

# 個人企業向けクレジットスコアリングモデルにおける業歴の有効性

尾木 研三<sup>1</sup>・内海 裕一<sup>1</sup>・枇々木 規雄<sup>2</sup>

(受付 2018 年 9 月 9 日；改訂 2019 年 3 月 16 日；採択 3 月 25 日)

## 要 旨

日本の多くの銀行は、融資審査業務の一部を自動化して審査コストを削減するために、クレジットスコアリングモデルを用いている。モデルにはさまざまなタイプがあるが、従業員数 20 人以下の小企業の信用力評価は、財務指標とデフォルトとの相関を利用したロジスティック回帰モデルの利用が一般的である。小企業には法人企業だけでなく個人企業も含まれるが、個人企業向けモデルの精度は期待するほど高くはない。主な理由は、個人企業は財務諸表を作成する義務がないため、資産や負債といった B/S に関する情報が不足しているからである。精度を上げるには、資産や負債に関する客観的な説明変数の使用が有効と思われるが、個人企業の大半は、事業主世帯の私的資産と事業の会計が混同している場合が多く、正確なデータを得ることが難しい。

尾木 他 (2016) は、小規模な法人企業向けモデルに業歴を用いることを提案し、業歴が重要な変数であることを示した。その理由は次の 3 つである。第 1 に、小企業の評価では、経営者の私的資産を企業の資産に追加することが効果的であるが、そのような情報は容易かつ正確に入手できない。第 2 に、小企業の実績デフォルト率が、業歴が長いほど低いのは、年々蓄積される経営者の私的資産が経営を支えるからと考えられる。第 3 に、経験豊富な銀行員は、融資審査で業歴を債務者の信用リスク評価の重要な要素に位置付けている。

本論文では、個人企業向けモデルにおいても業歴を用いることを提案する。そして、2007 年から 2014 年の間に日本政策金融公庫が融資した 68 万以上の個人企業のデータセットを利用して、業歴とデフォルト率の相関関係を分析する。本論文では、測定値として一般的な AR 値でモデルを評価する。分析の結果、業歴別デフォルト率はワイブル分布の密度関数もしくは「0-3 年」「4-25 年」「26 年以上」の 3 つの区分線形関数で定式化できることがわかった。いずれの関数を用いても、AR 値を約 9% 改善させることができ、実務的観点からモデルが有効であることを確認した。

キーワード：クレジットスコアリングモデル、クレジットリスク、個人企業、業歴、ロジスティック回帰モデル、ワイブル分布。

## 1. はじめに

金融機関は信用リスク管理だけではなく、融資審査業務の一部を自動化して審査コストを削

---

<sup>1</sup> 株式会社日本政策金融公庫 国民生活事業本部：〒100-0004 千代田区大手町 1-9-4

<sup>2</sup> 慶應義塾大学 理工学部：〒223-8522 横浜市港北区日吉 3-14-1

減するために、デフォルト確率(Probability of Default: PD)を推定するクレジットスコアリングモデルを使用している。モデルには、企業の財務諸表データを用いて PD を推定する統計モデル、株価を用いて推定する構造型モデル、社債価格を用いて推定する誘導型モデルなど、さまざまなタイプがある。しかし、中小企業のほとんどは株式や社債を発行していないため、中小企業を対象とする場合は、統計モデルを用いることが多い。なかでも金融機関の間では、ロジスティック回帰モデルが普及している。173 の金融機関(2018 年 4 月現在)が加入する一般社団法人 CRD 協会や、22 の銀行等が共同出資して設立した日本リスクデータバンク株式会社もロジスティック回帰モデルを提供している。

日本政策金融公庫国民生活事業本部(以下、公庫という)は、主に従業者数が 20 人未満の小企業に対する融資を行っている。公庫もロジスティック回帰モデルを使用しているが、融資先の半数近くは個人企業である。個人企業とは、法人としての登記をしていない個人が経営する経営体のことであり、税務申告においては、青色申告もしくは白色申告している個人事業者のことである<sup>1)</sup>(なお、本論文では個人企業と法人企業の経営者の混同を避けるため、個人企業に対しては「事業主」、法人企業に対しては「経営者」と呼ぶことにする)。

個人企業は、法人が作成するような決算書はなく、その代わりとなる税務申告書においても、貸借対照表(B/S)を作成しているケースが少ないうえ、資産や負債に関する情報の開示にも消極的である。したがって、融資審査にあたっては、金融機関同士が情報を共有するデータベースから負債に関する情報を取得し、審査担当者による事業主からのヒアリングや提出書類、帳簿、通帳などから資産に関する情報を把握して、簡便的に B/S を作成して評価している。このように、個人企業は審査の手間とコストがかかるうえ、法人企業に比べて融資金額が小さく採算ベースに乗りにくいいため、スコアリングモデルの精度向上による審査業務の一層の効率化は金融機関にとって長年の課題となっている。

先行研究をみると、尾木 他(2016)は小規模な法人企業向けモデルに業歴を用いることを提案し、業歴が重要な変数であることを示した。その理由は次の 3 つである。第 1 に、小企業の評価では、経営者の私的資産を加味することが効果的であるが、そのような情報は容易かつ正確に入手できない。第 2 に、小企業の実績デフォルト率が、業歴が長いほど低いのは、年々蓄積される経営者の私的資産の大きさが影響していると考えられる。第 3 に、経験豊富な銀行員は、業歴を債務者の信用リスク評価の重要な要素として扱っている。

さらに、業歴別デフォルト率は、決算書の数値には表れない小企業のパフォーマンスを示す指標と考えられるが、その一つとして経営者の私的資産の大きさが含まれ、大きなウェイトを占めている可能性があることも示した。具体的には、個人企業の業歴別デフォルト率と事業主の保有する預貯金額あるいは不動産評価額との相関係数がいずれも約 0.8 と高いことを明らかにしている<sup>2)</sup>。分析の目的は法人企業といえども、小企業の場合には経営者の私的資産額が重要なファクターと考えられる一方で、法人経営者の私的資産額のデータが存在しないため、個人企業のデータを用いて検証したとしている。

一連の分析の結果、尾木 他(2016)は、財務指標に加えて、経営者の私的資産を含む小企業のパフォーマンスを示す指標として、業歴別デフォルト率を定式化した 3 次関数を変数として投入すると、法人企業向けモデルの精度を示す AR 値<sup>3)</sup>が改善することを示した。

個人企業の場合、事業資産と事業主の私的資産が混在しており、それを明確に区分することは難しいが、個人企業の業歴別デフォルト率を、非開示となっている事業主の私的資産を含む個人企業のパフォーマンスを示す指標として利用できれば、それを定式化した業歴関数をモデルの変数に含めることによって、法人企業と同様に個人企業向けモデルの精度を改善できる可能性があると考えた。そこで、本研究では尾木 他(2016)の研究を参考にし、個人企業の業歴別デフォルト率を定式化した業歴関数の有効性について検証することを試みる。分析の結果、

