

信用スコアリングモデルにおけるマクロファクターの導入と 推定デフォルト確率の一致精度の改善効果

枇々木 規雄
慶應義塾大学

尾木 研三
日本政策金融公庫

戸城 正浩

(受理 2011 年 10 月 25 日; 再受理 2012 年 3 月 1 日)

和文概要 小企業向けの信用スコアリングモデルは主に財務指標から個別企業のデフォルト確率を推定する統計モデルである。デフォルトは固有ファクターだけではなく、すべての企業に共通するマクロファクターの影響も受ける。推定デフォルト確率(推定 PD)と実績デフォルト率(実績 DR)の一致精度を高めるにはマクロファクターを説明変数に加えることが望ましいが、デフォルトに関する時系列データの蓄積が不十分なため、回帰分析などによって具体的な指標を特定することが難しいという課題がある。一方で、2007 年頃から始まった急速かつ大幅な景気悪化によって、実績 DR が推定 PD を上回る状況が続いており、マクロファクターを加味することの必要性が高まっている。そこで、本研究では日本政策金融公庫国民生活事業本部が保有する約 54 万件の豊富なデータを用いて時系列データの不足を補い、具体的な指標の選択を行った。その結果、マクロファクターとして前月デフォルト率が有効であり、前月デフォルト率を説明変数に追加した新モデルを構築すると、推定 PD と実績 DR との乖離が最大で 0.72%ポイント縮小するなど、推定 PD の一致精度を改善することができた。

キーワード: 金融, 信用リスク管理, 信用スコアリングモデル, ロジットモデル, 推定デフォルト確率, マクロファクター

1. はじめに

金融機関では、個別企業の信用リスクを客観的に評価するために信用スコアリングモデルを活用している。モデルは、構造モデル、誘導モデル、統計モデルの三つのタイプに大別されるが、未上場の中小企業を対象にしたモデルは、主に財務指標とデフォルトとの相関関係を利用して個別企業のデフォルト確率を推定する統計モデル¹が一般的である。

デフォルトは企業固有のファクターだけではなく、景気などすべての企業に共通するマクロファクター、すなわちシステムティックリスクの影響も受ける。したがって、統計モデルの推定デフォルト確率(以降、推定 PD: Probability of Default)の精度を向上させるには、マクロファクターを追加することが望ましい。その一方で、株価や社債データがない未上場の中小企業に対するデフォルト関連の時系列データは長くても 10 年程度しか蓄積できていない。蓄積期間が短いと時系列データの数が少なくなるため、回帰分析の精度が低くなり、マクロファクターとなる具体的な指標の特定が難しくなる。しかし、米国のサブプライムローン問題を契機とした景気悪化によって、実績デフォルト率(以降、実績 DR: Default Rate)が推定 PD を上回る状態が続いており、モデルにマクロファクターを導入することの必要性が高まっている。

そこで、本研究では日本政策金融公庫国民生活事業本部(以降、公庫)が保有する約 54 万

¹ほとんどの中小企業は株式を上場しておらず、社債も発行していないので、構造モデルや誘導モデルを使用することが難しく、統計モデルが一般的である。モデルの詳細については、森平 [14] を参照されたい。

件の豊富なデータ²から月次デフォルト率を算出することによって時系列データの不足を補い、マクロファクターとなる具体的な指標を特定する。さらに、マクロファクターを説明変数に加えた新モデルを構築して、推定 PD と実績 DR の一致精度について分析を行う。

先行研究をみると、デフォルトとマクロファクターとの関係分析やモデルにマクロファクターを追加することの有効性についての分析はいくつか存在する。たとえば、Sommar and Shahnazarianka [19] はベクトルエラーコレクションモデルを使った時系列分析を行い、物価、工業生産、短期金利の三つの変数がデフォルト率と相関があり、なかでも短期金利のインパクトが強いことを示している。Simons and Rolwes [17] はオランダ企業のデータとロジットモデルを用いた分析を行い、デフォルト率が GDP 成長率とは負の相関、金利、為替レート、原油価格とは正の相関があることを示している。Figlewski, Frydman and Liang [7] はムーディーズの社債データベースとコックス回帰モデルを使って、景気がデフォルトや格付け遷移にどのような影響を及ぼすかを分析している。工業生産や GDP 成長率などがデフォルトと強い相関があり、モデルに変数として追加するとパフォーマンスが向上することを明らかにしている。Bonfim [2] はポルトガルの銀行の融資先の約 3 万社のデータを用いた分析を行い、マクロファクターを取り入れるとモデルの説明力が改善することを示している。Bhattacharjee et al. [1] は米英の上場企業のパネルデータを使い、企業や業界の固有ファクターに加えたハザードモデルを構築して、マクロファクターが企業のデフォルトや買収の危険性に影響を与えることを示している。2002 年版中小企業白書 [3] は中小企業の倒産件数と実質 GDP 成長率に有意な負の相関があることを述べている。白田 [18] は株価、為替相場、金利水準、地価の 4 つの要因について倒産率との関係を分析し、倒産率は金利、地価との間に高い逆相関の関係があることを示している。大橋 [16] は企業倒産との関係について重回帰分析を行い、倒産件数 (対数) は株価と有効求人倍率で、倒産率は新設住宅着工件数、マネーサプライ、政府支出、公的需要の 4 つの変数で説明されることを示している。柳沢他 [20] は中堅・中小企業向けスコアリングモデルの経年劣化 (AR 値³の低下) にマクロファクターが影響を与えていることを示したが、信用スコアリングモデルの提案までは行っていない。

マクロファクターを追加したモデルとして、金子・中川 [12] は市場で観察可能な金利の期間構造を使い、企業格付けの推移が景気の先行き見通しに影響を受ける信用リスクポートフォリオのリスク計量モデルを提案している。森平・岡崎 [15] は上場企業の財務データと景気動向指数や日経平均株価、原油先物価格指数などを用いて、マクロファクターを加味した期間構造型の信用スコアリングモデルを提案し、推定 PD の一致精度⁴ とその有効性について検証している。

以上のように、マクロファクターを加味したモデルの提案を行っている先行研究はあるが、株価や社債データを使用できる大企業を対象にしたものであり、本研究の対象としている未上場の中小企業や小企業についての実証研究は見当たらない。そこで、本研究では以下の点を明らかにするために、公庫が 2005 年～2008 年の間に融資した約 54 万件の小企業のデータを使用して分析を行う。

²日本政策金融公庫国民生活事業本部は主に従業者数 20 人未満の小零細企業に融資しており、2010 年 3 月末時点で 1 社あたり平均 602 万円の小口債権を約 102 万件保有している。融資先の業種や地域構成比はわが国の事業所・企業統計調査の構成比に比べて大差なく、公庫の融資ポートフォリオは分散度が高い。公庫の融資ポートフォリオと事業所・企業統計調査との比較については、枇々木・尾木・戸城 [10] および公庫のホームページを参照されたい。

³AR 値 (Accuracy Ratio) および信用スコアリングモデルの評価方法については山下・三浦 [21] が詳しい。

⁴ここでいう一致精度とは「標本サイズが大きくなると、標本統計量が母数に収束していく性質」を指す「一緻性」という概念ではない。

- (1) 小企業向け信用スコアリングモデルに有効なマクロファクターを明らかにする．先行研究を参考にしてさまざまな経済関連指標を使って分析を行い，標準化回帰係数，符号条件，順序性尺度である AR 値などの比較を行う．
- (2) マクロファクターを加味した新モデルを構築し，推定 PD と実績 DR の一致精度を検証する．その際，業種別や格付け別などさまざまな切り口で分析することによって新モデルの頑健性を確認する．

分析の結果，マクロファクターの候補として株価 (TOPIX) や原油先物価格指数 (WTI)，鉱工業生産指数など，さまざまな経済関連指標を分析した結果，小企業の前月デフォルト率 (ストックベース) が最も有効であることがわかった．前月デフォルト率はサンプル期間に関わらず，標準化回帰係数や符号条件が安定した結果が得られた．前月デフォルト率を説明変数に加えた新モデルを構築すると，推定 PD と実績 DR のギャップを最大で 0.72% ポイント縮小することができた．さらに，格付け別，業歴別，業種別といったさまざまな切り口で検証しても，AR 値の水準を低下させることなく推定 PD の一致精度が改善され，モデルの頑健性が確認できた．

本論文の構成は以下のとおりである．第 2 節では統計モデルの説明変数にマクロファクターを追加することの理論的な背景を解説したあと，具体的な指標として前月デフォルト率が有効であることを示し，新モデルを構築する．第 3 節ではマクロファクターを導入した新モデルの推定 PD の一致精度について，格付け別，業歴別，業種別などさまざまな切り口で現行モデルとの比較を行う．第 4 節では新モデルが金融機関の収益に与える効果をシミュレーションによって検証する．第 5 節では結論と今後の課題を述べる．

2. マクロファクターを導入した新モデルの構築

本節では推定 PD の一致精度を改善するため，マクロファクターを導入した新モデルを構築する．具体的には説明変数として経済関連指標を追加することによって，財務内容などが同じでも景気が悪いときは推定 PD が高くなり，景気が良いときは推定 PD が低くなるというモデルを構築する．2.1 節で現行の小企業向け信用スコアリングモデルの概要を説明し，2.2 節で現行モデルの一致精度の低下の要因について考察する．2.3 節ではマクロファクターの感応度と資産相関の関係を示し，統計モデルの説明変数にマクロファクターを追加することがシステムティックリスクを加味することになる理論的な背景を説明する．2.4 節では説明変数に用いる具体的な経済関連指標の選択を行い，マクロファクターを導入した新モデルを構築する．

2.1. 現行モデルの概要

信用スコアリングモデルに用いられる統計モデルには，判別モデル，ロジットモデル，ハザードモデルなど，さまざまなタイプがある．なかでもロジットモデルは最も一般的に用いられており，CRD 協会や RDB といった代表的な中小企業向けモデルでも採用されている．公庫も 2 項ロジットモデルをベースに，貸付時点における小企業のデータを説明変数として貸付後 1 年以内にデフォルトする確率を推定する 1 期間モデルを構築しているが，財務指標に加えて業歴を説明変数に用いている点に特徴がある．小企業のデフォルト率はほぼ業歴の 3 次式で表すことができるうえ，財務指標との相関が低いため，業歴を変数として追加すると AR 値が 5～10% 程度改善する．しかも，業種別や従業者規模別の AR 値の改善幅は財務指標モデルの AR 値が低いカテゴリーほど大きく，カテゴリー間の AR 値の格差を縮小させ

