR モデルへ拡張を行い、推定法を説明する。マクロ経済指標を含まない INAR モデルについては、Kirchner [7] や監物・中川 [11] を参照されたい。

#### 3.1 拡張 Hawkes 過程

適当なフィルター付き完備確率空間  $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  を与える。  $m \in \mathbb{N}$  種類のイベントを対象にする。ここで1から自然数  $n$  までの集合を  $[n] = \{1, 2, \dots, n\}$  のように表すことにする。

各イベント  $j \in [m]$  に対して、 $\mathbb{F}$ -適合な点過程 (停止時刻の増大列)  $\{\tau_k^j\}_{k \in \mathbb{N}}$  を考える。すなわち、 $\tau_k^j$  をタイプ  $j$  のイベントが  $k$  回目に発生する  $\mathbb{F}$ -停止時刻と見なし、 $\mathbf{P}$ -a.s. に  $0 (= \tau_0^j) < \tau_1^j < \tau_2^j < \dots$  が成り立つとする。

さらに、各イベント  $j \in [m]$  に対して、点過程  $\{\tau_k^j\}_{k \in \mathbb{N}}$  に付随する計数過程を  $N_t^j := \sum_{k \geq 1} \mathbf{1}_{\{\tau_k^j \leq t\}}$  と定義し、まとめて  $\mathbf{N}_t = (N_t^1, \dots, N_t^m)^\top$  と表す。

このとき、一般に、多次元 Hawkes 過程は次のような強度過程で特徴づけられる計数過程として定義される (Bacry et al. [2] の Definition 1 などを参照)。

定義 (多次元 Hawkes 過程).  $m$ -次元 計数過程のベクトル  $\mathbf{N}_t$  が  $m$ -次元 Hawkes 過程であると

は、各  $j \in [m]$  に対して  $N_t^j$  に付随する強度過程\*1が

$$\lambda_t^j = \mu^j + \sum_{i=1}^m \int_{-\infty}^t h_{i \rightarrow j}(t-s) dN_s^i \quad (1)$$

で与えられるものをいう。ただし、 $\mu^j \geq 0$  はタイプ  $j$  の「外生強度 (*exogenous intensity*)」、 $h_{i \rightarrow j}(u)$  は「カーネル関数 (*kernel function*)」と呼ばれる、任意の  $(i, j) \in [m]^2$  に対して非負値の確定的な可積分関数で、 $u < 0$  のときは  $h_{i \rightarrow j}(u) \equiv 0$  を満たすものとする\*2。

次に、 $k \in [\ell]$  種類のマクロ経済指標値を、時点  $t$  において  $\mathbb{F}$ -適合であり右連続かつ左極限を持つ確率過程  $M_t^k$  とすると、拡張 Hawkes 過程を次のような強度過程で特徴づけられる計数過程として定義する。

定義 (拡張 Hawkes 過程).  $m$ -次元 計数過程のベクトル  $\mathbf{N}_t$  が  $m$ -次元 拡張 Hawkes 過程であるとは、各  $j \in [m]$  に対して  $N_t^j$  に付随する強度過程\*3が

$$\lambda_t^j = \mu^j + \sum_{i=1}^m \int_{-\infty}^t h_{i \rightarrow j}(t-s) dN_s^i + \sum_{k=1}^{\ell} \beta_k^j M_t^k \quad (2)$$

で与えられるものをいう。ただし、 $\beta_k^j$  は、マクロ経済指標値  $M_t^k$  に付随する係数である。

(1) 式における積分項の被積分関数  $h_{i \rightarrow j}(t-s)$  は具体的には「時点  $s$  で発生したタイプ  $i$  のイベントが、タイプ  $j$  のイベント発生強度に与えたインパクトが時点  $t$  では (すなわち  $t-s$  だけ経過後に) どの程度残っているか」を表していると解釈できる。また、(2) 式のように、マクロ経済指標による説明変数項を加えることで、カーネル関数  $h_{i \rightarrow j}(t-s)$  だけでは捉えきれない影響を勘案できるようになる。したがって、Hawkes 過程は強度過程のカーネル関数の形のほか、どのようにマクロ経済指標を与えるかによって特徴づけることが可能である。

5 節においては、倒産発生の依存関係をモデル化する際に Hawkes 過程のカーネル関数  $h_{i \rightarrow j}(t-s)$  をの形は特定せず、Kirchner [7] が提唱した手法で推定する。より具体的には、「単位期間ごとの倒産件数」をデータとし、翌単位期間における倒産件数を整数値の AR モデル (「INAR モデル」と呼ばれる) で近似的に表現できると仮定して、「(単位期間ごとの) 離散時点上のカーネル関数の推定値」を線形代数の演算によって計算する。

また、Kirchner [7] が提唱した INAR モデルに対してマクロ経済指標を組み込み、Contagion 項だけでなくマクロ経済環境の変化に伴う影響も加味した方法により、2 節で述べた日本の関東地区の倒産履歴データに対する分析結果 (推定された依存関係および事後的なイベント発生強度の推移) を比較・考察する。

\*1 非負値 (発展的) 可測過程  $\lambda_t^j$  が計数過程  $N_t^j$  の強度過程であるとは、 $N_t^j - \int_0^t \lambda_s^j ds$  が  $\mathbb{F}$ - (局所) マルチンゲールになることと定義される。

\*2 Kirchner [7] では、 $\mu^j, h_{i \rightarrow j}(u)$  をそれぞれ *baseline intensity*, *excitement function* と呼んでいる。また、Embrechts and Kirchner [3] では  $\mu^j$  を *immigration intensity* と呼んでいる。本稿では Bacry et al. [2] の呼び方に即した訳語を用いている。

\*3 非負値 (発展的) 可測過程  $\lambda_t^j$  が計数過程  $N_t^j$  の強度過程であるとは、 $N_t^j - \int_0^t \lambda_s^j ds$  が  $\mathbb{F}$ - (局所) マルチンゲールになることと定義される。

## 4 多次元 Hawkes 過程のカーネル関数の特徴づけと推定法

当節では、Embrechts and Kirchner [3] が提唱した Kirchner [7] による INAR モデルに対して、マクロ経済指標を追加した場合におけるカーネル関数の推定手法の導出について説明する。Hawkes グラフ表現については、Embrechts and Kirchner [3] による方法と同様の手順で得られることから説明を省略する。

### 4.1 Embrechts-Kirchner の推定法

Embrechts and Kirchner [3] で提唱されている INAR モデルによるカーネル関数の推定をベースに、マクロ経済指標を加えた 拡張 Hawkes 推定量の導出についてまとめる。Embrechts and Kirchner [3] では Hawkes 推定量の後に Hawkes グラフ表現を推定しており、以後、この推定法を「Embrechts-Kirchner の推定法」と呼ぶことにする。

Embrechts-Kirchner の推定法においては、Hawkes 推定量と名付けられる「(分析上指定した) 単位期間に対応する離散時点上のカーネル関数の推定値」を成分にもつ行列、およびその標準誤差を与える行列を求めることが重要となる。そのためには、イベント発生時点データを指定した単位時間ごとの「イベント発生件数」のデータに変換する必要がある。<sup>\*4</sup>

#### 拡張 Hawkes 推定量

ここでは、Kirchner [7] および Embrechts and Kirchner [3] をもとに、「(分析上指定した) 単位期間に対応する離散時点上のカーネル関数の推定値」及びマクロ経済指標を成分にもつ行列である拡張 Hawkes 推定量の計算法を述べる。

拡張 Hawkes 推定量を用いる動機として、式 (1) のカーネル関数  $h_{i \rightarrow j}(u)$  に対して特定の関数形を与えず、 $p$ -次のラグをもつ整数値 (ベクトル) 自己回帰モデル (INAR( $p$ ) モデルと表される) で近似できると仮定することで、線形代数計算のみで推定できるという利点があるという点が挙げられる。指数減衰型カーネル関数を用いるようなパラメトリックモデルに対する最尤法では数値的最適化が必要となることが多いが、特に今回のモデルのようにマクロ経済指標を含むような高次元の場合には最適化計算の負荷が大きくなる。したがって、特定のカーネル関数を仮定すべき必然的な理由が無ければ、推定時の計算負荷を回避することができる点は手法的に大きな優位性と言える。