以下の点が明らかになった。

- (1) 個人企業の業歴別デフォルト率の形状は、法人企業とは異なることがわかった。具体的には、法人企業と同様に業歴 25 年頃まではデフォルト率が単調減少するが、業歴 35 年頃からデフォルト率が上昇するという特徴は個人企業ではみられず、業歴 25 年以降は概ね一定であった。
- (2) 業歴関数は法人企業で有効性が確認された 3 次関数よりも、ワイブル分布の密度関数もしくは、業歴が「0-3 年」「4-25 年」「26 年以上」で 3 区分した区分線形関数のいずれかを用いることが有効であり、どちらを用いてもパフォーマンスに大きな差はないことを確認した。
- (3) 業歴関数をモデルの変数に投入すると、医療・福祉業、不動産業、運輸業を除く業種で AR 値が改善し、特に飲食・宿泊業においては約 9% ポイント改善することを確認した。
- (4) 個人企業全体の AR 値の改善幅は 2.3% ポイントから 2.6% ポイントと、法人の 7.6% ポイントに比べて小さいことがわかった。要因として、開示されている事業資産に事業主の私的資産が相応に含まれている（非開示の資産が少ない）、もしくは、非開示の私的資産は事業に投入しない傾向にあるなどの理由により、業歴関数の効果が期待したほど高くならなかったと考えられる。

なお、本研究では、デフォルトの定義は破綻懸念先以下へのランクダウンおよび 3 か月以上延滞先とし、デフォルト率は融資企業数のうち融資した翌年度末までにデフォルトした企業の割合とする<sup>4)</sup>。ただし、廃業に伴い融資残高を一括返済した企業については成功企業と考え、デフォルト企業に含めていない。

本論文の構成は以下の通りである。2 章では、関連する先行研究を概観する。3 章では、個人企業の概要と特徴を述べ、4 章はモデルおよび使用データの概要と業歴別デフォルト率の特徴を確認する。5 章では、個人企業の業歴別デフォルト率の定式化を試みる。6 章では業歴関数のモデル投入によるパフォーマンスを検証し、7 章では結論と今後の課題を述べる。

## 2. 先行研究

デフォルト確率や倒産確率を推計するモデルの研究にはきわめて多くのものがあり、さまざまな手法が提案されている。山下 他 (2003) によれば、これらは、貸出先企業の財務データから判別分析や回帰分析などの統計手法を用いて推定する統計モデルと株価や社債金利などの市場データからオプション理論を用いて推定するオプションアプローチに大別される。この他にも、サポートベクターマシーン (SVM) やニューラルネットワークといった機械学習的な手法を用いて推定する方法がある。

本論文は、個人企業を対象にし、金融機関の間で最も普及しているロジスティック回帰モデルにおいて、その変数として業歴別デフォルト率を定式化した業歴関数の投入により AR 値が向上するかどうかを検証することを目的としている。そこで本章では、2.1 節でロジスティック回帰モデル、2.2 節では個人企業、2.3 節では業歴とデフォルトとの関係に関する先行研究について概観する。

### 2.1 ロジスティック回帰モデルに関する研究

ロジスティック回帰モデルに関する研究は、非常に盛んに行われている。Ohlson (1980) が行ったロジスティック回帰モデルを使って倒産確率を推定する研究に始まり、Altman and Sabato (2007) の中小企業の 1 年間のデフォルト確率を推定する研究など、国外には数多くの研

究がある。国内においても数多くの研究が行われている。柳澤 他 (2007) は、2000 年度から 7 年間の約 35,000 件の RDB(Risk Data Bank) データを用いて、複数のロジスティック回帰モデルを構築し、AR 値の推移について検証している。森平・岡崎 (2009) 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モデルの有効性を提案している。山下 他 (2003) は、スコアリングモデルの評価手法の特徴などについて詳細な分析を行っている。この他、変数選択とデータ量に関する研究 (山下 他, 2003)、財務変数に加えて非財務変数として業歴の有効性を検証した研究 (枇々木 他, 2010)、創業企業の信用リスクを評価するスコアリングモデルに関する研究 (尾木 他, 2017) などもある。

多数の研究に加えて、データの量と質の向上によって、モデルの精度は着実に上がってきている。今やほとんどの金融機関が統計モデルを使って中小企業のクレジットリスクを定量化し、審査やリスク管理などの実務に活用している。最近では、機械学習の手法を用いた人工知能(AI)型スコアリングモデルを使った審査に取り組む金融機関も増えており、実務におけるスコアリングモデルの活用範囲は着実に広がっている。ただ、尾木 他 (2017) を除いて、いずれの研究も個人企業ではなく、法人企業を対象にしている。

## 2.2 個人企業に関する研究

前述のように、個人企業のスコアリングモデルに関する先行研究はほとんど見当たらない。ただ、個人企業に関する調査・研究はいくつか存在する。国民生活金融公庫総合研究所 (2002) は、「自己雇用者(≡ 自営業者)に関する実態調査」を行い、個人企業の事業主と法人企業の経営者の違いについて分析している。上村 (2006) は、個人企業世帯の家計資産について、国勢調査や貯蓄動向調査をもとに分析し、個人企業世帯の家計は複雑・不安定で把握しにくいことを述べている。高川・亀田 (2008) は、事業所・企業統計調査や国勢調査などをもとに個人企業の企業数が減少している理由について、高齢化や規制緩和による法人との競争激化にあることを明らかにした。

Schreiner (2000) は、個人企業の信用評価は、定性情報の依存度が高い一方で、スコアリングは定量情報に依存しているため、貸し手がスコアリングモデルを使って業務の効率化を図ることは容易ではないことを述べている。Kneiding and Kritikos (2011) は、ドイツの個人企業世帯の資金移転について所得と消費に関する調査データを使って分析し、個人企業世帯はビジネスの目的で消費者ローンを使用しており、家計と事業の資金移転が行われている実態を明らかにした。Gweyi (2013) は、個人企業はデフォルトリスクが非常に高く、ケニアの銀行は、マイクロファイナンスの分野に進出することが困難である現状を解説している。以上のように、個人企業向けスコアリングモデルに関する研究はほとんどない。個人企業は家計と事業の会計が混在していて、実態が掴みにくいことがわかる。本研究は、個人企業向けスコアリングモデルの変数として、業歴別デフォルト率を定式化した業歴関数を利用する点に特徴がある。

## 2.3 業歴と倒産およびデフォルトとの関係に関する研究

業歴に関する研究は近年盛んに行われ、さまざまな研究成果が発表されている。まず、業歴と企業のパフォーマンスとの関係を分析した研究を紹介する。Cabral and Mata (2003) は、ポルトガルの製造業のデータを用いて、業歴が長いほど企業規模が大きくなることを示した。Sakai et al. (2010) は、一般社団法人 CRD 協会が保有する 20 万社以上のパネルデータを使用して中小企業の業歴と借入コストとの関係を分析し、業歴が長くなるほどパフォーマンスが上がり、借入コストが下がることを明らかにした。一方で、森川 (2008) は、従業員数 50 人以上の法人企業約 5 千社のデータを用いて同族企業の生産性について分析し、その中で企業年齢と生