る効果がある．以上のモデルから算出される推定 PD は概ね 0～100 点までのスコアに変換して，貸付時点における企業の信用力の相対的な評価にも用いている．具体的には以下の手順で信用スコアを算出している．

- (1) 企業 i の業歴変数 $(g_i)^k$ ($i = 1, \dots, n; k = 1, 2, 3$) と J 個の決算書 1 期分の財務指標を用いた変数 f_{ij} ($i = 1, \dots, n; j = 1, \dots, J$) を使用してロジットモデルを構築し，最尤法によってパラメータ α_j ($j = 1, \dots, J$) と β_k ($k = 1, 2, 3$) を推定する．財務指標の変数には全業種共通変数と業種固有の変数が含まれる⁵．パラメータ推定には SAS/STAT[®] の LOGISTIC プロシジャを使用している．ここで， p_i は倒産確率， n は企業数， g_i は業歴 (61 年以上は 61 年) を表す． z_i が大きければ大きいほど，倒産確率は低くなる．

$$p_i = \frac{1}{1 + e^{z_i}}, \quad z_i = \ln \left(\frac{1 - p_i}{p_i} \right) = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k \quad (i = 1, \dots, n) \quad (2.1)$$

- (2) 推定されたパラメータを用いて計算された z_i^* から企業 i の信用スコア CS_i

$$CS_i = \eta_0 + (\eta_1 - \eta_0) \left(\frac{z_i^* - Z(1\%)}{Z(99\%) - Z(1\%)} \right)$$

を計算する．ここで， $Z(1\%)$, $Z(99\%)$ はモデル構築時のインサンプルデータにおける z_i^* の 1 パーセント点，99 パーセント点を表す．信用スコアが $z_i^* = Z(1\%)$ ならば η_0 点， $z_i^* = Z(99\%)$ ならば η_1 点となるように基準化している．本研究では $\eta_0 = 10$, $\eta_1 = 90$ としている． z_i^* を直接用いても結果に影響を与えない．

2.2. 推定 PD の一致精度が低下した要因

2007 年頃の米国のサブプライムローン問題から始まった世界的な景気悪化とその後のリーマンショックを契機とした金融危機は，中小企業をかつてないほど厳しい状況に追い込んだ．図 1 に RDB 企業のデフォルト率を指数化 (2005 年 4 月 = 100) したグラフを示す．中小企業のデフォルト率はサブプライムローン問題が表面化した 2007 年ごろから上昇し始め，リーマンショック後のピークを迎えるまでの約 2 年間で 60% もの上昇を記録した．急速かつ大幅な景気悪化の影響で中小企業の実績 DR が急上昇していることがわかる．

今回の景気悪化が現行の信用スコアリングモデルに与えた影響を公庫のデータを用いて確認する．具体的には 2005 年～2006 年に融資した約 26 万件のデータを用いて，現行モデルのパラメータを再推計したモデルを構築し，2007 年と 2008 年のデータをアウト・オブ・サンプルとして AR 値と推定 PD の一致精度を検証した．AR 値の結果を表 1 に示す．2007 年はインサンプルとほぼ同じ水準だが，2008 年は 4.5% ポイント低下している．

次に推定 PD と実績 DR の一致精度を評価する．代表的な指標として (2.2) 式に示すブライア (Brier) スコアがある．ここで， N は企業数， p_i は個別企業の推定 PD， 1_i はデフォル

⁵業種を考慮して厳密に (2.1) 式を記述するならば，以下のように記述する必要がある．

$$z_i = \ln \left(\frac{1 - p_i}{p_i} \right) = \alpha_0 + \sum_{j \in V} \alpha_j f_{ij} + \sum_{s \in S} \sum_{j \in V_s} 1_i^s \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k \quad (i = 1, \dots, n)$$

ここで， V は全業種共通変数の集合， V_s は業種 s の固有の変数の集合 (これらの変数集合の数の合計は J である)， S は業種 s の集合を表す．また， 1_i^s は企業 i が業種 s に含まれていれば 1，含まれていなければ 0 となる指示関数である．煩雑さを避けるためにモデル上は財務変数を分けずに記述する．

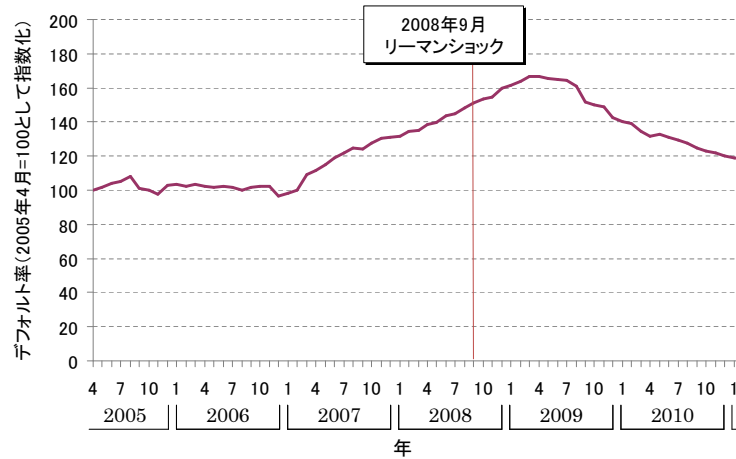


図 1: RDB 企業デフォルト率の推移

表 1: AR 値の推移 (IN : 2005 ~ 2006 年 , OUT : 2007 ~ 2008 年)

2005 年	2006 年	2007 年	2008 年
46.10%	49.43%	47.76%	43.23%

トならば 1 , デフォルトしていなければ 0 となる指示関数である . したがって , ブライアスコアは 0 に近いほど一致精度が高いと評価する .

$$\text{ブライアスコア} = \frac{1}{N} \sum_{i=1}^N (p_i - 1_i)^2 \quad (2.2)$$

表 2 にブライアスコアを示す . スコアが年々大きくなっており , 一致精度が悪化している可能性がある . ただ , ブライアスコアは推定 PD の水準が低いとかなり小さな値をとる . 差がわずかであるため , 2005 年から 2008 年は 0 の近傍でほぼ横ばいと評価することもできる . このようにブライアスコアではモデル間の比較や改善度合いを評価しにくい面もあるため , 本研究ではグラフと実数値によって一致精度の評価を行う .

表 2: ブライアスコア (IN : 2005 ~ 2006 年 , OUT : 2007 ~ 2008 年)

2005 年	2006 年	2007 年	2008 年
0.0053801	0.0090510	0.0142918	0.0171396

図 2 に推定 PD と実績 DR の推移を示す⁶ . 2007 年と 2008 年は実績 DR が推定 PD を大きく上回り , とりわけ , 2008 年は 0.96% ポイント乖離するなど , 推定 PD の一致精度が低下している .

本研究の分析結果と同じように , 本邦で代表的な中小企業向けモデルの CRD モデルにおいても , 推定 PD が実績 DR を下回るという過小評価が起きている . AR 値の水準が維持されたまま , 推定 PD の一致精度だけが低下しており , マクロファクターがモデルに加味されていない影響が出ていると推察される . CRD 協会は要因の一つとしてデフォルト情報の蓄

⁶デフォルト率 (変換後) は , 2005 年の実績 DR を 1% に変換して , 他をパラレルシフトした数値を示している .

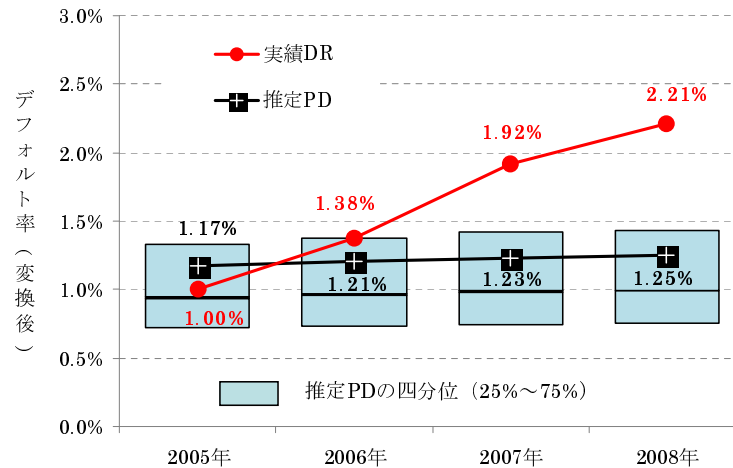


図 2: 推定 PD と実績 DR の推移 (IN : 2005 ~ 2006 年 , OUT : 2007 ~ 2008 年)

積が不十分であることを指摘し、定数項を上方修正して推定 PD を調整する対策を講じている [4-6] .

2.3. マクロファクターの導入とシステムティックリスクの関係

推定 PD の一致精度を改善するにはモデルにシステムティックリスクを考慮する必要がある . 本研究では統計モデルの説明変数に特定のマクロファクター x を追加する方法を採用する . 本節ではその統計モデルが示す企業のデフォルト構造を確認し、説明変数にマクロファクターを追加することがシステムティックリスクを加味することになる理論的な背景を説明する . 本研究では最も普及しているロジットモデルを用いているが、説明をわかりやすくするためにリンク関数として標準正規分布を用いているプロビットモデルに置き換えて確認する⁷ . まず、標準正規分布に従う誤差項 $\tilde{\varepsilon}_i$ を含めて表記した現行モデルの回帰式を (2.3) 式に示す .

$$\tilde{z}_i = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k + \tilde{\varepsilon}_i \quad (2.3)$$

\tilde{z}_i は企業 i の信用度を表す変数である . 現行モデルで推定した企業 i のデフォルト確率 p_i は (2.4) 式に示すとおり、 \tilde{z}_i が閾値 Q を下回った場合にデフォルトが発生すると想定している .

$$p_i = \Pr(\tilde{z}_i \leq Q) \quad (2.4)$$

ここで、 $\alpha'_0 = \alpha_0 - Q$ とすると、(2.4) 式は (2.5) 式に書き直すことができる .

$$p_i = \Pr \left[\tilde{\varepsilon}_i \leq - \left(\alpha'_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k \right) \right] \quad (2.5)$$

誤差項 $\tilde{\varepsilon}_i$ は企業 i の資産価値変動の代理変数と考えられ、

$$Q'_i = - \left(\alpha'_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k \right) \quad (2.6)$$

⁷マダラ [13] は標本が大きい場合、ロジットモデルとプロビットモデルは推定結果に差異がないことを示している .