しかし一方であくまでも離散時点上で近似的にカーネル関数を推定する手法であるので、近似が精度良く機能しているかどうかには注意を払う必要がある。

まず、拡張 Hawkes 推定量を求めるためには、イベント発生時点のデータを変換する必要がある。

そのために (十分に小さい)  $\Delta > 0$  を分析上定める単位期間として最初に固定する。  $k \in [n]$  (ただし  $n := \lfloor T/\Delta \rfloor$ ) に対して、時間区間  $((k-1)\Delta, k\Delta]$  内におけるタイプ  $j$  のイベント発生件数を  $X_k^{(j,\Delta)}$  と表すことにする<sup>\*5</sup>。すなわち  $X_k^{(j,\Delta)} = N_{k\Delta}^j - N_{(k-1)\Delta}^j$  である。

<sup>\*4</sup> 実際の推定においては、Hawkes グラフ表現を得る前に、やや弱い条件の下で「Hawkes スケルトン」という外生強度や辺の重みの情報をもたない有向グラフ表現を得て、それから Hawkes スケルトンの部分グラフとして Hawkes グラフを得るという 2 段階のステップで構成される

<sup>\*5</sup> このように、ある区間内でのイベント発生件数をデータとして用いることになるので、倒産データのように複数のイベント発生時点が同時点に記録されている (実際は正確な時点を特定できない) 場合でも、同時点データに対する

この  $X_k^{(j,\Delta)}$  を成分にもつ  $m$ -次元ベクトルを  $\mathbf{X}_k^{(\Delta)} = (X_k^{(1,\Delta)}, X_k^{(2,\Delta)}, \dots, X_k^{(m,\Delta)})^\top$  と定義する。このとき、 $\mathbf{X}_k^{(\Delta)}$  は、ある  $p \in \mathbb{N}$  に対して直前の過去  $p$  期分の観測  $\mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}$  を所与とし、時点  $k\Delta$  におけるタイプ  $n$  のマクロ経済指標を  $M_k^{(n,\Delta)}$  と表し、 $M_k^{(n,\Delta)}$  を成分にもつ  $q$ -次元ベクトルを  $\mathbf{M}_k^{(\Delta)} = (M_k^{(1,\Delta)}, M_k^{(2,\Delta)}, \dots, M_k^{(q,\Delta)})^\top$  と定義する。このとき、マクロ経済指標を含んだ場合の近似式を以下で表すことができる。

$$\mathbf{E} \left[ X_k^{(j,\Delta)} \mid \mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}, \mathbf{M}_k^{(\Delta)} \right] \approx \Delta \cdot \mu^j + \Delta \sum_{\ell=1}^p \sum_{i=1}^m h_{i \rightarrow j}(\ell\Delta) X_{k-\ell}^{(i,\Delta)} + \Delta \sum_{a=1}^q \beta_a^j M_k^{(a,\Delta)} \quad (3)$$

これらの近似式によって、 $\mathbf{X}_k^{(\Delta)}$  を次のような整数値（ベクトル）自己回帰モデルに対応させることが考えられる。そこで、ある固定された  $p \in \mathbb{N}$  に対して、 $m$ -次元ベクトル  $\alpha_0 := (\alpha_0^1, \dots, \alpha_0^m)^\top$ 、行列  $(\alpha_\ell^{j,i})_{1 \leq j, i \leq m, \ell \in [p]} \in \mathbb{R}_{\geq 0}^{m \times m}$  および  $q$ -次元ベクトル  $\beta := (\beta_1, \dots, \beta_q)^\top$  を推定すべきパラメータとする。 $\mathbf{X}_k^{(\Delta)}$  は、 $p$  次のラグをもつ次のような整数値自己回帰モデル（以後、INAR( $p$ )モデル）を満たすとす\*6。

$$k \in \{p+1, \dots, n\} \quad \mathbf{X}_k^{(\Delta)} = \sum_{\ell=1}^p \begin{pmatrix} \sum_{i=1}^m X_{k-\ell}^{(i,\Delta)} \sum_{u=1}^m \xi_u^{(\alpha_\ell^{1,i})} \\ \vdots \\ \sum_{i=1}^m X_{k-\ell}^{(i,\Delta)} \sum_{u=1}^m \xi_u^{(\alpha_\ell^{m,i})} \end{pmatrix} + \sum_{a=1}^q \varepsilon_k^{(a)} + \varepsilon_k. \quad (4)$$

ただし、 $\{\xi_u^{(\alpha_\ell^{j,i})}\}_{u=1,2,\dots}$  は、平均  $\alpha_\ell^{j,i}$  のポアソン分布  $Po(\alpha_\ell^{j,i})$  に従う i.i.d. の確率変数であり、 $\mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}$  とは独立である。行列  $\varepsilon_k^{(a)} = (\varepsilon_k^{(a,1)}, \dots, \varepsilon_k^{(a,m)})^\top \in \mathbb{R}^{m \times q}$  は他のすべての確率変数と独立であり、 $\varepsilon_k^{(a,j)} \sim Po(\Delta \beta_a^j M_k^{(a,\Delta)})$  とする。また、 $\varepsilon_k = (\varepsilon_k^1, \dots, \varepsilon_k^m)^\top$  は  $m$ -次元確率変数の列ベクトルであり、 $j \in [m]$  に対して  $\varepsilon_k^j$  は他のすべての確率変数と独立であり、 $\varepsilon_k^j \sim Po(\alpha_0^j)$  であるとする。

特別な工夫をする必要がなくなるメリットはある。

\*6 本稿の INAR( $p$ ) モデルの定式化は Kirchner [7] を参考にしている。ただし、原論文では reproduction operator (thinning operator などとも呼ばれる) というオペレーターを用いた簡潔な表現を採用しているが、ここではオペレーターを用いない表記で記述している。なお、INAR モデルについては、Fokianos [5] などで詳しく解説されている。また、基礎となる i.i.d. の確率変数列の分布も、今回は Hawkes モデルとの対応からポアソン分布を選択しているが、他の離散分布を仮定することも可能である。

(4) 式の第  $j$  成分について条件付き期待値をとると、

$$\begin{aligned}
& \mathbf{E} \left[ X_k^{(j,\Delta)} \mid \mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}, \mathbf{M}_k^{(\Delta)} \right] \\
&= \mathbf{E} \left[ \sum_{\ell=1}^p \sum_{i=1}^m \sum_{u=1}^{X_{k-\ell}^{(i,\Delta)}} \xi_u^{(\alpha_{\ell}^{j,i})} + \sum_{a=1}^q \varepsilon_k^{(a,j)} + \varepsilon_k^j \mid \mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}, \mathbf{M}_k^{(\Delta)} \right] \\
&= \sum_{\ell=1}^p \sum_{i=1}^m \mathbf{E} \left[ \sum_{u=1}^{X_{k-\ell}^{(i,\Delta)}} \xi_u^{(\alpha_{\ell}^{j,i})} \mid \mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)} \right] + \sum_{a=1}^q \mathbf{E}[\varepsilon_k^{(a,j)} \mid \mathbf{M}_k^{(\Delta)}] + \mathbf{E}[\varepsilon_k^j] \\
&= \sum_{\ell=1}^p \sum_{i=1}^m \mathbf{E}[\xi_1^{(\alpha_{\ell}^{j,i})}] \mathbf{E}[X_{k-\ell}^{(i,\Delta)} \mid \mathbf{X}_{k-1}^{(\Delta)}, \dots, \mathbf{X}_{k-p}^{(\Delta)}] + \sum_{a=1}^q \Delta \beta_a^j M_k^{(a,\Delta)} + \alpha_0^j \\
&= \alpha_0^j + \sum_{\ell=1}^p \sum_{i=1}^m \alpha_{\ell}^{j,i} X_{k-\ell}^{(i,\Delta)} + \sum_{a=1}^q \Delta \beta_a^j M_k^{(a,\Delta)}. \tag{5}
\end{aligned}$$