産性が有意に負の関係にあるという結果を示した。Loderer and Waelchli (2010)は、米国の上場企業においては、業歴が長いほど組織の硬直化や研究開発活動の低下が進行し、ROA(総資産利益率)や Tobin の  $q$  のパフォーマンスが悪化することを述べている。また、Akben-Selcuk (2016)は、302社のトルコの企業を対象に、業歴と ROA や ROE といった収益性との関係进行分析し、業歴が長くなるほど収益性が低下することを明らかにした。このように、業歴が長いほど企業のパフォーマンスが上がるという研究がある一方で、パフォーマンスが下がるとする研究もある。

次に、業歴とデフォルトもしくは倒産との関係进行分析した研究をいくつか紹介する。清水 (1985)は、企業年齢と倒産について分析し、倒産する企業数だけをみると、企業年齢が3~10年未満の企業が倒産しやすく、10年以上の企業は倒産し難いものの、生存する企業との比較では単純にそうとも言えないことを指摘している。Evans (1987)は米国の中小企業は、業歴が長いほどデフォルトが減ると述べている。一方、中小企業庁の中小企業白書(2002)は、東京商工リサーチのデータを用いて業歴とデフォルトとの関係についてプロビット分析し、業歴はデフォルトを説明する有意な変数にはならないとしている。ただ、白書のなかに記載されている生存企業と倒産企業の業歴別構成比を用いて、その比率(倒産企業のうち業歴  $a$  年の企業の割合  $\div$  生存企業のうち業歴  $a$  年の企業の割合)を計算すると業歴と倒産率との関係は、業歴が長いほど倒産率が低くなるが、この点は指摘されていない。

また、鶴田 (2005)はノンバンクの融資先を対象に、業歴が長いほど経営は安定しているという仮説を検証した。債務超過企業をデフォルト企業とした場合は有意ではないが、インタレストカバレッジレシオが1未満の企業をデフォルト企業とした場合には有意になることを示している。東京商工リサーチの友田 (2008)は業歴別倒産件数構成比の推移を示し、業歴30年以上の老舗企業の構成比が年々上昇し、業歴10年未満の新興企業の構成比が下がっていることを示している。枇々木 他 (2010)は、日本政策金融公庫が保有する約48万件のデータを用いて、小企業の業歴とデフォルトとの関係は3次関数で表現できることを明らかにした。このように、業歴とデフォルトとの間に関係があるかどうかについては、必ずしも明確になっていない。さらに、個人企業の業歴とデフォルトとの関係进行分析した研究は、筆者たちの知る限り存在しない。

### 3. 個人企業の概要と特徴

本章では、個人企業の概要と特徴について、国内の個人企業の実態に関する先行研究をもとに、法人企業との比較を交えながら概観する。

#### 3.1 個人企業の概要

個人企業の一般的なイメージは、商店主や工場主など、夫が事業主で妻が家族従業者という形である。具体的な数値でみてみよう。

平成26年の経済センサス(総務省統計局, 2016)をみると、個人企業の事業所数の割合は、38.4%と全体の約4割を占める。業種構成を表1に示す。卸・小売業の割合が最も高く21.8%となっている。法人企業と比較すると、飲食業の割合が、法人企業が6.1%であるのに対し、個人企業が19.4%と高くなっているほか、理美容業の割合が、法人企業が3.6%であるのに対し、個人企業が14.4%と高いという点が特徴的である。製造業の減少とともに、工場主は減り、現在では商店主、あるいは、飲食店や理美容店の店主が多くなっている。

次に、個人企業の規模について確認する。2002年に実施したものであることには注意を要するが、表2に国民生活金融公庫総合研究所(現:日本政策金融公庫総合研究所)(2002)「自己雇

表 1. 個人企業の業種構成比。数値は事業所数に対する割合。値の小さな業種は除いている。  
したがって、合計は 100% にならない。出典：総務省統計局（2016）「平成 26 年経済  
センサス-基礎調査」。

| 業種                | 個人企業    | 法人企業   |
|-------------------|---------|--------|
| 建設業               | 7.2%    | 10.7%  |
| 製造業               | 6.9%    | 10.0%  |
| 情報通信業             | 0.1%    | 1.9%   |
| 運輸業，郵便業           | 0.9%    | 3.4%   |
| 卸売業，小売業           | 21.8%   | 27.8%  |
| 金融業，保険業           | 0.3%    | 2.4%   |
| 不動産業，物品賃貸業        | 7.0%    | 6.9%   |
| 学術研究，専門・技術サービス業   | 4.9%    | 3.7%   |
| 宿泊業，飲食サービス業       | 20.8%   | 8.4%   |
| ( 飲食店 )           | (19.4%) | (6.1%) |
| 生活関連サービス業，娯楽業     | 15.1%   | 4.9%   |
| ( 理美容業 )          | (14.4%) | (3.2%) |
| 教育，学習支援業          | 4.4%    | 2.2%   |
| 医療，福祉             | 8.2%    | 7.2%   |
| サービス業（他に分類されないもの） | 2.4%    | 8.5%   |

表 2. 個人企業の概要（法人企業との比較）。出典：国民生活金融公庫総合研究所（2002）「自  
己雇用者（≡自営業者）に関する実態調査」（以下，図 1～図 3 も同じ）。

|                | 平均<br>従業員数 | 平均売上高<br>( 年間 ) | 創業者の<br>割合 | 経営者の<br>平均年齢 |
|----------------|------------|-----------------|------------|--------------|
| 個人企業 (n=2,380) | 1.9人       | 2,116万円         | 84.5%      | 53.8歳        |
| 法人企業 (n=1,426) | 36.9人      | 8億6,314万円       | 55.0%      | 56.9歳        |

用者（≡自営業者）に関する実態調査」（以下，公庫調査という）の結果を示す。平均従業員数は本人を含めて 1.9 人，配偶者が家族従業員となっている割合は 57.2% で，夫婦で事業を営んでいる企業が大半である。平均売上高は年間 2,116 万円と，法人企業の 8 億 6,314 万円に比べて規模はかなり小さい。また，経営者または事業主が創業者である割合は，法人企業が 55.0% であるのに対し，個人企業が 84.5% となっており，事業承継が少ないことがわかる。2014 年版中小企業白書（中小企業庁，2014）によると，廃業した企業のうち個人企業が約 9 割を占めている。日本政策金融公庫総合研究所が 2016 年に行った「中小企業の事業承継に関するインターネット調査」（日本政策金融公庫総合研究所，2016）でも廃業予定企業の 68.4% が個人企業であることから，法人企業に比べて一代限りで廃業する企業が多いことがわかる。

### 3.2 個人企業の特徴

3.1 節のとおり，個人企業の平均像は小売店，飲食店，理美容店などを営む経営者と配偶者で構成される年商 2 千万円程度の小規模な企業であることがわかった。本節では，個人企業の主な特徴についてさらに 4 つの視点から確認する。

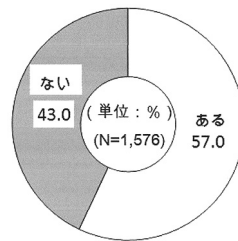


図 1. 個人企業世帯の別収入. 自己雇用者(個人事業主)本人, 配偶者, 同居家族のそれぞれについて, 別収入の有無を尋ね, 集計したものである. 出典: 表 2 に同じ.