とすると、(2.5) 式は信用度の変動 $\tilde{\varepsilon}_i$ が企業固有の閾値 Q'_i を下回った場合にデフォルトが発生すると考えることができる。

次に現行モデルにシステムティックリスクを考慮するために、企業 i の信用度の変動 $\tilde{\varepsilon}_i$ はマクロファクター \tilde{X} と企業固有のファクター $\tilde{\xi}_i$ によって変動すると考え、(2.7) 式に示すような1要因モデルで記述されると想定する⁸。

$$\tilde{\varepsilon}_i = b\tilde{X} + \sqrt{1-b^2}\tilde{\xi}_i \quad (2.7)$$

ここで、 \tilde{X} と $\tilde{\xi}_i$ はそれぞれ基準化を行った値とし、お互いに独立な標準正規分布に従うと仮定する。(2.5) 式に (2.6), (2.7) 式を代入すると

$$p_i = \Pr\left(b\tilde{X} + \sqrt{1-b^2}\tilde{\xi}_i \leq Q'_i\right) = \Pr\left(\tilde{\xi}_i \leq \frac{Q'_i}{\sqrt{1-b^2}} - \frac{b}{\sqrt{1-b^2}}\tilde{X}\right) \quad (2.8)$$

マクロファクター \tilde{X} が特定の値 x をとったときの条件付きデフォルト確率 $p_i(t|\tilde{X} = x)$ は (2.9) 式で表される。

$$p_i = \Pr\left(\tilde{\xi}_i \leq \frac{Q'_i}{\sqrt{1-b^2}} - \frac{b}{\sqrt{1-b^2}}x\right) = \Phi\left(\frac{Q'_i}{\sqrt{1-b^2}} - \frac{b}{\sqrt{1-b^2}}x\right) \quad (2.9)$$

ここで、 $\Phi(\cdot)$ は標準正規分布の累積分布関数を表す。システムティックリスクを考慮した統計モデルで推計したデフォルト確率は代表的なモデルである1ファクター・マートン型モデルと同じ形になることがわかる。Gordy [8], Gordy and Heitfeld [9] がその枠組みを拡張させ、BIS II 規制のリスクウェイト関数の理論的な背景にもなっている。1ファクターモデルでは b はマクロファクターに対する感応度、 b^2 は企業間の資産相関を表し、重要なパラメータとなる。(2.7) 式を (2.3) 式に代入し、企業固有の信用度 \tilde{z}_i を基準化すると (2.10) 式のようなになる。

$$\tilde{z}'_i = \frac{\alpha_0}{\sqrt{1-b^2}} + \sum_{j=1}^J \frac{\alpha_j}{\sqrt{1-b^2}} f_{ij} + \sum_{k=1}^3 \frac{\beta_k}{\sqrt{1-b^2}} (g_i)^k + \frac{b}{\sqrt{1-b^2}}\tilde{X} + \tilde{\xi}_i \quad (2.10)$$

統計モデルにおけるマクロファクターの回帰係数は、(2.11) 式のように1ファクターモデルの資産相関 b^2 で表現できる。

$$\text{マクロファクターの回帰係数} = \frac{b}{\sqrt{1-b^2}} \quad (2.11)$$

ここではプロビットモデルで示したが、ロジットモデルを用いても同じである。標準正規分布の分散は1、ロジスティック分布の分散は $\pi^2/3$ であるので、ロジットモデルを用いた場合は、回帰係数を $\sqrt{3}/\pi$ 倍した標準化回帰係数にする必要があるものの、説明変数にマクロファクターを追加することによって、1ファクターモデルにおける資産相関 b^2 を推計することができる。

⁸マクロファクターは時間の関数 $\tilde{X}(t)$ と考えることもできるが、ここでは i.i.d. を仮定し、 \tilde{X} とする。

2.4. マクロファクターの特定と新モデルの構築

統計モデルの説明変数に GDP や金利といった具体的な指標としてマクロファクターを特定し、デフォルト率との関連を調べるためにロジスティック回帰分析をする必要がある⁹。しかし、年ベースのデフォルト率のデータは蓄積が不十分で分析が困難という課題がある。

そこで、本研究では経済関連指標とのロジスティック回帰分析を可能にするために、公庫が保有する約 54 万件の豊富なデータから月次のデフォルト率を算出することによって時系列データの不足を補う。ロジスティック回帰分析によってマクロファクターとなる経済関連指標を特定したあと、説明変数にマクロファクターを加えた 1 期間の信用スコアリングモデルを構築する。具体的には 2005 年～2008 年に公庫が融資した小企業（法人）のデータを使用し、(2.1) 式の現行モデルの説明変数に経済関連指標 $h_{l,t-1}$ ($l = 1, \dots, L$) を加えた (2.12) 式のような新モデルを構築する。 $h_{l,t-1}$ は前月の経済関連指標であり、 L は経済関連指標の数を表す。

$$z_i = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k + \sum_{l=1}^L \gamma_l h_{l,t-1} \quad (2.12)$$

マクロファクターとなる具体的な経済関連指標の選択を行う。デフォルトに関する時系列データの不足を補うため、月次ベースの指標を用いることにする。月次でデータが提供されている経済関連指標のなかから、デフォルトと相関が高いと考えられる約 20 個の指標を使用して有効な指標を抽出する。具体的には、株価 (TOPIX)、原油先物価格指数 (WTI)、鉱工業生産指数 (IIP)、国債利回り、日本株投資家の不安心理の度合いを示す恐怖指数 (VXJ: Volatility Index Japan)¹⁰などのデータの原数値のほか、適宜、対数、階差、変化率などに変数変換したデータを作成する。経済関連指標はお互いに相関が強く、複数の変数を用いると多重共線性が発生しやすいため、指標を売上（生産・販売）動向、資産価格、原材料価格、金利、恐怖指数の 5 つのカテゴリーに分類したうえで、相関と符号条件を勘案して各カテゴリーから代表的な指標を一つずつ選択する。相関や散布図などで分析した結果、最終的に鉱工業生産指数 (IIP, 季節調整値)、株価 (TOPIX, 月中平均値)、原油先物価格指数 (WTI, 月中平均値)、金利 (新発 10 年国債利回り, 月中平均値)、恐怖指数 (VXJ, 月中平均値) を選択した。

さらに、これらの経済関連指標に加えてストックベースの前月デフォルト率 (前月 DR) を用いることにする。今井・清水 [11] は中小・中堅企業のデフォルト率が株式 (TOPIX) に対して先行性をもつ景気指標であることを示している。また、森平・岡崎 [15] は基底デフォルト確率を考慮すると、株価 (日経平均株価, 日経平均株価ボラティリティ) や原油先物価格指数 (WTI) などが有意でなくなることから、基底デフォルト確率と経済関連指標は代替関係にあることを示している。さらに、森平 [14] は信用保証協会の件数ベースの弁済率が 1 階の自己回帰モデルで相当程度説明できることを述べている。以上の先行研究からマクロファクターとして前月デフォルト率が有効な変数になる可能性があると考え変数候補とする。具体的には融資実行月の前月の公庫のローン・ポートフォリオの実績 DR を用いる。

多重共線性を回避するために変数を単独で採用したモデルの推定結果を表 3 に示す。パラ

⁹現行モデルと同様に、SAS/STAT® の LOGISTIC プロシジャを使用して、最尤法を用いたパラメータ推定を行っている。

¹⁰大阪大学 金融・保険教育研究センターが、大阪証券取引所で取引された日経 225 オプションの日次終値から算出しているインプライド・ボラティリティの指標である。計算方法については、大阪大学 金融・保険教育センターのホームページを参照されたい。

メータとAR値、符号条件を比較する。鉱工業生産指数，TOPIXの回帰係数の符号がマイナスになり，整合性が得られない。前月デフォルト率とWTIは同程度のパフォーマンスを示しているが，金利(国債利回り)，VXJは両者に比べて標準化回帰係数やAR値のパフォーマンスがやや劣ることがわかる。

表 3: マクロファクターのAR値と標準化回帰係数の比較

	景気変数	なし	IIP	TOPIX	金利	WTI	VXJ	前月DR
2005年1月～ 2006年12月貸付 ($n=263,715$)	AR値	48.43	49.86	49.43	49.64	49.09	49.37	49.00
	(差)		(1.43)	(1.00)	(1.21)	(0.66)	(0.94)	(0.57)
	回帰係数 (p値)		-0.139 (0.00)	-0.119 (0.00)	-0.130 (0.00)	-0.102 (0.00)	-0.089 (0.00)	-0.086 (0.00)
	符号条件		×	×				
2005年1月～ 2007年12月貸付 ($n=398,010$)	AR値	49.07	51.89	50.74	50.27	50.39	50.20	50.90
	(差)		(2.82)	(1.67)	(1.20)	(1.33)	(1.13)	(1.83)
	回帰係数 (p値)		-0.206 (0.00)	-0.166 (0.00)	-0.128 (0.00)	-0.136 (0.00)	-0.119 (0.00)	-0.154 (0.00)
	符号条件		×	×				
2005年1月～ 2008年12月貸付 ($n=541,423$)	AR値	47.51	48.64	47.51	47.68	49.36	48.75	50.23
	(差)		(1.13)	(0.00)	(0.17)	(1.85)	(1.24)	(2.73)
	回帰係数 (p値)		-0.135 (0.00)	- (0.00)	-0.042 (0.00)	-0.137 (0.00)	-0.098 (0.00)	-0.178 (0.00)
	符号条件		×	-				