ただし、3つ目の等号では Wald の等式を用いている。

この結果と (3) の近似式を比較することで、INAR( $p$ ) モデルのパラメータと、拡張 Hawkes 推定量の成分として得られる外生強度  $\mu^j$  およびカーネル関数  $h_{i \rightarrow j}(t)$  の間には、 $\alpha_0^j \approx \mu^j \Delta$  ( $j \in [m]$ ) および  $\alpha_{\ell}^{j,i} \approx h_{i \rightarrow j}(\ell \Delta) \Delta$  ( $(i, j) \in [m]^2, \ell \in [n]$ ) という近似関係が成り立つことに注意する。

INAR( $p$ ) モデルのパラメータ  $(\alpha_0^j)_{j \in [m]}$ 、 $(\alpha_{\ell}^{j,i})_{(i,j) \in [m]^2, \ell \in [p]}$  および  $(\beta_a^j)_{j \in [m], a \in [q]}$  は「条件付き最小二乗法 (CLS)」によって推定することができる\*7。要するに、離散時点グリッド上のカーネル関数の値に  $\Delta$  をかけた  $(\Delta h_{i \rightarrow j}(\Delta), \dots, \Delta h_{i \rightarrow j}(p\Delta))$  は、線形の重回帰式の回帰係数に対応したものであるので、グリッド上のカーネル関数の値を行列形式にまとめたものは、線形重回帰モデルにおける OLS 推定量が行列演算式で与えられることのアナロジーで次のような行列演算で与えられることが説明できる。

次の行列演算式で与えられる  $p$ -期間分の強度過程に関する推定量の行列  $\hat{\mathbf{H}}_M^{(\Delta,p)} := (\hat{\mathbf{H}}_1^{\top} \dots \hat{\mathbf{H}}_p^{\top} \hat{\boldsymbol{\beta}} \hat{\boldsymbol{\mu}})^{\top} \in \mathbb{R}^{(mp+q+1) \times m}$  から  $\hat{\boldsymbol{\beta}}$  を除いた  $\hat{\mathbf{H}}^{(\Delta,p)} := (\hat{\mathbf{H}}_1^{\top} \dots \hat{\mathbf{H}}_p^{\top} \hat{\boldsymbol{\mu}})^{\top} \in \mathbb{R}^{(mp+1) \times m}$  は Embrechts-Kirchner の推定法における「Hawkes 推定量」となる。ただし  $k \in [p]$  に対して  $\hat{\mathbf{H}}_k := (\hat{h}_{i \rightarrow j}(k\Delta))_{(i,j) \in [m]^2} \in \mathbb{R}^{m \times m}$  とする。

$$\hat{\mathbf{H}}_M^{(\Delta,p)} = \frac{1}{\Delta} (\mathbf{Z}_M^{\top} \mathbf{Z}_M)^{-1} \mathbf{Z}_M^{\top} \mathbf{Y}. \tag{6}$$

ただし、

$$\mathbf{Z}_M := \begin{pmatrix} (\mathbf{X}_p^{(\Delta)})^{\top} & (\mathbf{X}_{p-1}^{(\Delta)})^{\top} & \dots & (\mathbf{X}_1^{(\Delta)})^{\top} & (\mathbf{M}_{p+1})^{\top} & 1 \\ (\mathbf{X}_{p+1}^{(\Delta)})^{\top} & (\mathbf{X}_p^{(\Delta)})^{\top} & \dots & (\mathbf{X}_2^{(\Delta)})^{\top} & (\mathbf{M}_{p+2})^{\top} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ (\mathbf{X}_{n-1}^{(\Delta)})^{\top} & (\mathbf{X}_{n-2}^{(\Delta)})^{\top} & \dots & (\mathbf{X}_{n-p}^{(\Delta)})^{\top} & (\mathbf{M}_n)^{\top} & 1 \end{pmatrix} \in \mathbb{N}^{(n-p) \times (mp+q+1)}. \tag{7}$$

この拡張 Hawkes 推定量は、強度過程の事後的なパスを描く際にも用いられる。

\*7 これらのパラメータ推定には、最尤法や Yule-Walker (YW) 法によるモーメント法的な推定法も考えられる。Kirchner (2016) では、パラメータ数が多くなるとナイーブな最尤法は難しく、YW 型は単位根に近い場合には推定バイアスが大きくなる、といったことがコメントされている。

## 4.2 Embrechts-Kirchner の推定法におけるチューニング・パラメータの決め方

前節で述べた Embrechts-Kirchner の推定法に基づく拡張 Hawkes 推定量にはマクロ経済指標が含まれており、倒産発生の余波の残る期間をコントロールする  $p$  とともにマクロ経済指標を決める必要がある。ここで、本来であれば  $p$  とマクロ経済指標は同時に選択すべきであるが、計算負荷の観点から本研究では  $p$  とマクロ経済指標は独立して決定する。

さて、マクロ経済指標の選択に当たっては様々な手法があるが、本研究においてはポアソン回帰により AIC を統計基準として、最も AIC が低くなる説明変数の組み合わせを選択することを考える。具体的には、総当り法により業種毎に説明変数の組み合わせに対して AIC を取得する。Hawkes グラフ表現においては全ての業種で共通のマクロ経済指標を用いる必要があることから、業種に対して平均を取り、最も低い説明変数の組み合わせを Hawkes グラフ表現で用いるマクロ経済指標とする。

なお、変数選択で採用する際に、各指標最大 1 つまでとする。すなわち、同じ指標で変数変換法が異なる変数は同時には採用しないこととする。

次に、倒産発生の余波の残る期間をコントロールする  $p$  について、Kirchner [7] では、適当な事前分析用の単位期間  $\Delta_0 > 0$  および最大のラグ  $p_0$  をそれぞれ固定したうえで AIC の最小化を基準に決定する方法を提案している。マクロ経済指標を加えると以下の通り与えられる。

$$\text{AIC}^{\Delta_0}(p) := \log \left( \det \hat{\Sigma}^{(\Delta_0)}(p) \right) + \frac{2m(pm + q)}{n_0 - p}, \quad n_0 = \lfloor T/\Delta_0 \rfloor, \quad p \in \{1, \dots, p_0\},$$

ただし、 $\mathbf{U}_k$  を Hawkes スケルトン表現の推定過程で得られる  $\text{INAR}(p)$  のホワイトノイズ項ベクトルとして、 $\hat{\Sigma}^{(\Delta_0)}(p)$  は次で定義されるものとする：

$$\hat{\Sigma}^{(\Delta_0)}(p) := \frac{1}{n_0 - p} \sum_{k=p+1}^{n_0} \mathbf{U}_k \mathbf{U}_k^T.$$

今回の倒産サンプルに対しては、マクロ経済指標の最小粒度に合わせて、単位期間を四半期として単位期間  $\Delta_0$  を採用することにする。また、 $p_0$  については、倒産イベントの伝播が最大 2 年間残留するとして単位期間を採用する。

その結果、 $\Delta_0 = 1/4$  および  $p_0 = 8$  を固定して、上記の式に基づいて AIC 最小化を行うことになり、 $p = 6$  が得られた。なお、本研究では  $\alpha$  を 0.05 に設定している（片側検定の意味で 5% 有意水準を設定していることになる）。

## 5 分析結果

### 5.1 Hawkes グラフ表現と倒産発生の伝播構造について

本節では、2 節に述べた 2003 年初～2015 年末の関東地区の倒産履歴データを業種カテゴリに分類したものをもとに、Hawkes 過程に基づく倒産強度モデルを 2 種類の特徴づけ（Embrechts-Kirchner の推定法）によって分析を行い、Hawkes グラフ表現として視覚化したうえで、倒産リスクの依存構造について比較・考察を行う。

本研究で用いたマクロ経済指標については以下の図 1 のとおりである。マクロ経済指標を用いない INAR モデルを除き、ポアソン回帰モデルとマクロ経済指標を組み込んだ INAR モデルの 2