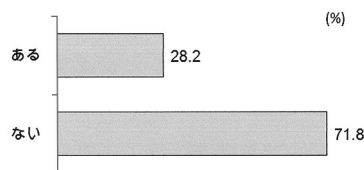


図 2. 個人企業経営者の副業の有無. 出典: 表 2 に同じ.

#### (1) 家族経営で別収入がある

内閣府(2011)「年次経済財政報告」によると, 労働政策研究・研修機構「副業者の就労に関する調査」を用いて, 副業の日数を本業の就業形態や性別などの要因で回帰分析した結果, 個人企業の事業主とその家族従業者は, 正社員に比べて, 1 カ月あたり 3 日程度副業従事日数が多いということがわかった. 公庫調査でも, 図 1 のとおり, 本人も含めた家族に別収入がある割合は 57% となっている.

事業主自身が他に本業を持っていることもある. 図 2 のとおり, 個人企業の事業主自身に別収入がある割合は 28.2% となっている. たとえば, プロゴルファーが経営するゴルフショップや勤務者が経営する不動産賃貸業など, 別収入の方が多い場合もある. 個人企業は規模が小さいため, 事業の赤字を別収入で補てんする企業は少ない.

#### (2) 家計と事業の会計が混在している

上村(2006)は「個人企業(自営業)の家計は事業経営との関係が緊密で, 勤労者の家計に比べて複雑・不安定で把握がしにくく, その内部構造を明らかにした研究は少ない」「全くの井勘定で家族員の生活に必要な費用がその日の家業の売上から支出されてしまうタイプもある」と述べている. 前(1)の特徴も踏まえると, 個人企業は, 別収入があるうえ帳簿類が未整備で, 家計と事業の会計が混在していることがわかる.

#### (3) 業種や経営形態が多様である

平成 23 年の年次経済財政報告に, 個人企業は「日銭を稼いで暮らしていくための「生業」がある一方, イノベーションのけん引役としての「ベンチャービジネス」がある」という記述がある. 一口に個人企業といっても, 大宗を占める商店や飲食店, 理美容店だけでなく, 開業医, 弁護士などの専門職業, 在宅ワークやネット関連ビジネスといったものまで, 多様な業種や形態がある.

#### (4) 法人成りする

個人企業はデフォルトや廃業以外に, 法人化することによっても企業数が減る. 企業規模を維持したまま個人企業として事業を継続する企業も少なくないが, 一般に, 規模が大きくなる

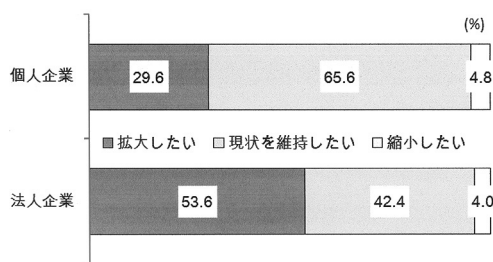


図 3. 事業規模の拡大意欲. 出典: 表 2 に同じ.

と法人化する傾向がある. 図 3 のとおり, 公庫調査によると, 個人企業の拡大志向は法人企業より低いものの, 3 割程度は事業規模の拡大に意欲的であることがわかる.

#### 4. モデル構築手順の概要とデータの特徴

枇々木 他 (2010) や尾木 他 (2016) は, 小規模な法人企業向けモデルにおける業歴の有効性について示した. 本章では, 尾木 他 (2016) の研究をもとに, モデル構築手順およびデータの概要と業歴別デフォルト率の特徴について述べる.

##### 4.1 モデル構築手順の概要

尾木 他 (2016) が示した業歴を考慮したロジスティック回帰モデル (以下, 業歴モデルという) は, 財務指標を用いた一般的なモデル<sup>5)</sup>をベースに, 説明変数として業歴別デフォルト率の形状を定式化した業歴関数 (3 次関数) を加えたものである.

尾木 他 (2016) が示した業歴関数  $\sum_{k=1}^3 \beta_k (g_i)^k$  ( $i = 1, \dots, N$ ) を加えた法人企業向けの業歴モデルを (4.1) 式で示す. ここで  $g_i$  は企業  $i$  の業歴 (61 年以上は 61 年) である.

$$(4.1) \quad p_i = \frac{1}{1 + e^{z_i}}, \quad z_i = \ln \left( \frac{1 - p_i}{p_i} \right) = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij} + \sum_{k=1}^3 \beta_k (g_i)^k \quad (i = 1, \dots, N)$$

尾木 他 (2016) が示した業歴モデルの特徴は, (4.1) 式の右辺第 3 項の業歴関数にある. すなわち, 個人企業の業歴モデル構築のポイントは, 業歴別デフォルト率をどのように定式化するかにある. 本研究におけるモデルは, (4.1) 式のモデルをベースに, 右辺第 3 項の業歴関数にさまざまなタイプの関数の当てはめを行い, 修正  $R^2$  によるフィッティングの良さ, AR 値を使ったモデルの精度を比較検討しながら, モデルを構築する.

##### 4.2 使用データの概要

表 3 に本研究で使用する個人企業データの概要を示す. まず, 年度別の企業数については, 対象となる融資年度は異なるが, 特徴を比較するために, 法人企業データを用いた尾木 他 (2016) の分析データおよび結果も示す<sup>6)</sup>. 本研究で使用するデータは公庫が 2007 年度～2014 年度の 8 年間に融資した個人企業の延べ 684,994 社の年度別データである. また, 尾木 他 (2016) の分析データは, 2004 年度～2011 年度の 8 年間に融資した法人企業, 延べ 1,089,362 社<sup>7)</sup>である.

年度ごとの企業数は重複を許している. 例えば, 業歴  $n$  年の企業が  $x$  年度と  $x + 3$  年度に 1 回ずつ合計 2 回の融資を受けた場合,  $x$  年度は業歴  $n$  年の企業,  $x + 3$  年度は業歴  $n + 3$  年の企業としてカウントしている. 延べサンプルを使用しているのは, 同一の企業でも, 時点が異なれば, 企業の財務内容, 経営者もしくは事業主の資質や私的資産額などが異なるためである.



表 3. 使用データの概要.