AR値の単位は%であり，AR値の下段の括弧内は景気変数なしとの差分である

1ファクターモデルを使ってモーメント法により推定した資産相関から導出した b の値¹¹と(2.11)式の標準化回帰係数から導出した b の値を前月デフォルト率とWTIについて比較すると，水準はどちらも近似しているが，WTIの方が感応度の変化が小さいことがわかる。

表 4: b の値の比較

貸付期間(貸付件数)	モーメント法	前月DR	WTI
2005年1月～2006年12月貸付($n=263,715$)	-11.08%	-8.54%	-10.13%
2005年1月～2007年12月貸付($n=398,010$)	-15.35%	-15.23%	-13.46%
2005年1月～2008年12月貸付($n=541,423$)	-16.22%	-17.50%	-13.55%

この5種類の変数すべてを説明変数として含め，有意水準5%のp値を基準として変数増減法によりステップワイズにパラメータを推定した。しかし，マクロファクターはお互いに相関が強く多重共線性が発生しやすいため，サンプルを変えると採用される変数が変わったり，符号条件が合わないなど安定的な結果が得られなかった。そこで，本研究ではマクロファクターを単独で追加することにする。

¹¹モーメント法による資産相関の推定方法は付録Aを参照されたい。

単独で使用する場合，WTIと前月デフォルト率が良好なパフォーマンスを示したが，前月デフォルト率は，業種別や地域別など様々な切り口でも算出可能であることから，今後の拡張性を勘案し，本研究では前月デフォルト率を採用する．最終的なモデルは(2.13)式のとおりである．ここで，融資時点の年月を t とすると， p_i は t 月に融資した企業の1年後のデフォルト確率(年次)， h_{t-1} は1カ月前のデフォルト率¹²， γ は前月デフォルト率のパラメータである．

$$p_i = \frac{1}{1 + e^{z_i}}, \quad z_i = \ln\left(\frac{1 - p_i}{p_i}\right) = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k + \gamma h_{t-1} \quad (2.13)$$

表 5: 新モデルのパラメータ

貸付時点(月)		2005年1月～ 2006年12月		2005年1月～ 2007年12月		2005年1月～ 2008年12月	
サンプル数		263,715		398,010		541,423	
AR 値		49.00%		50.90%		50.23%	
業種	変数	回帰係数	p 値	回帰係数	p 値	回帰係数	p 値
全業種 共通	財務比率 1	-0.0735	0.0000	-0.0796	0.0000	-0.0869	0.0000
	財務比率 2	-0.2206	0.0000	-0.2173	0.0000	-0.2027	0.0000
	財務比率 3	0.1498	0.0000	0.1770	0.0000	0.1883	0.0000
	財務比率 4	-0.0902	0.0000	-0.0828	0.0000	-0.0885	0.0000
	財務比率 5	0.0718	0.0000	0.0416	0.0001	0.0264	0.0014
	財務比率 6	-0.1266	0.0000	-0.1322	0.0000	-0.1260	0.0000
	財務比率 7	0.0716	0.0000	0.0779	0.0000	0.0780	0.0000
	財務比率 8	-0.0699	0.0000	-0.0734	0.0000	-0.0667	0.0000
	財務比率 9	-0.1522	0.0000	-0.1477	0.0000	-0.1244	0.0000
業種 1	財務比率 10	—	—	—	—	0.0424	0.0000
業種 2	財務比率 11	-0.0790	0.0000	-0.0743	0.0000	-0.0734	0.0000
業種 3	財務比率 12	-0.0362	0.0013	-0.0305	0.0000	-0.0252	0.0000
業種 4	財務比率 13	—	—	—	—	—	—
業種 5	財務比率 14	—	—	—	—	-0.0264	0.0018
業種 6	財務比率 15	0.1080	0.0000	0.0745	0.0000	0.0599	0.0000
全業種 共通	業歴(1乗)	1.6970	0.0000	1.7396	0.0000	1.5903	0.0000
	業歴(2乗)	-2.9035	0.0000	-2.9748	0.0000	-2.7020	0.0000
	業歴(3乗)	1.4741	0.0000	1.5106	0.0000	1.3849	0.0000
	前月 DR	-0.0857	0.0000	-0.1541	0.0000	-0.1778	0.0000

表中の「—」は変数選択で採用されなかった変数である

前月デフォルト率の標準化回帰係数は景気が安定していた2005年～2006年をインサンプルとしたモデルでは-0.0857であるのに対し，リーマンショック前後の2007年，2008年の

¹² $t-2$ 月時点のローン・ポートフォリオ(ストック)から $t-1$ 月に発生したデフォルト率(月次)とする．

データを含めたモデルでは -0.1541 , -0.1778 と高くなっている．景気後退期においてマクロファクターがデフォルトに与える影響が大きくなることが確認できる．景気変動するとデフォルト率も変動し，その結果，両者の相関が高くなるからである．このようにマクロファクターの感応度は一定ではない．マクロファクターを導入した統計モデルを構築する際には時間とともに景気状況が変わり，それによってマクロファクターの感応度も変わることを勘案する必要がある．本研究においてはデータ期間の制約からこの点についての分析が困難であるため，今後の検討課題としたい．

3. 新モデルの有効性

本節では，新モデルの推定 PD の一致精度について現行モデルと比較しながら検証する．その際，AR 値 (序列性) が低下していないのかも併せて確認する．本来，説明変数にマクロファクターを加えても，すべての企業に同一のインパクトを与えるので AR 値には影響しないはずである．しかし，本モデルは，説明変数に用いるマクロファクターは月単位である一方，デフォルトの観測は年単位で行っているため，AR 値が乱れる可能性がある．検証にあたっては，格付け別，業歴別，業種別にも検証し，新モデルの推定 PD の一致精度が頑健であることを示す．

3.1. 推定 PD の一致精度の改善効果

新モデルの推定 PD と分散の推移を図 3 に示す．比較のために図 2 をみると，現行モデルでは推定 PD の水準も分散も徐々に上昇・拡大してはいるものの，変化は極めてゆるやかで小さく，実績 DR との乖離は広がるばかりである．一方，図 3 (左) の新モデルによる推定 PD の水準は，2007 年が 1.55% ，2008 年が 1.87% と大きく上昇し，実績 DR とのギャップは 0.37% ポイント， 0.34% ポイントに縮小した．グラフの四分位範囲 ($25\% \sim 75\%$) も拡大しており，デフォルト率の水準が高くなるほど，分散も大きくなる様子が確認できる．

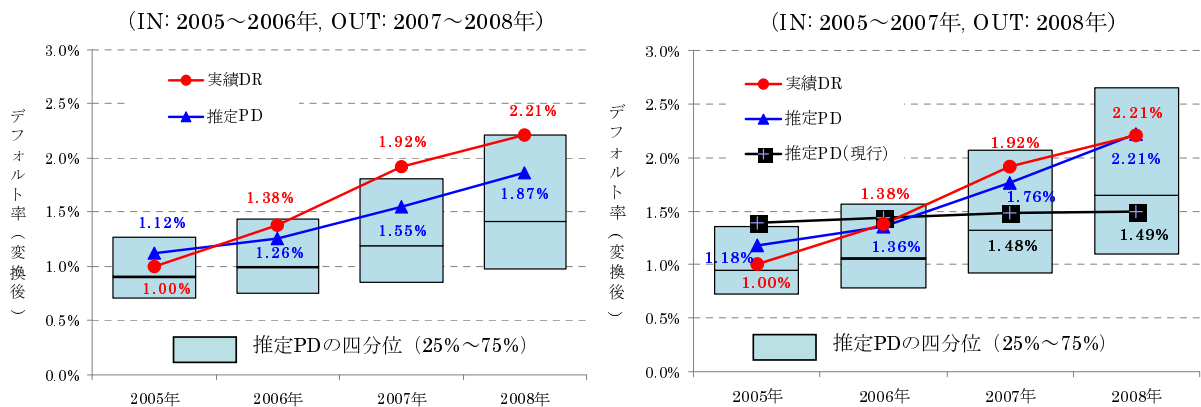


図 3: 新モデルの推定 PD と実績 DR の推移

次に，2005 年～2007 年のデータをインサンプルにしたモデルを構築し，2008 年のデータでアウト・オブ・サンプルテストした結果を図 3 (右) に示す．2008 年の推定 PD は 2.21% に修正され， 0.72% ポイントあった実績 DR とのギャップがほぼ解消されている．

表 6 に AR 値の推移を示す．新モデルになっても AR 値に大きな変化はなく，AR 値の水準を維持したまま，PD の一致精度だけが改善されている．

表 6: AR 値の変化

	インサンプル	アウトオブサンプル	
	2005年～2006年	2007年	2008年
現行モデル	48.4%	47.8%	43.2%
新モデル	49.0%	47.5%	42.4%
	2005年～2007年	2008年	
現行モデル	49.1%	43.6%	
新モデル	50.9%	42.9%	

データ期間が短いため効果の大きさには留意する必要があるものの、マクロファクターを導入した新モデルを使用することにより、AR 値の水準を維持したまま景気悪化時に発生する推定 PD の過小評価を修正できることがわかった。また、マクロファクターの感応度は時間とともに変化するため、できる限り近時のデータを使って推計することが望ましい。

3.2. デフォルト感応度の改善

図 3 をみると、景気が悪化すると推定 PD の水準が高くなるだけでなく、分散も拡大している様子が確認できる。スコアが低くなるほど、スコア 1 単位あたりの変動が推定 PD に与える影響が大きくなる。これは図 4 に示したように、推定 PD の分布はロジスティック分布に従っているため、景気の悪化によってスコアの水準が低くなるほど、横軸のスコアの分布が左方にシフトする結果 (①→②)、縦軸の推定 PD の分布が拡大する (①→②)。図 2、図 3 において四分位範囲の 50% 線が下方に位置しているのは、スコアの水準が低くなるほど推定 PD の変動が大きくなるためである。