つのモデルはこの指標を用いて推計される。その結果が図 2 である。

この結果を見ると、いずれのモデルも倒産件数がピークを迎えた時期の動きについては捉えることが出来ているが、ポアソン回帰の場合では、その他の時点では大きく推移から乖離しており倒産過程のトレンドを捉えきることが出来ていないことが確認できる一方で、2つの INAR モデルでは倒産過程のトレンドを上手く捉えている。また、INAR モデル同士でのモデルを比較するため AIC 及び平方平均二乗誤差の 2 つの評価指標で見ると、マクロ経済指標を含んだ INAR モデルの方が優れていることが分かる。

次に、図 3 及び図 4 による寄与度分解の結果を見ると、全体的に製造関係とサービス関係が各業種の内訳に占める割合が大きく、マクロ経済指標の有無に関わらず構造が大きく変わらないことが分かる。より細かく見ていくと、製造関係におけるインフラ・流通関係の符号反転やサービス関係におけるインフラ・流通関係の負の影響の強まりといった点が、両者の構造の差異として確認できる。また、マクロ経済指標を含んだ INAR モデルの結果を見ると、サービス関係を除く倒産件数がピークを迎えるリーマン期付近ではマクロ経済指標の効果が強まり、倒産過程に寄与している様子が確認できる。以上のことから、倒産過程における伝播構造については両モデルとも大きく変わらないことが寄与度分解の結果から確認でき、細かい部分ではマクロ経済指標の影響とトレードオフ関係になっていることが考えられる。

図 5 では、Embrechts-Kirchner の推定法により倒産発生伝播構造を Hawkes グラフ表現として表したものである。この図から、製造関係とサービス関係が大きく占めている構造であり、サービス関係が特に強く全体の伝播構造に寄与していることが確認できる。Hawkes グラフ表現による結果からも、マクロ経済の有無に関わらず INAR モデルが倒産発生伝播構造を上手く捉えていることが確認できる。

以上の結果から、ポアソン回帰では捉えられないトレンドを INAR モデルでは捉えることができ、INAR モデルにより倒産発生伝播構造は概ね捉えることができる。ただし、マクロ経済指標をモデルに組み込むことで、INAR モデルだけでは捉えきれない影響を捉えられることを AIC や平方平均二乗誤差による定量的な結果のほか、実績対比図や寄与度分解図からも評価することができ、マクロ経済指標を含んだ INAR モデルの有効性を示すことができた。

一方、監物・中川 [11] の研究対比では、製造関係が多くの業種に影響を与えている点は同じであるものの、サービス関係の影響についてが本研究のような影響力を持たない点が異なっている。監物・中川 [11] では週次データを用いて分析を行っており、本研究では捉えきれない粒度で倒産の動きを把握できるが、マクロの影響は捉えられないなど、双方にメリット・デメリットがあり、直接は比較できない。しかし、本研究のように予測を主眼とした分析では、マクロ経済指標との関連性が重要となることから、本研究のようにマクロ経済指標を組み込んだモデルの方が扱いやすいと考えられる。

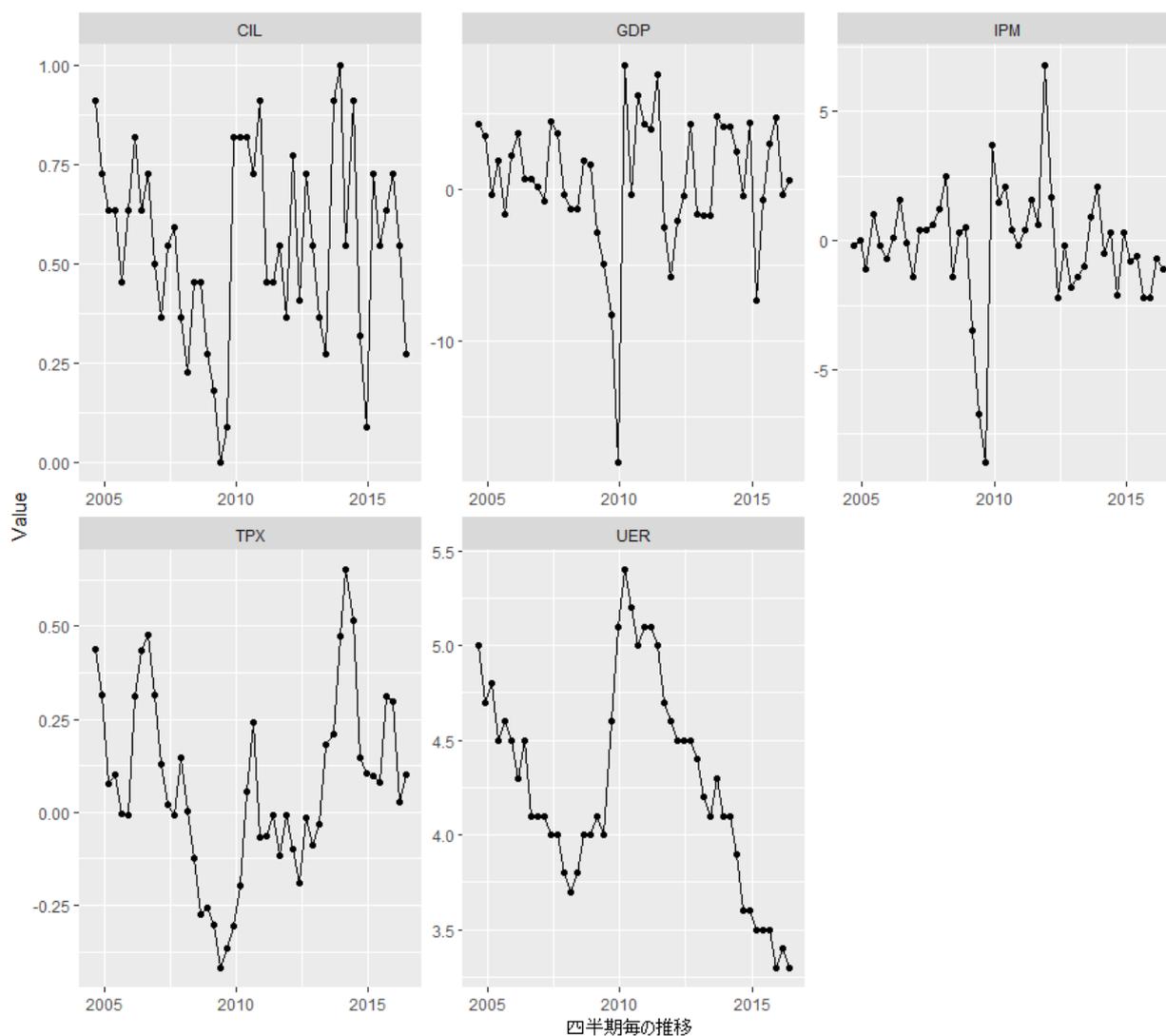


図1 四半期ごとのマクロ経済指標の推移。CILは「景気動向指数」、IPMは「鉱工業生産指数」、UERは「完全失業率」を表す。GDPとTOPIXについては変化率で表示している。なお、TOPIXについては、四半期ごとの期末値を用いており、変化率はこの四半期に対する変化率である。

表3 AICと平方平均二乗誤差によるモデルの評価。平方二乗誤差については業種別に算出している。

	INAR(マクロあり)	INAR(マクロなし)
AIC	35.5	37.3
製造(1)	5.30	7.09
イン・流通(2)	4.54	4.84
金融(3)	2.91	3.57
サービス(4)	3.53	4.32