(年度別データ数)

| 年 度   | 2007    | 2008    | 2009    | 2010    | 2011    | 2012    | 2013    | 2014    | 合計        |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|-----------|
| 個人企業数 | 88,053  | 93,779  | 99,999  | 95,642  | 83,210  | 77,590  | 74,762  | 71,959  | 684,994   |
| 年 度   | 2004    | 2005    | 2006    | 2007    | 2008    | 2009    | 2010    | 2011    | 合計        |
| 法人企業数 | 133,643 | 127,496 | 123,174 | 130,821 | 140,181 | 162,616 | 147,718 | 123,713 | 1,089,362 |

(業種別データ数)

|      | 製造     | 建設      | 卸小売     | 飲食宿泊   | 医療福祉   | サービス    | 不動産    | 運輸     | その他    |
|------|--------|---------|---------|--------|--------|---------|--------|--------|--------|
| 個人企業 | 44,305 | 111,876 | 145,022 | 85,801 | 43,979 | 123,725 | 47,148 | 14,054 | 35,527 |

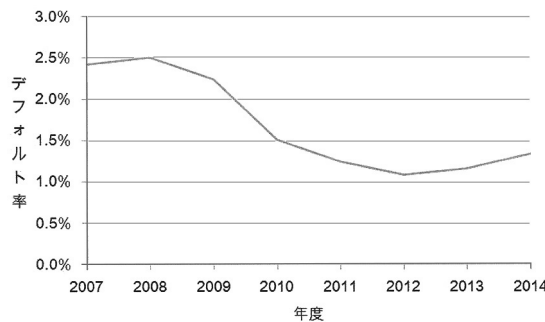


図 4. 年度別デフォルト率.

次に、業種別のデータ数をみると、卸・小売業が約 14 万 5 千件と最も多く、次にサービス業が約 12 万 3 千件、建設業が約 11 万 1 千件と続く。製造業と医療・福祉業は 4 万件台、運輸業は約 1 万 4 千件と業種間の差が大きくなっている。

最後に、図 4 で年度別のデフォルト率を見てみると、リーマンショックがあった 2008 年度前後のデフォルト率が 2% 台と高くなっている。2010 年度以降は、おおむね 1%～1.5% の間で推移している。デフォルトや倒産は、企業個別の要因だけではなく、マクロ経済要因の影響も受けるため、時系列での分析も重要であるが、データの制約もあり、この点については今後の課題と認識している。

#### 4.2.1 業歴別企業数の特徴

業歴別企業数を図 5 に示す。業歴が長くなるほど個人企業数は減少する。業歴 15 年頃から 30 年頃にかけてやや減少のペースが遅くなるものの、業歴 15 年ぐらいいまでと、業歴 30 年頃から 50 年頃にかけて、おおむね一定のペースで減少する。創業時に個人企業であっても、業歴が長くなるにつれて法人化する企業が増えるため、個人企業数は一貫して減少する傾向にある。業歴 15 年頃から 30 年頃にかけては、法人化する企業が一段落するため、減少のペースが緩やかになるが、事業承継期もしくは廃業時期を迎えて業歴 30 年以降は再び減少のペースが上がると考えられる。一方、法人企業は、尾木 他 (2016) が示したとおり、業歴 10 年ぐらいいまで低下したあと、しばらく安定した期間が続くが、業歴 40 年ぐらいいから大きく減少する。なお、業歴が 60 年を超えると企業数が大幅に減少するため、図 5 では示していないが、業歴 100 年以上の個人企業は 3,500 社程度存在する。

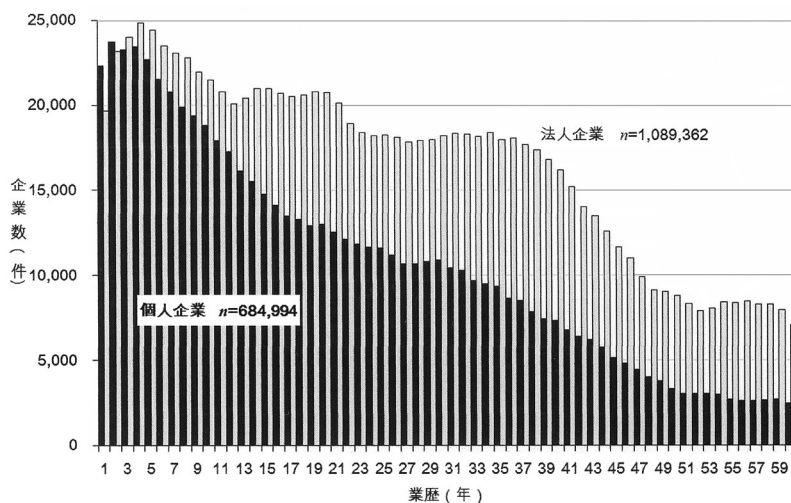


図 5. 業歴別企業数.

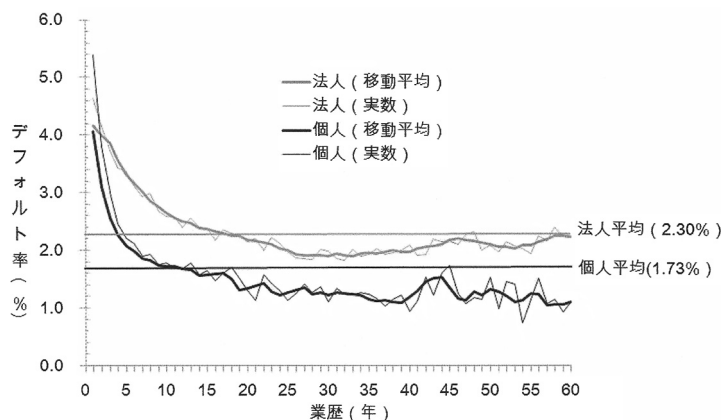


図 6. 業歴別デフォルト率.

#### 4.2.2 業歴別デフォルト率の特徴

尾木 他 (2016)の分析した法人企業は 2004 年度から 2011 年度までのサンプルをそれぞれプールした状態でデフォルト率を計算している. 本研究でも同様の条件で, 2007 年度から 2014 年度までの個人企業のサンプルを使って計算した.

デフォルト率を業歴別に計算した結果を図 6 に示す. デフォルト率の傾向を示すために, 前後 2 年を含む 5 年間 (-2, -1, 0, +1, +2 年)のデフォルト率を平均した移動平均(太線)<sup>8)</sup>も示す. 平均値をみると, 個人企業が 1.73% であるのに対して法人企業は 2.30% と, 個人企業の方が低い. 観測期間が異なる影響を除去するため, 共通する年度(2007~2011 年度)の結果も表 4 に示す. 同じ年度で比べても, 個人企業は法人企業に比べてデフォルト率が低い傾向にある.

業歴とデフォルト率との関係を見ると, 個人企業, 法人企業とも業歴 5 年ぐらいまでの企業のデフォルト率は高く, 業歴 15 年ではほぼ平均まで低下(個人 1.64%, 法人 2.38%), その後も低下を続けて業歴 25 年から 35 年ぐらいまではおおむね横ばいとなる.