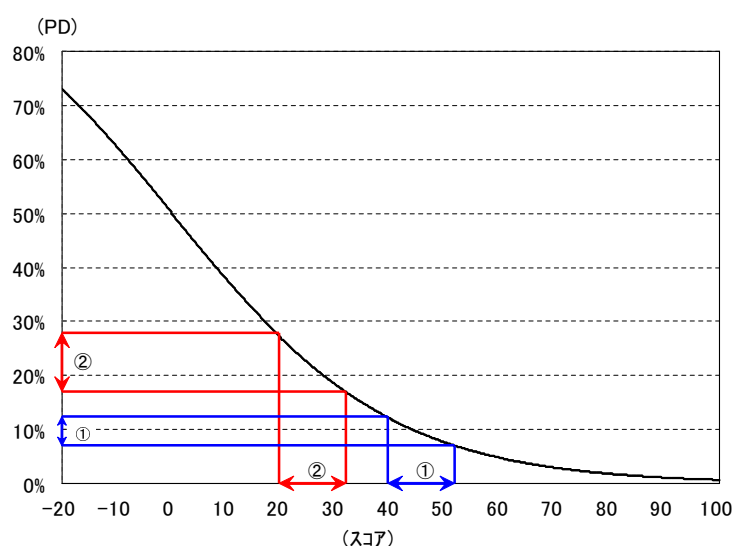


図 4: スコアと推定 PD の関係 (模式図)

森平 [14] はロジスティック回帰におけるパラメータ α_j は j 番目の財務指標が 1 単位変化したときの推定 PD に与える影響を示している。しかし、分布関数が非線形であるので PD

に対する企業 i における j 番目の財務指標が微小変化したときの影響は

$$\Delta_{ij} \equiv \frac{\partial p_i}{\partial f_{ij}} = \frac{\partial}{\partial z_i} \left(\frac{1}{1 + e^{z_i}} \right) \times \frac{\partial z_i}{\partial f_{ij}} = -\frac{e^{z_i}}{(1 + e^{z_i})^2} \cdot \alpha_j = -\alpha_j p_i (1 - p_i)$$

となる．推定PDのデルタは係数 α_j だけではなく，推定PDの水準にも依存することを示している．つまり，スコアの水準が低くなる（推定PDが高くなる）と個別企業の財務指標の変化がデフォルトに与える影響が大きくなるのである．これは財務内容の悪い企業ほど景気の影響を受けてデフォルトしやすくなるという現場感覚とも整合的である．

3.3. カテゴリー別検証からみた新モデルの頑健性

新モデルの有効性は業種別などのカテゴリーごとに異なる可能性がある．そこで，さまざまな切り口で推定PDの一致精度を検証して新モデルの頑健性を確認する．

3.3.1. 格付け別推定PDの一致精度

格付け別の推定PDと実績DRを図5～図7に示す．格付けはインサンプルの企業を推定PDの低い企業から順番に並べたあと，全体を均等に8区分して1格から8格まで付与した．図は推定PDと実績DRの折れ線グラフおよびそのギャップを棒グラフで示したもので，左側が現行モデル，右側は新モデルである．

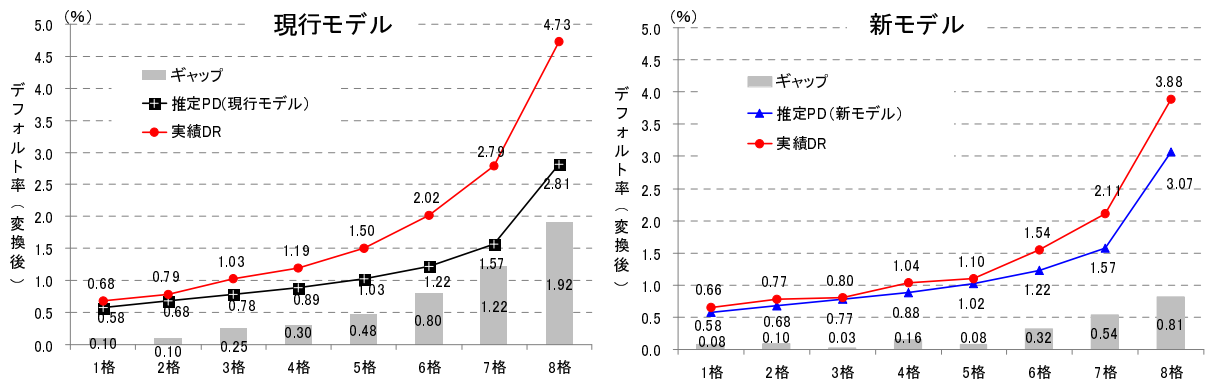


図 5: 格付け別推定PDと実績DRの推移 (その1: IN: 2005～2006年, OUT: 2007年)

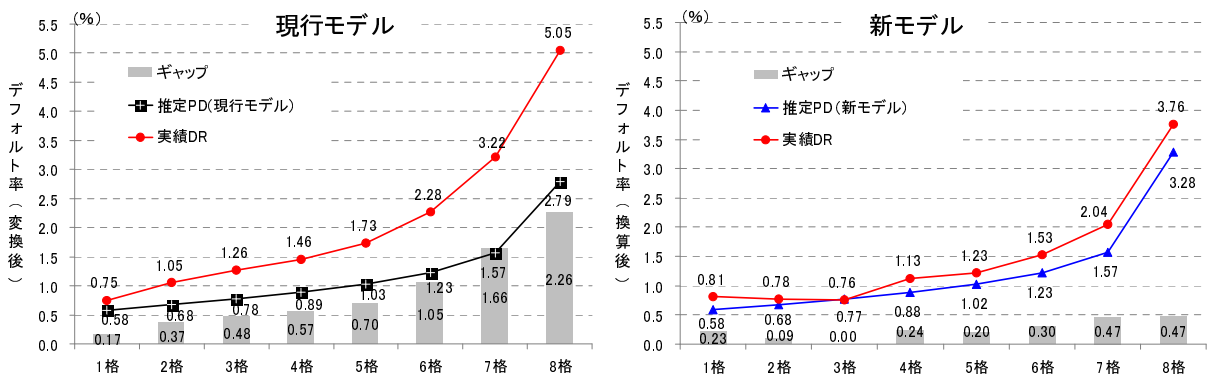


図 6: 格付け別推定PDと実績DRの推移 (その2: IN: 2005～2006年, OUT: 2008年)

2005年～2006年のデータをインサンプルにしてモデルを構築し，2007年と2008年のデータでアウト・オブ・サンプルテストした結果をそれぞれ図5，図6に示す．インサンプル期

間の景気は比較的安定的でマクロファクターの感応度は高くないが、それでもアウト・オブ・サンプルテストにおいては現行モデルに比べて実績 DR とのギャップがかなり縮小されていることが確認できる。図 6 においては 1 格においてギャップがやや拡大している。インサンプル期間とアウト・オブ・サンプル期間に 1 年の空白があることが影響していると思われる。もっとも、水準は低いので実務での影響はほとんどないと考えられる。ギャップを示す棒グラフをみると、現行モデルに比べて新モデルの方のギャップが小さいことは明らかである。8 格を例にとると、図 5 では 1.92%ポイントから 0.81%ポイントに、図 6 では 2.26%ポイントから 0.47%ポイントにギャップが縮小している。

次に 2005 年～2007 年をインサンプルにしてモデルを構築し、2008 年のデータでアウト・オブ・サンプルテストした結果を図 7 に示す。リーマンショック直前のデータをインサンプルに含んでいることからマクロファクターの感応度は図 5 と図 6 に比べて高い。そのため、新モデルの推定 PD の曲線の方が実績 DR の曲線にフィットしている様子が確認できる。ギャップを示す棒グラフをみると、図 7 ではギャップがほぼ解消していることが確認できる。このように格付け別にみると、推定 PD の一致精度が改善されている様子がよくわかる。

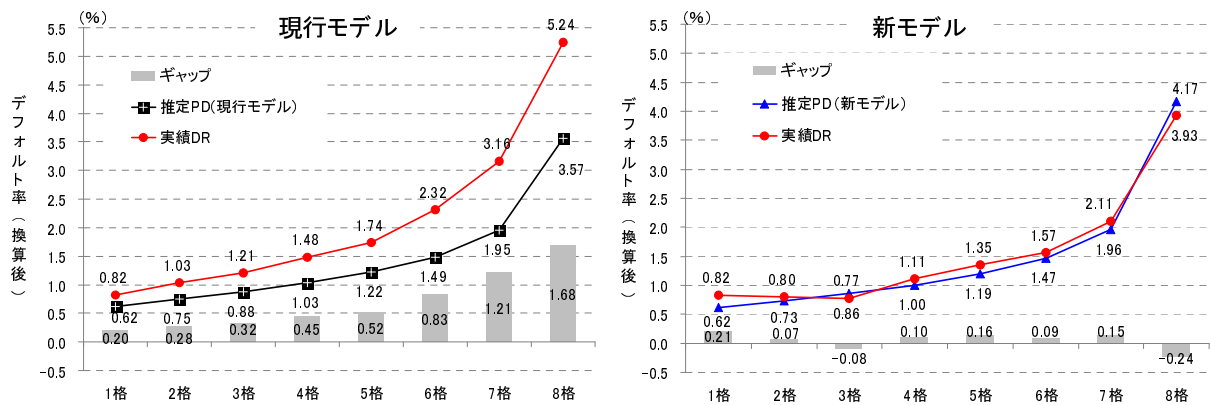


図 7: 格付け別推定 PD と実績 DR の推移 (その 3: IN : 2005～2007 年, OUT : 2008 年)

3.3.2. 業歴別推定 PD の一致精度

業歴別に推定 PD の一致精度を検証したグラフを図 8～図 10 に示す。業歴は 10 年間隔で区分している。左図は現行モデルの推定 PD、新モデルの推定 PD および実績 DR を業歴区分ごとに示した折れ線グラフである。右図は現行モデルと新モデルについて実績 DR とのギャップをそれぞれ折れ線グラフで示し、ギャップの縮小幅 (新モデルのギャップから現行モデルのギャップを引いた値) を棒グラフで示したものである。

2005 年～2006 年のデータをインサンプルにしてモデルを構築し、2007 年、2008 年のデータでそれぞれアウト・オブ・サンプルテストした結果である図 8、図 9 の左図をみると、どの業歴区間においても新モデルの方が実績 DR に近い値になっていることが確認できる。