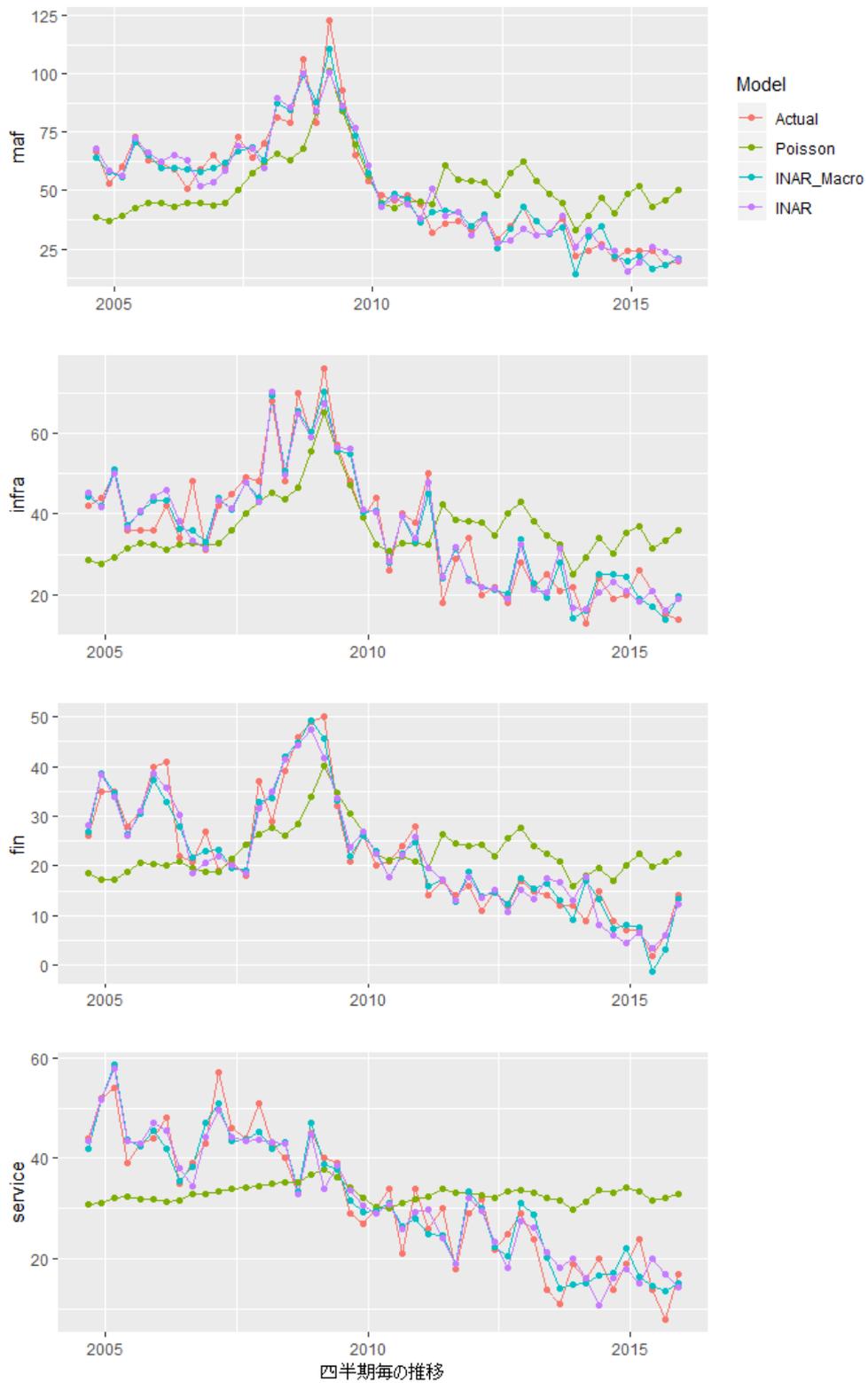


図2 3つのモデルによる比較。マクロ経済指標については、総当り法により景気動向指数（移動平均（1年））とTOPIX（ラグ1年）が選択される。ポアソン回帰については別々に選択することもできるが、INARモデルに合わせて共通のマクロ経済指標を用いている。

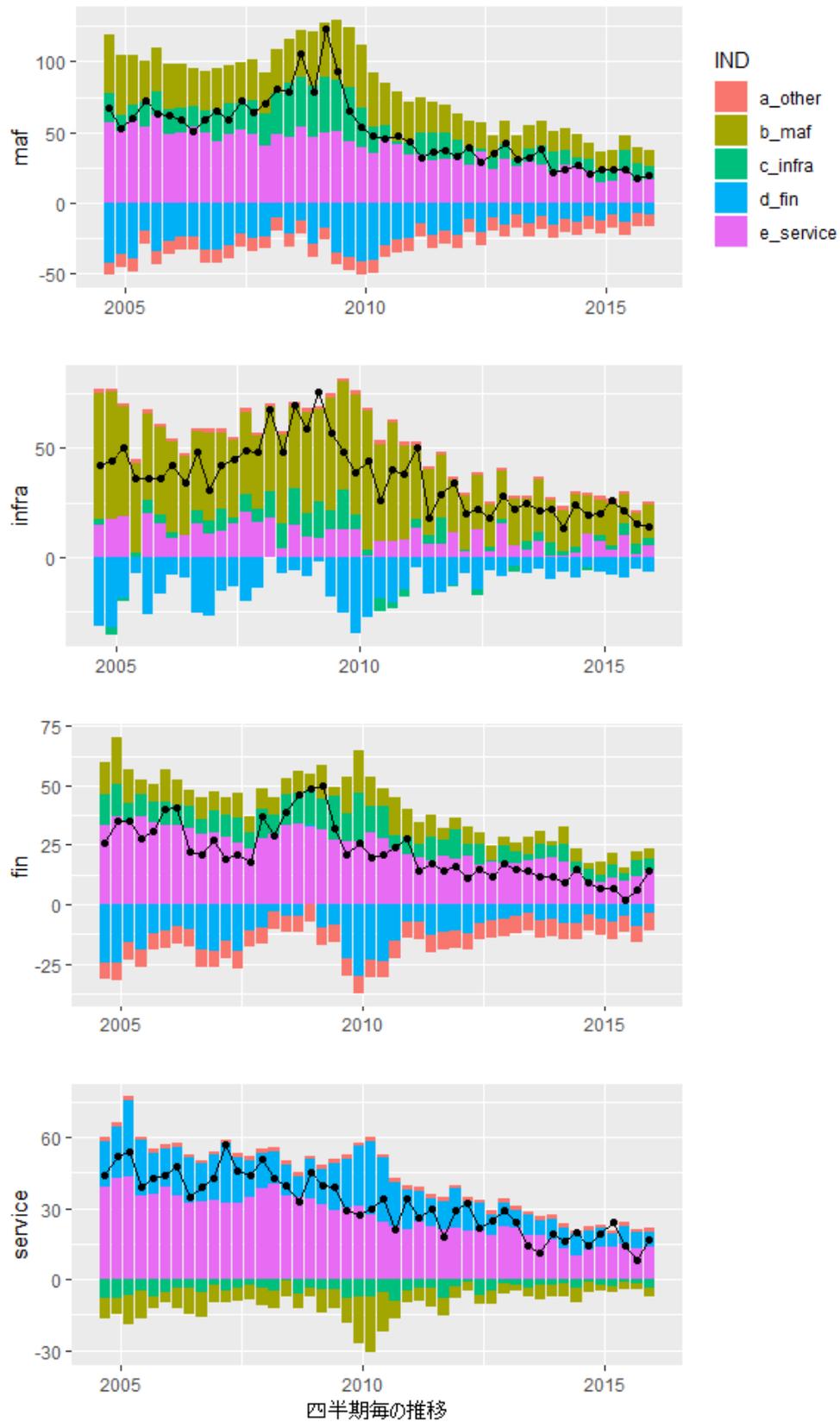


図3 INAR モデルによる寄与度分解。Hawkes 推定量算出時の係数を用いて業種別・ラグ別に推計し、ラグを全て足し合わせたものを業種別の寄与度としている。なお、縦軸は倒産発生件数を表し、黒線は倒産発生件数の実績値である。また、other は切片を表している。