表 4. 個人・法人企業別年度別デフォルト率.

| 年度   | 2007  | 2008  | 2009  | 2010  | 2011  |
|------|-------|-------|-------|-------|-------|
| 個人企業 | 2.42% | 2.50% | 2.24% | 1.51% | 1.25% |
| 法人企業 | 3.75% | 3.20% | 2.22% | 1.54% | 1.57% |

35 年頃(個人 1.23%, 法人 1.93%)から法人企業のデフォルト率は徐々に上昇しはじめ, 45 年(2.18%)を過ぎると再び低下する. 一方, 個人企業は 44 年と 45 年に高い時期もあるが, これを除くと 40 年を過ぎてもおおむね横ばいか, やや低下傾向にある. 業歴別デフォルト率の動きにおける個人企業と法人企業との違いについて, 現場感覚や中小企業研究における指摘などと整合的かという視点を交えながら, 業歴区間ごとの特性を詳しくみていくことにする.

#### (1) 業歴 5 年未満

業歴 1 年のデフォルト率は, 個人企業が 5.39%, 法人企業が 4.63% といずれも高い水準にある. 開業後 5 年を過ぎるまではデフォルト率が高い時期がしばらく続くことがわかる. 安田他 (2007) は, 開業したての経営者には資質に欠く者が一定数含まれており, 開業後間もない時期にこうした企業が自然淘汰されるため退出率が高くなると指摘している.

個人企業を法人企業と比較すると, 法人企業は業歴 5 年に 3.37% に低下するのに対し, 個人企業は 2.20% と半分以下の水準まで急激に低下する. 低下のペースは個人企業の方が早いことがわかる. この理由はいくつか考えられる. 一つ目は, 3.2 節(4)のとおり, 事業が拡大すると法人化するが, 拡大期はデフォルトしやすい傾向があるため, 法人成りした後でデフォルトしている企業があると考えられる. 図 5 をみると, 業歴 5 年未満の企業数は, 個人企業が横ばいであるのに対し, 法人企業は 5 千社程度増加しており, 個人企業が法人成りしている可能性がある. 二つ目に, 個人企業は規模が小さく小回りが利くため, 事業が一旦軌道に乗るとデフォルトしにくい傾向にあるということである. 3.1 節のとおり, 飲食店や理美容店のような規模が小さく現金商売の業種が多いうえ, 3.2 節(1)(2)のとおり, 事業収入の減少を家族の別収入で補ったり, 経営者自らが副業を行ったりするケースは珍しくない. 個人企業のデフォルト率が法人企業のデフォルト率より低い理由もこの点にあると考えられる.

#### (2) 業歴 5 年以上 15 年未満

徐々にデフォルト率が低下していく. 生き残った企業の経営者は経営経験を積むことによって経営ノウハウを身につけるため, 業歴 10 年を過ぎるころからデフォルト率が平均に近づき, 15 年ではほぼ平均の水準になる. 中小企業白書 (2002) では, 誕生期の危機を乗り越えた開業者は, 十数年で既存事業者との差がさほど見られなくなるという分析がある.

個人企業もデフォルト率が低下していく傾向は法人企業と同じであるが, デフォルト率低下のペースが緩やかである. 拡大志向のある個人企業は業歴 10 年ぐらいいまでに法人化し, 規模の小さな個人企業が残るため, 先述のとおり, 小回りの利く経営でデフォルトしにくいと考えられる.

#### (3) 業歴 15 年以上 35 年未満

デフォルト率が平均を下回り, 個人企業は業歴 25 年頃まで, 法人企業は業歴 28 年頃までデフォルト率が低下して, それ以降はしばらく横ばいとなる.

#### (4) 業歴 35 年以上 50 年未満

この業歴区間の個人企業は法人企業と大きな違いがある. その違いを明らかにするためにまず, 法人企業の特徴およびその理由を説明する.

法人企業は業歴 35 年を過ぎたころからデフォルト率が少しずつ上昇し始め, 45 年ごろにピークを迎え, その後, また下降する. 日本政策金融公庫総合研究所の新規開業実態調査 (2017) で

は、開業時の経営者の平均年齢は 42.6 歳となっており、業歴 35 年の経営者の年齢は 70 歳代後半になる。事業承継を考える時期であるが、小企業は後継者難という構造問題を抱えており、事業承継の失敗によるデフォルトが増えている可能性がある。

一方、個人企業のデフォルト率は 44 年 (1.61%) と 45 年 (1.73%) にやや高い時期もあるが、これを除くと 40 年を過ぎてもおおむね横ばいである。その理由としては、個人企業の業歴別企業数をみると、業歴が長くなるほど企業数が減少していることから、①表 2 のとおり、大半の個人企業においては事業承継が行われていない、②事業承継する企業は拡大志向があり、法人化する、ことが背景にあると思われる。つまり、個人企業の組織形態を維持しているということは、そもそも事業承継を考えていないと思われる。

#### (5) 業歴 50 年以上

法人企業は業歴 45 年ごろから再びデフォルト率が低下に転じたあと、55 年ごろをボトムに再び上昇するが、事業承継が少ない個人企業にはこの傾向はみられない。もっとも、個人企業の企業数はかなり少なくなるため評価が難しい。

### 5. 業歴別デフォルト率の定式化

4 章で議論した個人企業の業歴別デフォルト率の特徴を踏まえて、本章では業歴別デフォルト率の定式化を試みる。枇々木 他 (2010) と尾木 他 (2016) において行われた法人企業に対する定式化と一部、同じ枠組みで分析を行う。そこで、その分析内容を最初にやや詳しく考察したあとで、個人企業に関して追加的に行った新たな定式化および分析について説明する。

#### 5.1 先行研究における法人企業の業歴別デフォルト率の定式化の考察

枇々木 他 (2010) と尾木 他 (2016) は、法人企業のデフォルト率の形状と業歴区間ごとの特性から、多項式関数 ( $n$  次関数) を業歴関数として定式化を試みた。手順は以下のとおりである。

まず、(5.1) 式の 1~8 次関数のパラメータを最小二乗法により推定している。

$$(5.1) \quad p_g = \beta_0 + \sum_{k=1}^n \beta_k g^k \quad (g = 1, \dots, G)$$

ここで、 $g$  は融資を受けた企業の融資年度時点の業歴、 $p_g$  は業歴  $g$  年で融資を受けた翌年度時点のデフォルト率、 $\beta_k$  は業歴の  $k$  乗に対する回帰係数、 $\beta_0$  は定数項を表している。(5.1) 式において業歴関数の次数  $n$  を 1~8 まで試したときの回帰係数 ( $\beta_0 \sim \beta_8$ ) の  $p$  値と修正  $R^2$  が示されている。尾木 他 (2016) の表 3 をみると、1 次から 8 次へと次数が上がるほど修正  $R^2$  は上昇しているが、7 次関数は 5 乗以上、8 次関数は 2 乗以上の  $p$  値が 5% を超えている。また、同表に示されている (4.1) 式の業歴モデル<sup>9)</sup> (右辺第 3 項の業歴関数の次数  $n$  を 1~8 まで試した場合) の AR 値も合わせて見てみると、6 次関数まではすべての  $p$  値が 1% 以下であるものの、AR 値は 3 次関数で 42.7%、8 次で 43.0% と 3 次以上ではほとんど上昇していない。