図 8 と図 9 の右図でギャップの縮小幅を確認すると、現行モデルでは業歴の浅い区間のギャップが比較的大きくなっている。しかし、新モデルではどの業歴においても 0.6%ポイント以内に収まり、業歴による差は解消されている。図 10 は 2005 年～2007 年のデータをインサンプルにしてモデルを構築し、2008 年のデータでアウト・オブ・サンプルテストした結果である。左図をみると推定 PD のグラフと実績 DR のグラフがほぼ重なっている。右図をみるとどの業歴の層でもギャップは 0.3%ポイント以内に収まっており、ギャップの縮小

が明らかである．以上のように業歴別にみても推定PDの一致精度が確認できる．

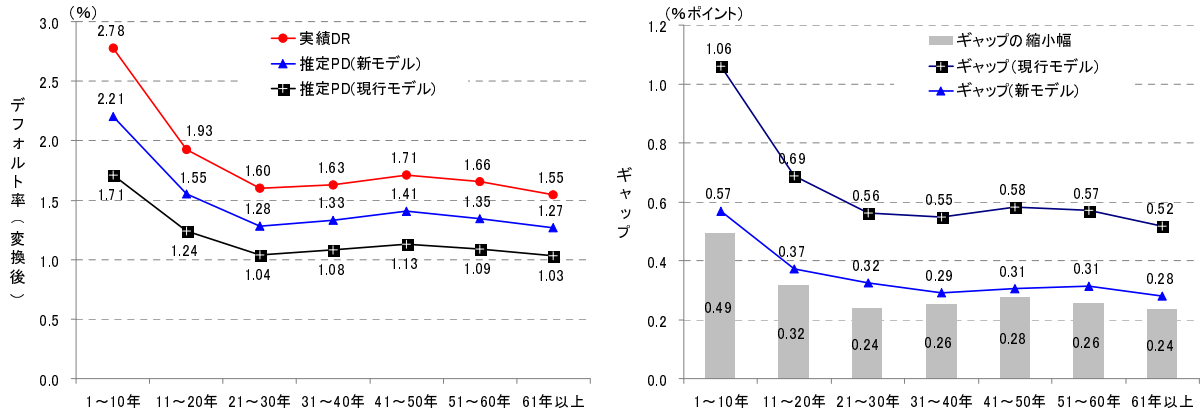


図 8: 業歴別推定PDと実績DRおよびギャップ (その1: IN: 2005~2006年, OUT: 2007年)

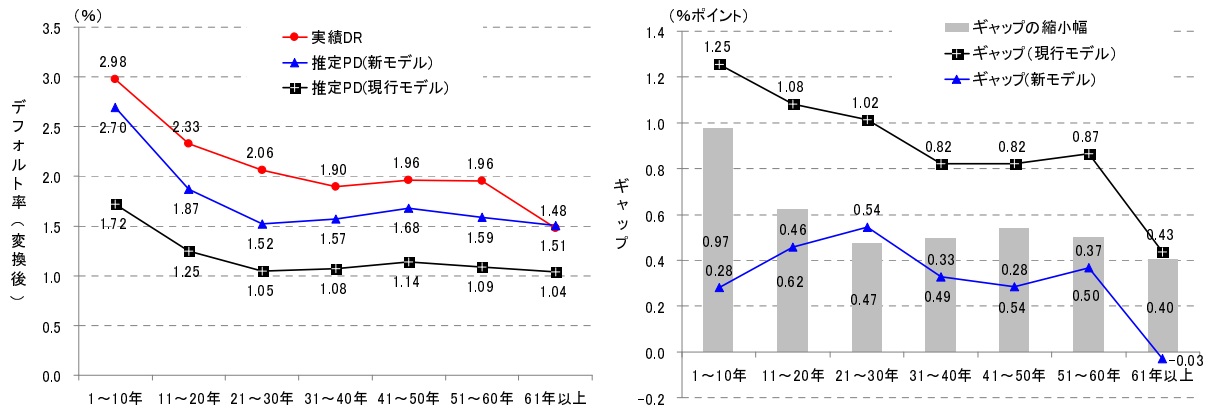


図 9: 業歴別推定PDと実績DRおよびギャップ (その2: IN: 2005~2006年, OUT: 2008年)

3.3.3. 業種別推定PDの一致精度

業種別に検証した結果を図11, 図12に示す．サービス業と卸・小売業のギャップをみても、図11の2005年から2006年のデータをインサンプルにしたモデルではいずれも0.4%ポイント以内となっている．図12の2005年から2007年をインサンプルにしたモデルではそれぞれ-0.10%ポイント, -0.37%ポイントと逆に過大評価となっているが, 0.4%ポイント以内とギャップはおおむね解消されている．

景気の影響を受けやすいと思われる建設業と製造業のギャップをみても、現行モデルではかなり大きいことがわかる．しかし、新モデルにおいては最大で0.7%ポイント残るものの、図11では最大で0.87%ポイント縮小するなど大幅な改善がみられる．図12ではそれぞれ0.22%ポイント, 0.23%ポイントとギャップはおおむね解消されている．

3.4. 業種別モデルの検討

3.3.3項において、業種別の一致精度の改善効果を検証した結果、建設業と製造業はサービス業や卸・小売業に比べて景気の影響を受けやすいことがわかった．景気の影響は業種によって異なる可能性がある．そこで本節では、(3.1)式のように業種別前月デフォルト率を

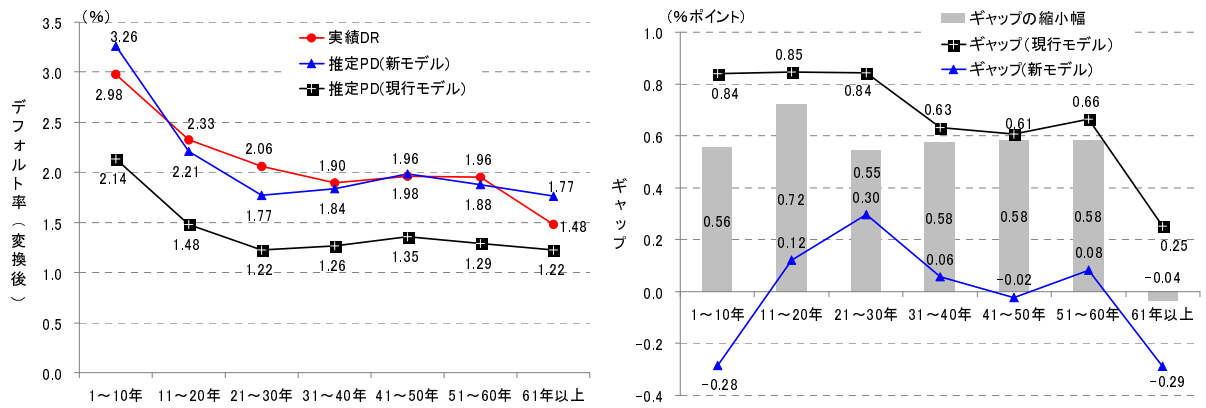


図 10: 業歴別推定 PD と実績 DR およびギャップ (その 3: IN: 2005 ~ 2007 年, OUT: 2008 年)

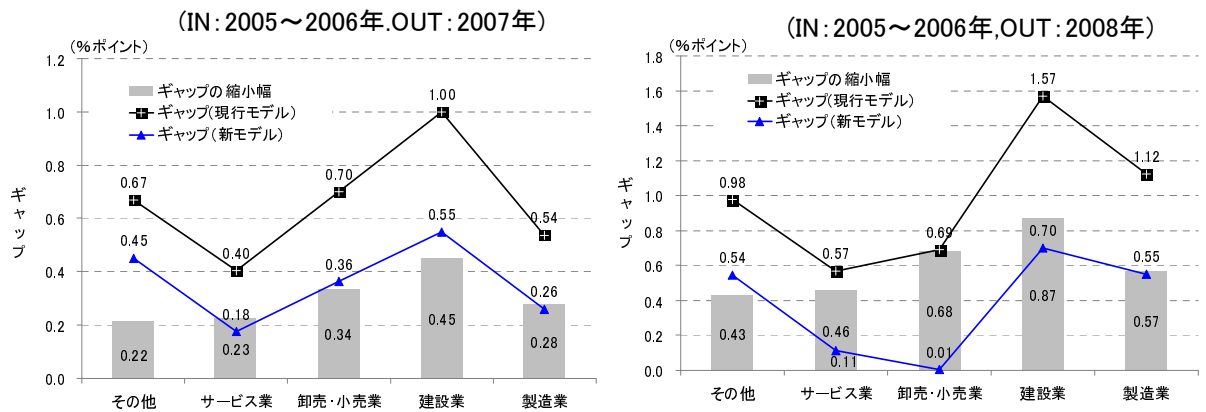


図 11: 業種別推定 PD と実績 DR のギャップ (その 1: IN: 2005 ~ 2006 年)

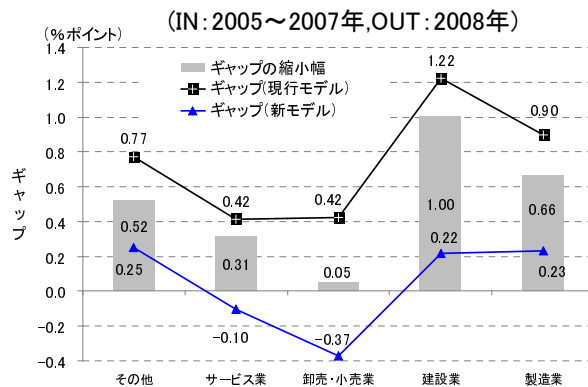


図 12: 業種別推定 PD と実績 DR のギャップ (その 2: IN: 2005 ~ 2007 年, OUT: 2008 年)

用いたモデルを構築し、その有効性について検証する。ここで h_{t-1}^s は融資時点 t の1カ月前の業種 s のデフォルト率、 γ^s は前月デフォルト率のパラメータである。

$$z_i = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k + \sum_{s \in S} \mathbf{1}_i^s \gamma^s h_{t-1}^s \quad (3.1)$$

ここで、 S は業種 s の集合を表す。また、 $\mathbf{1}_i^s$ は企業 i が業種 s に含まれていれば1、含まれていなければ0となる指示関数である。