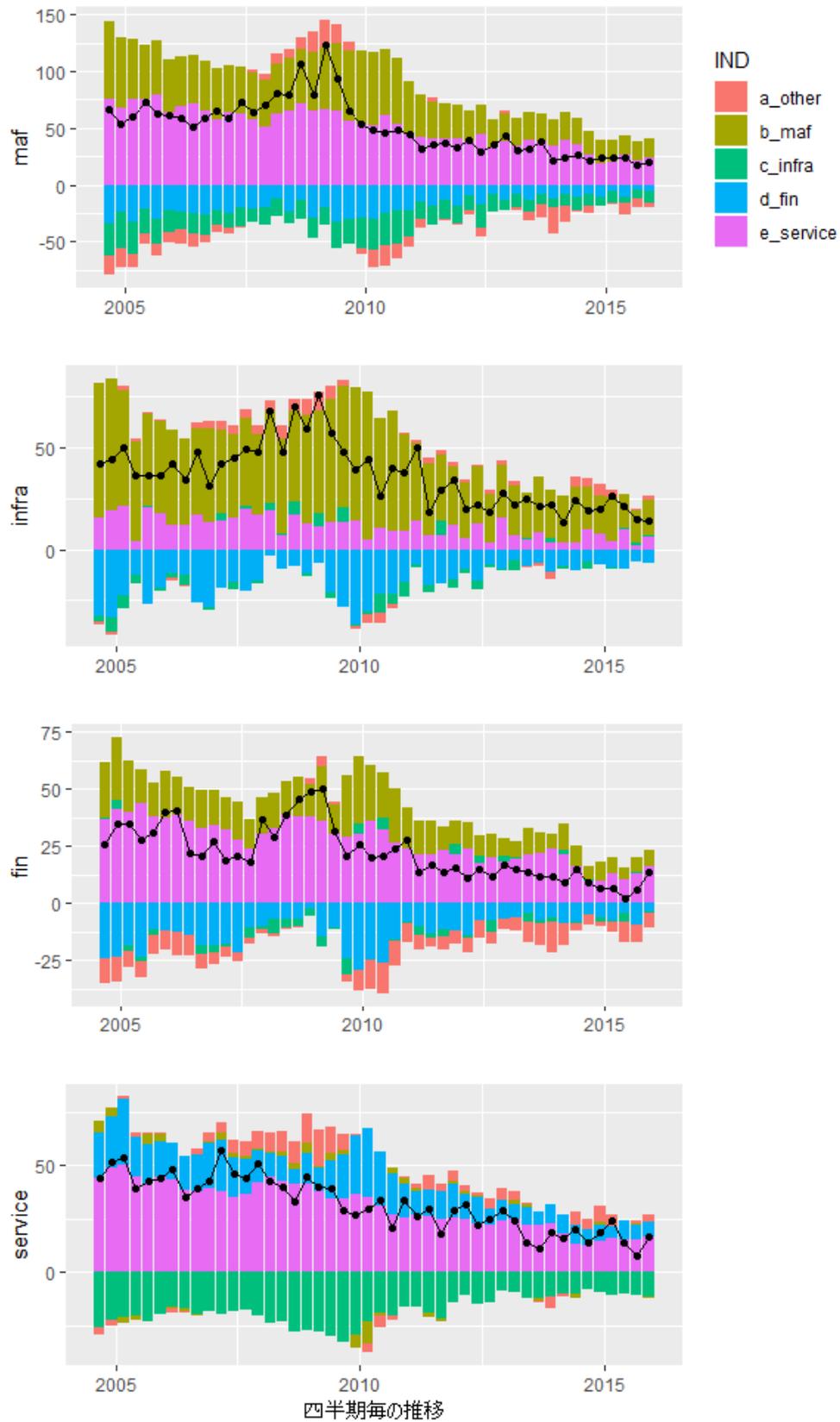


図4 マクロ経済指標を含んだ INAR モデルによる寄与度分解。Hawkes 推定量算出時の係数を用いて業種別、ラグ別、マクロ経済指標毎に推計し、ラグを全て足し合わせたものを業種別の寄与度としている。なお、縦軸は倒産発生件数を表し、黒線は倒産発生件数の実績値である。また、other はマクロ経済指標と切片の和を表している。

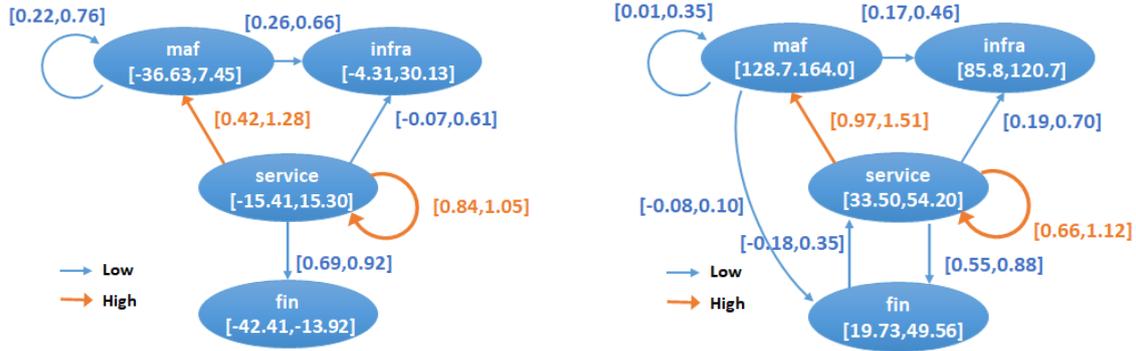


図5 業種カテゴリに対して Embrechts-Kirchner の推定法で得られた Hawkes グラフ表現。矢印に付随する区間は  $\hat{a}_{i,j}$  の 90% 信頼区間。各カテゴリの楕円内の区間は「頂点の重み」に当たる外生強度  $\mu^j$  の信頼区間（左：マクロ経済指標なし、右：マクロ経済指標あり）。90% 信頼区間の上側が 1 を超え、特に影響が大きいと考えられる矢印を High として強調している。

## 5.2 ストレストテストについて

本研究において、前節の議論によりマクロ経済指標を組み込むことによるモデルの有効性が示された。このモデルを実務において適用する際には本研究のモデルを基に予測値を算出することが考えられるが、予測値を利用したリスク管理への応用としてストレストテストが考えられる。一般的に倒産発生はマクロ経済環境に影響を受け、マクロ経済環境が悪化した場合には倒産件数も増加することが考えられる。しかし、マクロ経済指標を含まない INAR モデルについては、こうしたマクロ経済環境を勘案せずに予測を行うことから、図 6 のように平均値が固定されてしまうため、仮にマクロ経済環境が悪化した場合においてもこの挙動は変わらず、予測値と実績値に乖離が生じることが考えられる。

こうした事象に対して適切にリスク管理を行うため、金融機関においては将来考えられるマクロ経済環境をいくつか仮定し、仮定したマクロ経済環境の下でポートフォリオがどの程度毀損するかシミュレーションを行っている。ポートフォリオの毀損度合いについてはデフォルト率が大きな役割を果たし、本研究における倒産過程の将来予測を行うことで、業種間の伝播構造により倒産が波及していく様子をシミュレーションすることができる。

本研究においては、シナリオ値を定式化し、倒産過程のシミュレーションを行う。ストレストテスト実施に当たって、倒産事例データの終点である 2015 年 10 月 1 日～12 月 31 日を 4.1 節のマクロ経済指標の定義を用いて、 $M_n^{(\Delta)}$  とし、2016 年 1 月 1 日～2018 年 12 月 31 日をストレストテスト期間として定義してシナリオの種類を  $s \in [S]$  とすると、ストレストテスト期間におけるマクロ経済指標のシナリオ値は  $u \in \mathbb{N}$  に対して  $M_{n+u}^{(s,\Delta)}$  と表すことができる。

本研究では  $S = 5$ 、 $u \leq 12$  とし、 $M_{n+4}^{(s=5,\Delta)}$  で最小値を取るよう設定し、その他の時点を以下のように表すこととする。

$$\begin{cases} M_{n+u}^{(s=i,\Delta)} = M_n^{(\Delta)} - \frac{iu}{20}(M_n^{(\Delta)} - M_{n+4}^{(s=5,\Delta)}), & (1 \leq u \leq 4) \\ M_{n+u}^{(s=i,\Delta)} = M_n^{(\Delta)}(1 - \frac{i}{5} + \frac{u}{8}) + \frac{i}{5}M_{n+4}^{(s=5,\Delta)} - \frac{u}{8}M_{n+4}^{(s=1,\Delta)}, & (4 < u \leq 8) \end{cases} \quad (8)$$

このような設定の下、図 7 のようにマクロ経済指標のシナリオを仮定し、倒産過程のシミュレーションを行う。シナリオの仮定は結果の解釈する上で非常に重要であるため、本来であれば深度や影響を与える期間など、細かく設定する必要があるが、本研究ではモデルのマクロ経済指標に対す