加えて、業歴が長くなるにつれて、デフォルト率が低下する傾向があるという大きな特徴もあるので、尾木 他 (2016) は指数関数についても検討した。分析の結果は、どちらの関数も修正  $R^2$  は遜色ないが、時系列での頑健性が若干劣るほか、現場感覚や中小企業研究における指摘などを踏まえると業歴とともにデフォルト率が低下し続ける指数関数は選択しにくいなど、総合的に判断して 3 次関数を選択している。

以上のような先行研究の結果を踏まえて、個人企業についても、まず法人企業と同様に多項式関数と指数関数の当てはめを行い、次に個人企業に対する新たな業歴関数として、ワイブル分布の密度関数および区分線形関数を使ったアプローチを試みる。パラメータの推定は

表 5. 個人企業の業歴別デフォルト率の多項式関数のモデル化.

|                                    | $n=1$                  | $n=2$                  | $n=3$                  | $n=4$                  | $n=5$                   | $n=6$                   | $n=7$                   | $n=8$                    |
|------------------------------------|------------------------|------------------------|------------------------|------------------------|-------------------------|-------------------------|-------------------------|--------------------------|
| $\beta_0$                          | 0.023<br>( $<0.001$ )  | 0.031<br>( $<0.001$ )  | 0.039<br>( $<0.001$ )  | 0.044<br>( $<0.001$ )  | 0.052<br>(0.010)        | 0.059<br>( $<0.001$ )   | 0.065<br>( $<0.001$ )   | 0.069<br>( $<0.001$ )    |
| $\beta_1$<br>[ $\times 10^{-3}$ ]  | -0.251<br>( $<0.001$ ) | -1.015<br>( $<0.001$ ) | -2.534<br>( $<0.001$ ) | -4.214<br>( $<0.001$ ) | -7.423<br>( $<0.001$ )  | -11.901<br>( $<0.001$ ) | -15.891<br>( $<0.001$ ) | -20.123<br>( $<0.001$ )  |
| $\beta_2$<br>[ $\times 10^{-5}$ ]  |                        | 1.253<br>( $<0.001$ )  | 7.428<br>( $<0.001$ )  | 19.581<br>( $<0.001$ ) | 55.198<br>( $<0.001$ )  | 125.002<br>( $<0.001$ ) | 207.303<br>( $<0.001$ ) | 317.899<br>(0.166)       |
| $\beta_3$<br>[ $\times 10^{-6}$ ]  |                        |                        | -0.675<br>( $<0.001$ ) | -3.757<br>( $<0.001$ ) | -19.130<br>( $<0.001$ ) | -63.925<br>( $<0.001$ ) | -136.506<br>(0.004)     | -263.656<br>(0.800)      |
| $\beta_4$<br>[ $\times 10^{-8}$ ]  |                        |                        |                        | 2.526<br>( $<0.001$ )  | 30.757<br>( $<0.002$ )  | 167.128<br>( $<0.003$ ) | 489.113<br>( $<0.004$ ) | 1252.392<br>( $<0.005$ ) |
| $\beta_5$<br>[ $\times 10^{-9}$ ]  |                        |                        |                        |                        | -1.851<br>( $<0.000$ )  | -21.455<br>( $<0.001$ ) | -96.872<br>( $<0.002$ ) | -353.780<br>(0.001)      |
| $\beta_6$<br>[ $\times 10^{-10}$ ] |                        |                        |                        |                        |                         | 1.071<br>( $<0.001$ )   | 9.973<br>( $<0.002$ )   | 58.782<br>(0.003)        |
| $\beta_7$<br>[ $\times 10^{-12}$ ] |                        |                        |                        |                        |                         |                         | -4.170<br>(0.001)       | -53.026<br>(0.007)       |
| $\beta_8$<br>[ $\times 10^{-14}$ ] |                        |                        |                        |                        |                         |                         |                         | 20.023<br>(0.013)        |
| AR値                                | 39.7%                  | 42.0%                  | 42.7%                  | 42.8%                  | 42.9%                   | 43.0%                   | 43.0%                   | 43.0%                    |
| 修正 $R^2$                           | 0.383                  | 0.612                  | 0.765                  | 0.814                  | 0.873                   | 0.918                   | 0.933                   | 0.941                    |

SAS/STAT を用いている。

## 5.2 多項式関数と指数関数の当てはめ

個人企業の業歴別デフォルト率をみると、法人企業と異なる点はあるものの、業歴 5 年前後までデフォルト率が大きく低下し、業歴 15 年以降は平均値を下回って安定するという傾向は共通している。定式化にあたっては、まず、尾木 他 (2016) の分析と同様に多項式関数と指数関数による当てはめを行った。

### 5.2.1 多項式関数

多項式関数を当てはめた結果を表 5 に示す。個人企業は法人企業と異なり、業歴 35 年以降にデフォルト率が上昇する傾向がみられないことから、3 次関数の修正  $R^2$  は 0.765 となり、法人企業の 0.932 を下回る結果となった。次数が上がるにつれて修正  $R^2$  は上昇するが、AR 値は 2 次以降ほとんど上昇しない。デフォルト企業数は業歴が長くなるほど減少し、業歴 20 年を過ぎると、業歴 1 年～2 年の 1/8 程度の数になるため精度が安定しないことが背景にあると思われる。また、尾木 他 (2016) が示したような審査担当者の経験やノウハウとの整合性、中小企業研究における指摘、オーバーフィッティングにより経年劣化が大きくなる可能性などを考慮すると、高次の関数は採用しにくい。

### 5.2.2 指数関数

次に、業歴 35 年以降にデフォルト率が上昇しない傾向を踏まえて、指数関数による当てはめを行った。最小二乗法を用いてパラメータを推定すると (5.2) 式が得られる。

$$(5.2) \quad p_g = 0.0511e^{-0.3158g} + 0.0130 \quad (g = 1, \dots, G)$$

指数関数の  $R^2$  は 0.898 となった。図 7 のとおり、開業から業歴 5 年までにデフォルト率が大きく低下したあと、業歴 25 年ぐらいまで緩やかに低下する期間があるため、指数関数では

この期間がうまく当てはまっていない可能性がある。

### 5.3 新たな業歴関数の提案

法人企業では有効であった多項式関数は、個人企業においては当てはまりが良いとはいえない結果となった。また、指数関数の修正  $R^2$  は 0.898 と、比較的良好な結果となったが、業歴 5 年から 25 年あたりのフィッティングが良くない。