表7に業種別モデルのパラメータを示す。2005年から2006年のデータをインサンプルにしてステップワイズにパラメータを推定すると、業種別の前月デフォルト率は選択されなかった。一方、2005年から2007年のデータをインサンプルにすると、業種別の前月デフォルト率が有意になったので、その結果のみ表7に示している。AR値は50.71%である。2007年ごろから景気の悪化が始まり実績DRに動きが出たため、業種特性が表れた可能性がある。

表7: 業種別新モデルの推定結果

業種	変数	回帰係数	p 値	変数	回帰係数	p 値	
全業種共通	財務比率 1	-0.0800	0.0000	業歴 (1 乗)	1.6970	0.0000	
	財務比率 2	-0.2122	0.0000	業歴 (2 乗)	-2.9035	0.0000	
	財務比率 3	0.1794	0.0000	業歴 (3 乗)	1.4741	0.0000	
	財務比率 4	-0.0827	0.0000	前月 DR			
	財務比率 5	0.0333	0.0023		製造業	-0.3574	0.0000
	財務比率 6	-0.1315	0.0000		建設業	-0.4348	0.0000
	財務比率 7	0.0815	0.0000		卸売・小売業	-0.5034	0.0000
	財務比率 8	-0.0732	0.0000		サービス業	-0.3833	0.0000
	財務比率 9	-0.1441	0.0159		その他	-0.4187	0.0000
業種 1	財務比率 10	-0.0475	0.0000				
業種 2	財務比率 11	-0.0704	0.0000				
業種 3	財務比率 12	-0.0304	0.0000				
業種 4	財務比率 13	—	—				
業種 5	財務比率 14	—	—				
業種 6	財務比率 15	0.0907	0.0000				

業種別新モデルの推定PDと実績DRとのギャップを図13に示す。図12の新モデルと比較して、景気の影響を受けやすい建設業と製造業のギャップは縮小し、ほとんど解消した。卸・小売業のギャップも同様に縮小したが、サービス業は逆に拡大した。その他業種もギャップが拡大しており、安定した結果が得られなかった。

以上の結果を総合的に勘案し、本研究では業種別の前月デフォルト率は採用せずに全体の前月デフォルト率を採用する。

4. 推定PDの一致精度の改善をもたらす収益効果

推定PDの一致精度の改善はどの程度ローンビジネスの収益に効果を与えるのであろうか。ローンビジネスの収益はモデルの精度だけではなく、営業や審査、債権管理などさまざまな

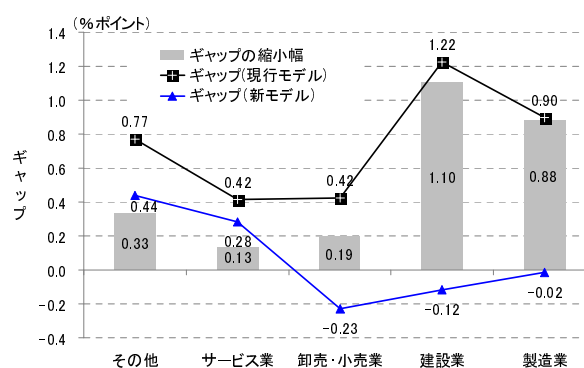


図 13: 業種別推定 PD と実績 DR およびギャップ (IN: 2005~2007 年, OUT: 2008 年)

要素によって決定される。それを踏まえたうえで、本節では、アウト・オブ・サンプルを用いた収益分析を行うことで、推定 PD の一致精度の改善をもたらす収益の改善効果を確認する。推定 PD が過小評価されると、信用コストは少なく見積もられることによって貸出金利は低く設定されることになり、金融機関は金利収入で予想損失額 (EL: Expected Loss) をカバーできなくなる。モデルの改善によって過小評価が修正できれば目標収益が確保できるようになるので、その差額をモデル精度の向上による貢献 (推定 PD の一致精度の改善効果) と考える。

具体的には、2005 年から 2007 年のデータをインサンプルとした現行モデルと新モデルを使用し、2008 年のデータをアウト・オブ・サンプルとした収益分析を行う。目標収益を 0 (収支均衡) と置いた場合、現行モデルの推定 PD を使って金利を設定すれば、個別企業の信用リスクが過小評価されるので収益はマイナスになることが予想される。一方、新モデルの推定 PD を使えば、信用リスクに応じた金利が設定できるので、収益は 0 近傍になると推察される。したがって、現行モデルでアウト・オブ・サンプルを用いた収益分析のマイナス金額と新モデルを使用したときの収益 (0 と予想) の差額が推定 PD の一致精度の改善による収益効果と考えられる。以上の想定のもと、以下の前提および方法でアウト・オブ・サンプルを用いた収益分析を行う。

- ① 1 社あたりの融資金額 (EAD) は一律 1,000 万円, 1 年後に元利一括返済とする。
- ② 貸出金利は一般に調達コスト, 経費, 信用コスト, 信用リスク, 目標収益などを勘案して決定される。ただし, ここでは調達コストと経費は一定, 信用リスクは織り込まず, 目標収益は 0 と置く。したがって, 信用コストの変動のみが収益に影響を与える。
- ③ 収益 I は (4.1) 式によって算出する。 EAD_i は企業 i への融資金額, r_i は企業 i に対する貸出から得られる収益率で, 前提より企業 i の推定 PD に等しい。

$$I = \sum_{i=1}^n EAD_i \times r_i \times (1 - \mathbf{1}_i) - \sum_{i=1}^n EAD_i \times \mathbf{1}_i \quad (4.1)$$

ここで, $\mathbf{1}_i$ は企業 i がデフォルトした場合は 1, デフォルトしなかった場合は 0 となる指示関数である。ただし, アウト・オブ・サンプルを用いた収益分析では, 前提として EAD_i を一律にしているので, i がとれて,

$$I = EAD \left\{ \sum_{i=1}^n r_i \times (1 - \mathbf{1}_i) - D \right\} \quad (4.2)$$

となる．ここで， $D = \sum_{i=1}^n 1_i$ であるので， D はデフォルト数を表す．

- ④ 2005 年から 2007 年のデータをインサンプルにしてモデルを構築し，2008 年のデータでアウト・オブ・サンプルを用いた収益分析を行う．

アウト・オブ・サンプルを用いた収益分析のサンプル数と格付け別の推定 PD および実績 DR を表 8 に示す．インサンプルの企業を推定 PD の低い順番に並べ，8 等分したうえで，2008 年のデータを使って格付けしている．現行モデルはどの格付けにおいても実績 DR が推定 PD を上回っており，過小評価が生じていることがわかる．一方，新モデルは景気の悪化を反映して推定 PD の高い企業が増えるため，6～8 格の企業割合が増加している．

表 8: アウト・オブ・サンプルを用いた収益分析の概要 (IN: 2005～2007 年, OUT: 2008 年)

格付	現行モデル				新モデル			
	n	推定 PD	実績 DR	差	n	推定 PD	実績 DR	差
1 格	16,564	0.62%	0.82%	-0.20%	4,646	0.62%	0.82%	-0.21%
2 格	16,409	0.75%	1.03%	-0.28%	7,571	0.73%	0.80%	-0.07%
3 格	17,118	0.88%	1.21%	-0.32%	10,125	0.86%	0.77%	0.08%
4 格	17,466	1.03%	1.48%	-0.45%	12,939	1.00%	1.11%	-0.10%
5 格	18,083	1.22%	1.74%	-0.52%	16,247	1.19%	1.35%	-0.16%
6 格	18,673	1.49%	2.32%	-0.83%	20,507	1.47%	1.57%	-0.09%
7 格	18,997	1.95%	3.16%	-1.21%	27,143	1.96%	2.11%	-0.15%
8 格	20,103	3.57%	5.24%	-1.68%	44,235	4.17%	3.93%	0.24%
全体	143,413	1.49%	2.21%	-0.72%	143,413	2.21%	2.21%	0.00%

差=推定 PD- 実績 DR

アウト・オブ・サンプルを用いた収益分析の結果を表 9 に示す．現行モデルでは 14 万 3 千件，約 1 兆 4 千億円のローンポートフォリオに対して，収益は約 108 億円のマイナスとなっている．一方，新モデルではマイナスがほぼ解消されて約 8 億円のマイナスに縮小しており，収益が約 100 億円改善していることがわかる．金利水準によってその効果は異なることを踏まえたうえであえて収益率に換算すると，新モデルにすることで 0.7% 程度の収益改善に貢献するという計算になる．前提を置いていることに注意を要するが，アウト・オブ・サンプルを用いた収益分析によって推定 PD の一致精度の向上は収益改善につながる可能性が高いことを確認できた．

5. おわりに

本研究ではマクロファクターを導入した統計モデルを構築し，推定 PD の一致精度の改善効果を検証した．分析にあたっては大規模で多数に分散された約 54 万件に及ぶデータを用いて，個別企業固有の信用リスクを可能な限り排除した．さらに，マクロファクターとなる具体的な経済関連指標を特定するには時系列データの蓄積が不十分という課題があるが，豊富なデータから月次デフォルト率を算出することによって対応した．

株価や原油先物価格指数，金利，恐怖指数など，さまざまな月次の経済関連指標を用いて分析した結果，前月デフォルト率が最も有効であることがわかった．前月デフォルト率を説明変数に加えた新モデルを構築すると，現行モデルに比べて推定 PD と実績 DR のギャップ

表 9: シミュレーションの結果 (IN : 2005 ~ 2007 年 , OUT : 2008 年)

(単位 : 百万円)

	現行モデル		新モデル		収益改善 効果 (B-A)
	<i>n</i>	収益 I(A)	<i>n</i>	収益 I(B)	
1 格	16,564	-337	4,646	-97	240
2 格	16,409	-458	7,571	-52	406
3 格	17,118	-562	10,125	83	644
4 格	17,466	-794	12,939	-138	655
5 格	18,083	-967	16,247	-264	703
6 格	18,673	-1,586	20,507	-216	1,370
7 格	18,997	-2,377	27,143	-467	1,909
8 格	20,103	-3,716	44,235	353	4,069
全体	143,413	-10,796	143,413	-800	9,996

が最大で 0.72%ポイント縮小した。全体の一致精度だけではなく、格付け別、業歴別、業種別の切り口でも一致精度が改善された。ここで業種によって景気感応度にやや違いがみられたため、業種別モデルを構築して検証したが、サンプルによって有意にならないケースがあるうえ、有意になったケースでも効果は限定的であった。このようにカテゴリー別の分析によって、新モデルの頑健性が確認できた。