る感応度を分析するためだけに使用するため、シンプルなシナリオに設定をしている。なお、シンプルでありつつも現実を踏まえたシナリオ設定にすべく、シナリオのボトムは1年後に設定している。また、経済環境が著しく悪化したとしてもその状態が維持されるケースは過去を見てを存在しないため、ボトムを迎えた後は上昇するような設定にしている。当該シナリオにより算出された結果は図8である。例えば、過去最悪値よりも深いシナリオであるシナリオ5では、サービス関係を除いて過去最大の倒産件数となったリーマン期に近い数字となり、インフラ・流通関係では過去最悪値を若干超える値となることが確認できる。過去の実績値との比較からも、極端に大きい件数や小さい件数になっていないことから、マクロ経済環境に対するモデルの感応度は実務で活用する上で適度な水準であることが確認できる。

こうしたストレステストの結果からも、倒産過程のモデルにはマクロ経済指標を含むことが好ましく、本研究におけるマクロ経済指標を含んだINARモデルによる分析が有用であることが分かる。

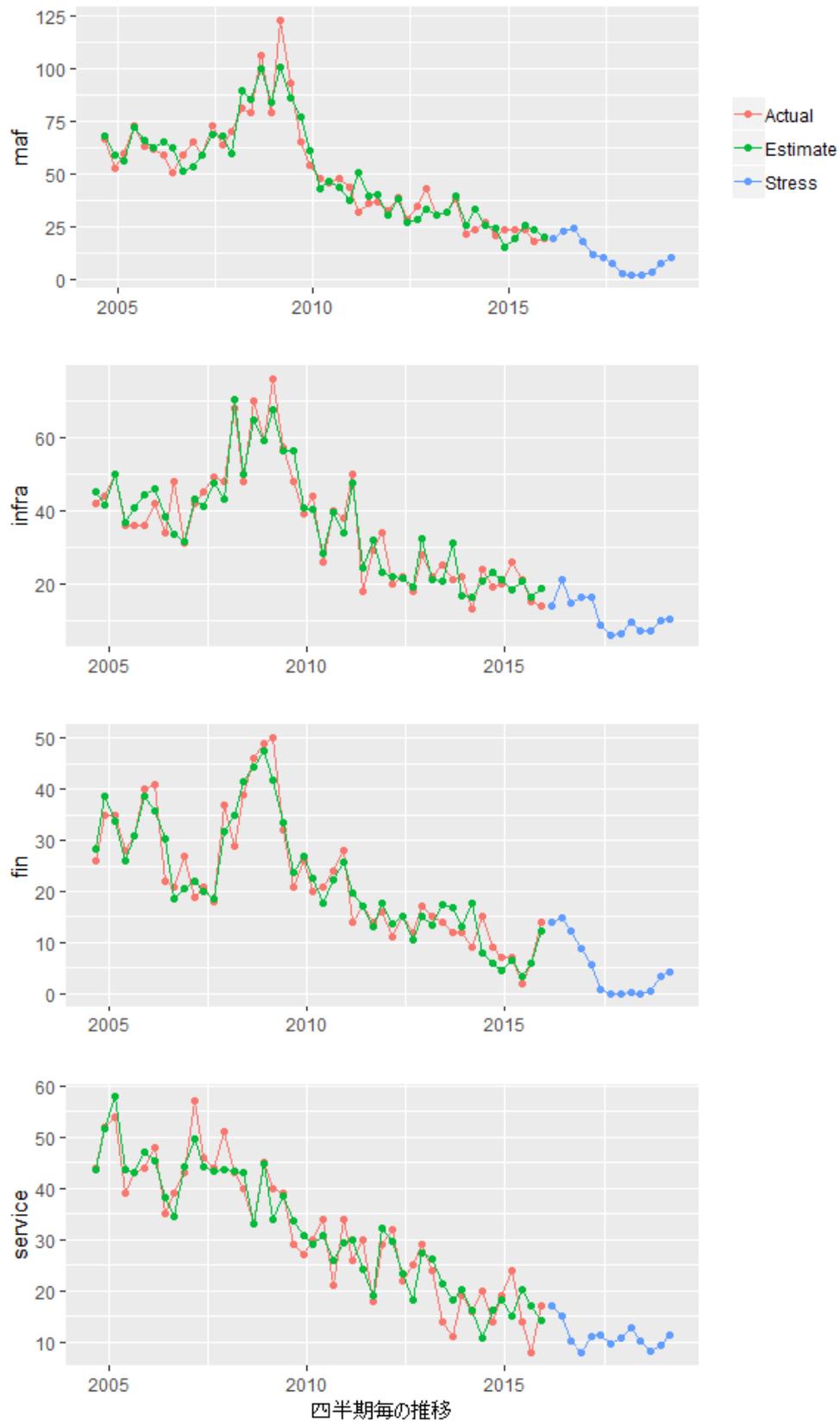


図6 マクロ経済指標を含まない INAR モデルにより、3 年後まで四半期ごとに予測を行ったもの。予測値は Hawkes 推定量算出時の係数を用いて時点毎に逐次的に算出される。

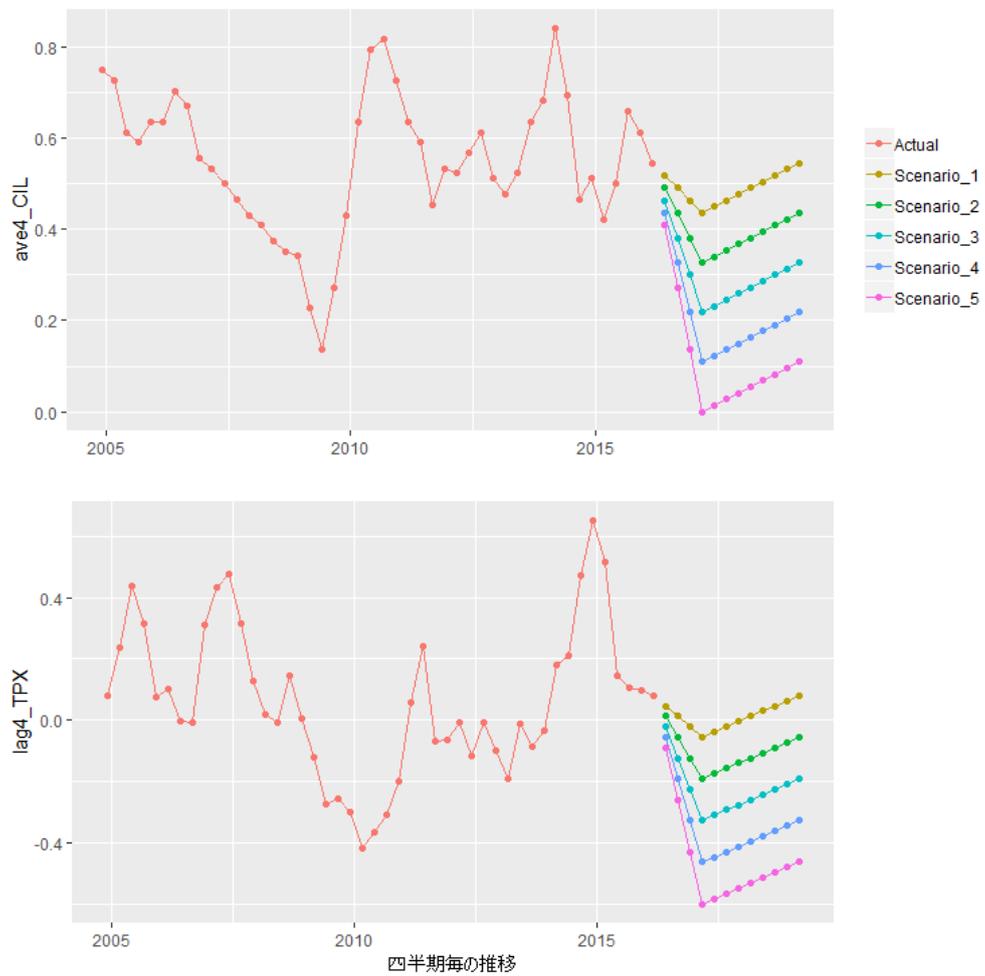


図7 ストレステスト実施に当たったストレスシナリオの一覧。景気動向指数（移動平均（1年））については、0を最小値としている。また、TOPIX（ラグ1年）については-0.6を最小値として定義している。