また、新モデルの導入によって金融機関の収益がどの程度改善するのかをアウト・オブ・サンプルを用いた収益分析で確認した。その結果、モデルの推定 PD を用いて貸出金利を設定している金融機関の場合、マクロファクターを考慮していない現行モデルを用いると、景気変動が収益にマイナスの影響を与える可能性が高いことがわかった。しかし、新モデルを活用すれば、景気変動に応じて適切な金利設定が可能になる。さらに、景気が悪化しているときはデフォルト率も高くなり、推定 PD が高く算出されるため人的審査を慎重に行うケースが増え、デフォルトを抑制する効果が期待できる。一方、景気が上昇しているときは推定 PD が低くなり、効率的に人的審査を行うケースが増えるので審査コストの低減効果が期待できる。アウト・オブ・サンプルを用いた収益分析によって、新モデルを活用することによる間接的な効果を含めて収益の改善効果を確認することができた。

今後の課題として以下の 2 点について研究を進める予定である。

- (1) データ不足を補うために月次データを使用した。月次で公表されている経済関連指標は限られている。GDP をはじめ、年次や四半期で提供されている代表的な経済指標を用いた検証を試みる。
- (2) マクロファクターの感応度は一定ではない。マクロファクターを導入した統計モデルを構築する際には、時間とともに景気状況が変わり、それによってマクロファクターの感応度も変わることを勘案したモデルを検討する。

モデルにマクロファクターをどのように勘案すべきか悩んでいる金融機関は少なくない。本研究によって、マクロファクターとして前月デフォルト率を導入した信用スコアリングモデルを構築することによって推定 PD の一致精度が改善されることが確認できた。本研究が公庫のみならず、他の金融機関の参考になれば幸いである。

参考文献

- [1] A. Bhattacharjee, C. Higson, S. Holly, and P. Kattuman: Business failure in UK and US quoted firms; Impact of macroeconomic instability and the role of legal institution. *Cambridge Working Papers in Economics* (2004).
- [2] D. Bonfim: Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Financial Stability Report 2006, Banco de Portugal* (2006).
- [3] 中小企業庁: 2002年版中小企業白書 (2002).
- [4] CRD協会: CRDモデル3(法人)の推計デフォルト確率水準の調整に関する評価報告書—概要版— (2009).
http://www.crd-office.net/CRD/img/model3-a_houkoku-NEW.pdf
- [5] CRD協会: CRDモデル3・モデル4(BSモデル)・モデル3-aの定期検証に関する評価報告書—概要版— (2010).
http://www.crd-office.net/CRD/img/model343-a_houkoku.pdf
- [6] CRD協会: CRDモデル3・モデル3-a・モデル4の品質に係る定期検証に関する評価報告書—概要版— (2011).
http://www.crd-office.net/CRD/img/model33-a4_houkoku2011.pdf
- [7] S. Figlewski, H. Frydman, and W. Liang: Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *Stern Finance Working Paper No. FIN-06-007* (2006).
- [8] M. Gordy: A comparative anatomy of credit risk models, *Journal of Banking and Finance*. **24-1&2** (2002), 119–149.
- [9] M. Gordy and E. Heitfield: Estimating default correlations from short panels of credit rating performance data. *Working Paper, Federal Reserve Board* (2002).
- [10] 枇々木規雄, 尾木研三, 戸城正浩: 小企業向けスコアリングモデルにおける業歴の有効性. 津田博史, 中妻照雄, 山田雄二 (編集): 定量的信用リスク評価とその応用 (朝倉書店, 2010), 83–116.
- [11] 今井健太郎, 清水信宏: 中小・中堅企業の経済指標を用いることで株式の予測可能性は高まるか. 日本金融・証券計量・工学学会 2010年夏季大会予稿集 (2010), 227–237.
- [12] 金子拓也, 中川秀敏: 信用ポートフォリオのリスク計量: 金利変化見通しと個別企業価値変動を考慮したトップダウンアプローチ. *金融研究*, **29-3** (2010), 19–44.
- [13] G.S. マダラ (和合肇 訳): 計量経済分析の方法 (シーエーピー出版, 1996).
- [14] 森平爽一郎: 信用リスクモデリング—測定と管理— (朝倉書店, 2009).
- [15] 森平爽一郎, 岡崎貴治: マクロファクターを考慮したデフォルト確率の期間構造推定. 早稲田大学大学院ファイナンス総合研究所ワーキングペーパーシリーズ (2009).
- [16] 大橋亨: 企業倒産とマクロ経済要因—企業倒産についての重回帰モデルの構築—. 現代社会文化研究 No.28, 新潟大学大学院現代社会文化研究科 (2003).
- [17] D. Simons and F. Rolwes: Macroeconomic default modeling and stress testing. *International Journal of Central Banking*, **5-3** (2009), 177–204.
- [18] 白田佳子: 企業倒産予知モデル (中央経済社, 2003).

- [19] P. A. Sommar and H. Shahnazarian: Interdependencies between expected default frequency and the macro economy. *International Journal of Central Banking*, 5-3 (2009), 83–110.
- [20] 柳澤健太郎, 下田啓, 岡田絵理, 清水信宏, 野口雅之: RDB データベースにおける信用リスクモデルの説明力の年度間推移に関する分析. 日本金融・証券計量・工学学会 2007 年夏季大会予稿集 (2007), 249–263.
- [21] 山下智志, 三浦翔: 信用リスクモデルの予測精度 —AR 値と評価指標— (朝倉書店, 2011).

付録

A. モーメント法による資産相関の推定方法

森平 [14] (pp.167–168) を参考にして, モーメント法による資産相関の推定方法を具体的に示す. 企業 i の資産成長率の変動を $\tilde{Y}_i \sim N(0, 1)$, マクロファクターを $\tilde{X} \sim N(0, 1)$, 企業間の相関を b_i^2 , 固有ファクターを $\tilde{\xi}_i \sim N(0, 1)$ とすると, 1ファクター・マートン型モデルは

$$\tilde{Y}_i = b_i \tilde{X} + \sqrt{1 - b_i^2} \tilde{\xi}_i \quad (\text{A.1})$$

と記述される. デフォルトは \tilde{Y}_i がある閾値 Q_i を下回ったときに発生し, そのデフォルト確率 $\Pr(\tilde{Y}_i < Q_i)$ は長期的な平均デフォルト確率 \overline{PD}_i と等しいと仮定する. すなわち,

$$\Pr(\tilde{Y}_i < Q_i) = \overline{PD}_i \quad (\text{A.2})$$

である. 前述の仮定の下で, PD のヒストリカルデータを基にしたモーメント法による b_i^2 の具体的な推計方法を示す.

$PD_i(\tilde{X})$ をマクロファクター \tilde{X} の確率変数とする. \tilde{X} をある値 x に固定した場合のデフォルト確率は (A.3) 式のように計算できる.

$$PD_i(x) = \Pr(\tilde{Y}_i < Q_i | \tilde{X} = x) = \Phi \left(\frac{Q_i - b_i x}{\sqrt{1 - b_i^2}} \right) \quad (\text{A.3})$$

n 個のヒストリカルデータを用いてパラメータを推定する. 具体的には, t 期の実績 DR (PD_{it}) は t 期のマクロファクターの実現値 x_t に対応して決まると考え, $PD_i(\tilde{X})$ の平均や分散が実績 DR の平均や分散と等しくなるようにモデルのパラメータの Q_i と b_i^2 を求める.

長期的な平均デフォルト確率 \overline{PD}_i を各期の実績 DR の平均とすると,

$$\overline{PD}_i = E[PD_i(\tilde{X})] = \frac{1}{n} \sum_{t=1}^n PD_{it} \quad (\text{A.4})$$

となる. したがって, (A.2) 式より,

$$Q_i = \Phi^{-1}(\overline{PD}_i) \quad (\text{A.5})$$

を得る. 一方, 分散は

$$V[PD_i(\tilde{X})] = E[PD_i(\tilde{X})^2] - \overline{PD}_i^2 = \frac{1}{n-1} \sum_{t=1}^n (PD_{it} - \overline{PD}_i)^2 \quad (\text{A.6})$$

より (A.7) 式が得られる .

$$\int_{-\infty}^{\infty} \Phi \left(\frac{Q_i - b_i x}{\sqrt{1 - b_i^2}} \right)^2 \phi(x) dx - \overline{PD}_i^2 = \frac{1}{n-1} \sum_{t=1}^n (PD_{it} - \overline{PD}_i)^2 \quad (\text{A.7})$$

ここで, $\phi(\cdot)$ は正規分布の確率密度関数を表す . (A.7) 式で未知なのは b_i^2 だけであるので , 適当な数値計算手法を用いれば b_i^2 を求めることができる .

枇々木 規雄

慶應義塾大学 理工学部 管理工学科

〒 223-8522 横浜市港北区日吉 3-14-1

E-mail: hibiki@ae.keio.ac.jp

ABSTRACT

ESTIMATING THE PROBABILITY OF DEFAULT IN THE CREDIT
SCORING MODEL WITH MACROECONOMIC VARIABLES

Norio Hibiki Kenzo Ogi Masahiro Toshiro
Keio University Japan Finance Corporation

Probability of default (PD) of a small company is estimated by the credit scoring model which mainly includes financial indices. Default is affected by not only a specific factor but also common factors. It is desirable to include the macroeconomic factors as explanatory variables in order to improve the accuracy of the estimated PDs. However, we have a serious problem that there are not enough time series data of default to determine the macroeconomic indices by the regression model. Recently, we begin to recognize a strong need to model the credit scoring with macroeconomic variables because the actual default rates (DRs) are higher than the estimated PDs by the serious downturn in economy from about 2007. In this paper, we determine the macroeconomic indices by using about 540,000 of loan data in Micro Business and Individual Unit of Japan Finance Corporation, and compensating for the lack of the time series data of macroeconomic factors. As a result of the analysis, we find that the previous default rate in a month is significant. We improve the accuracy of the estimated PDs by using the modified credit scoring model with the previous default rate in a month. The difference between the estimated PDs and the actual DRs can be reduced at a maximum of 0.72%.