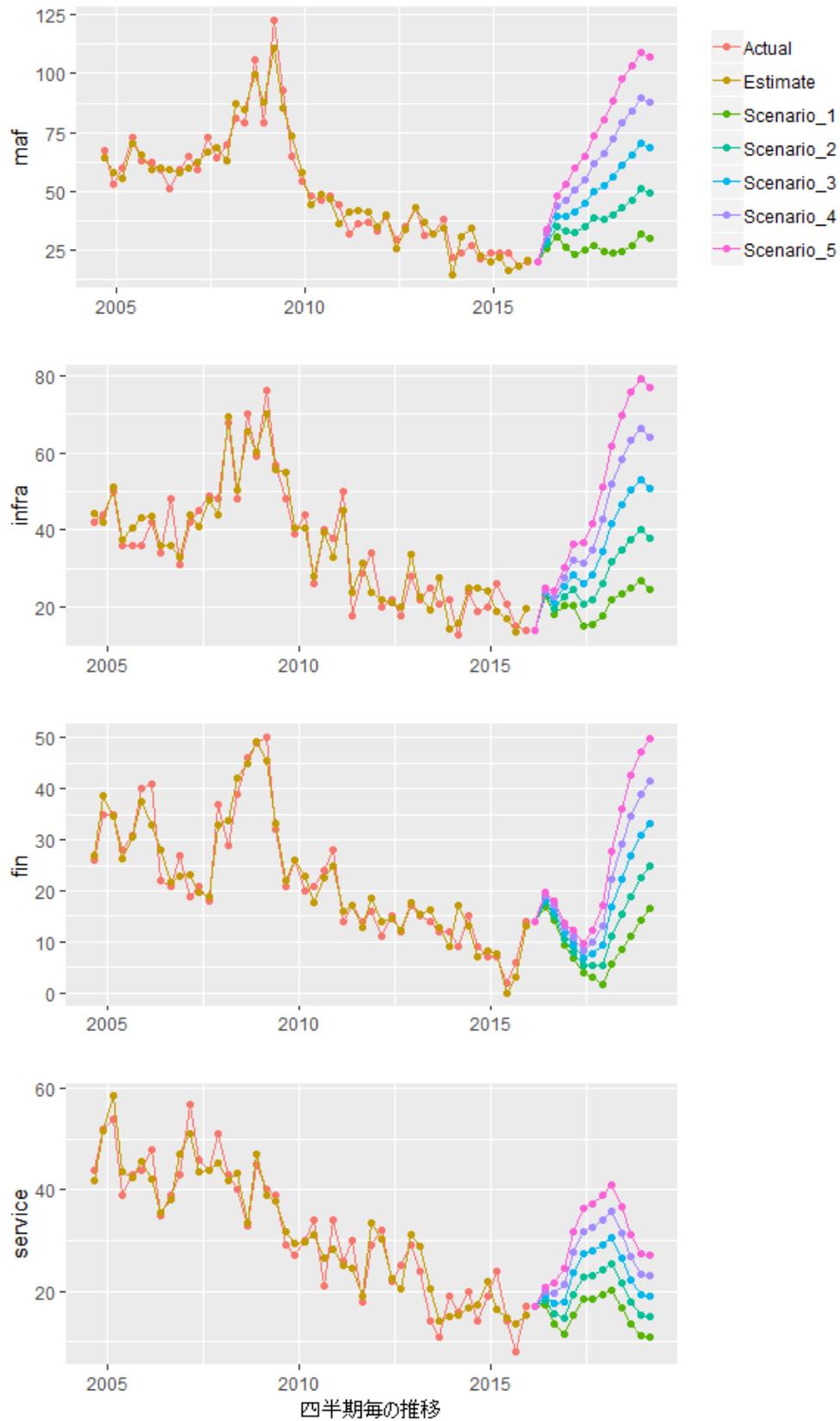


図8 マクロ経済指標を含んだ INAR モデルにより、3年後まで四半期ごとに予測を行ったもの。シナリオについては、図7を適用し、それぞれのシナリオ値を用いて時点毎に逐次的に予測値を算出している。

## 6 まとめ

本研究では、特に関東地区企業の比較的長期の倒産履歴データを対象に、定義した業種に基づいて区分し、倒産リスクの依存関係の分析を行った。分析では、Embrechts and Kirchner [3] により提唱されている INAR モデルによる分析を行ったほか、これらのモデルにマクロ経済指標を組み込んだ手法により比較を行い、実務の観点からもマクロ経済指標を含んだモデルのほうが有用性が高いことを示すことを目指した。

まず、倒産リスクの伝播構造を示す Hawkes グラフ表現に関して結果に大きな違いがなく、マクロ経済指標に関係なく伝播構造を捉えていることが確認できた。ただ、倒産件数がピークとなるリーマン期前後ではマクロ経済指標の効果が反映され、マクロ経済指標を含んだ INAR モデルの方が、AIC や平方二乗誤差といった評価指標に対して説明力が高いことが示された。

また、倒産件数の予測という観点からはマクロ経済指標を含むべきであり、複数のシナリオによりマクロ経済環境を仮定した際の倒産の予測件数は概ねリーマン期前後のピーク時の水準まで上昇し、モデルのマクロ経済指標に対する感応度も十分にあることが確認され、実務の観点からも有用性が示された。

一方で、Azizpour et al. [1] などで示唆されている結果を踏まえると、さらに潜在変数を組み込んだモデルに拡張することが考えられ、これらの事項については、今後の課題としたい。

## 参考文献

- [1] S. Azizpour, K. Giesecke, G. Schwenkler (2016), “Exploring the Sources of Default Clustering”, *Working paper*, <http://people.bu.edu/gas/papers/exploring.pdf> (accessed on April, 26, 2017). (To appear in *Journal of Financial Economics*.)
- [2] E. Bacry, I. Mastromatteo and J.-F. Muzy (2015), “Hawkes processes in finance”, *Working paper*, arXiv:1502.04592.
- [3] P. Embrechts and M. Kirchner (2016), “Hawkes graphs”, *Theory of Probability & Its Applications*, forthcoming.
- [4] E. Errais, K. Giesecke and L. R. Goldberg (2010), “Pricing credit from the top down with affine point processes”, *SIAM Journal on Financial Mathematics*, 1, 642–665.
- [5] K. Fokianos (2012), “Count time series models”, *Time series Analysis: Methods and applications, Handbook of Statistics 30* (edited by T.S. Rao, S. S. Rao, and C.R. Rao. Elsevier, Amsterdam), 315–347.
- [6] A.G. Hawkes (1971), “Spectra of self-exciting and mutually exciting point processes”, *Biometrika*, 58, 83–90.
- [7] M. Kirchner (2016), “An estimation procedure for the Hawkes process”, *Quantitative Finance*, forthcoming.
- [8] Y. Ogata (1981), “On Lewis’ simulation method for point processes”, *IEEE Transactions on Information Theory*, 27(1), 23–31.
- [9] Y. Ogata (1988), “Statistical models for earthquake occurrences and residual analysis for point processes”, *Journal of American Statistical Association*, 83(401), 9–27.

- [10] S. Yamanaka, M. Sugihara and H. Nakagawa, “Modeling of contagious credit events and risk analysis of credit portfolios,” *Asia-Pacific Financial Markets*, 19(1), 43–62.
- [11] 監物 輝夫・中川 秀敏 (2019), 「多次元 Hawkes 過程を用いた倒産リスク伝播構造の推定–Hawkes グラフ表現による視覚化–」, ジャフイージャーナル, 第 17 巻, 15–44.
- [12] 中川 秀敏 (2010), 「相互作用型の格付変更強度モデルによる格付変更履歴データの分析」, 日本応用数理学会論文誌, 20(3), 183–202.