そこで、図 8 の灰色の実線の指数関数に近いワイブル分布の密度関数、さらには、業歴 5 年まで急激に低下、業歴 25 年ぐらいまで緩やかに低下、それ以降はほぼ横ばいという業歴区間ごとの特徴を踏まえて、新たに図 7 の黒色の実線のような全体を 3 つに区分(急激な低下区間、緩やかな低下区間、横ばい区間)した区分線形関数による当てはめを行った。

#### 5.3.1 ワイブル分布の密度関数

信頼性工学の分野では、機械などの故障率は、縦軸に故障率、横軸に時間をとると、初期不良に起因する故障が多く発生し、時間とともに減少する時期、偶発的な事故が発生するのみで故障率が一定で安定する時期、寿命が近づいて時間と共に故障率が増加する時期の 3 つの時期

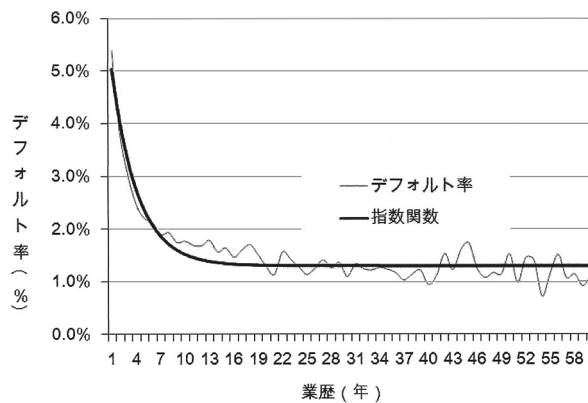


図 7. 指数関数の当てはめ。

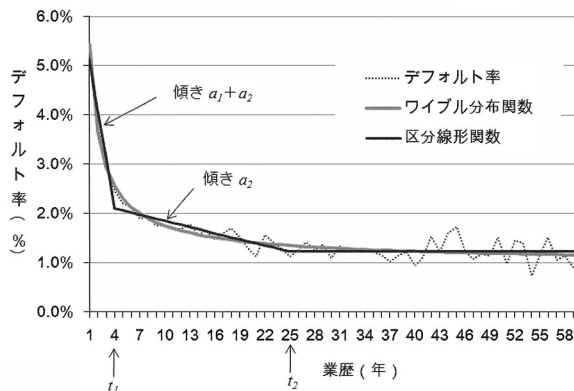


図 8. ワイブル分布の密度関数および区分線形関数の当てはめ。

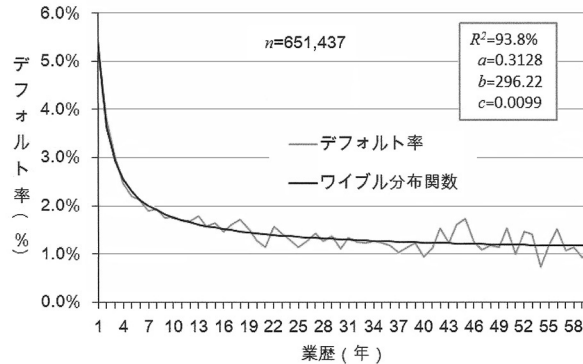


図 9. ワイブル分布の密度関数.

があり、バスタブのような曲線を描くと言われている。個人企業の多くが一代限りで事業承継しないとすれば、個人企業には寿命があると考えられることもできる。とすれば、デフォルト率もバスタブ曲線を描くと仮定することは可能である。このバスタブ曲線にはワイブル分布を仮定することが多い。ワイブル分布の密度関数を(5.3)式に示す。

$$(5.3) \quad f(x) = \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} \exp \left[ -\left(\frac{x}{b}\right)^a \right]$$

ここで、 $a$  は分布の形状、 $b$  は分布の幅を決めるパラメータ、 $x$  は経過時間を入力する。次に、(5.3)式に定数項を加えた(5.4)式を用いてフィッティングする。

$$(5.4) \quad f(x) = \frac{a}{b} \left(\frac{g}{b}\right)^{a-1} \exp \left[ -\left(\frac{g}{b}\right)^a \right] + c \quad (g = 1, \dots, G)$$

ここで、 $g$  は業歴年数である。

結果を図 9 に示す。修正  $R^2$  は 0.938 と、法人企業の業歴関数と同水準となった。

### 5.3.2 区分線形関数

次に、区分線形関数による当てはめを行う。再度、図 8 を見てほしい。区分数を考えるにあたって、業歴 25 年以降の区分数については、法人企業と同様に 35 年から 45 年にかけて上昇し、45 年以降に減少するという傾向に見えなくもない。ただ、4.2.2 項(4)で述べたように、個人企業は事業承継せず 1 代限りの企業が多いこと、5.2.1 項で述べたように、業歴 20 年以降の業歴別デフォルト率はデフォルト企業数が少ないため不安定なことを踏まえて 3 区分とした。

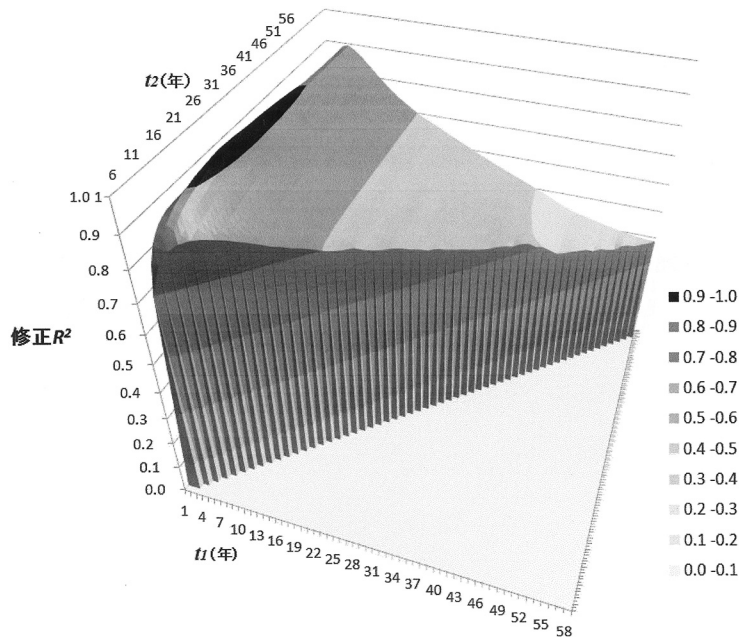
まず、デフォルト率を区分する業歴は、急激な低下が続く創業から  $t_1$  年までと、緩やかな低下が続く  $t_2$  年までの 2 点とし、 $t_2$  年以降は定数とする。次に、 $t_1 = 2, \dots, 59$ ,  $t_2 = 3, \dots, 60$  の 1,711 通りについて、 $y_t = \min(t - t_1, 0)$  と  $\min(t - t_2, 0)$  で回帰し、最も決定係数の高い組み合わせを求める。具体的には以下のとおり、デフォルト率  $y_t$  を区分線形関数で近似した  $\hat{y}_t$  を求める。なお、業歴 60 年超については、業歴 60 年として扱う。

$$\begin{aligned} 1 \leq \text{業歴 } t \leq t_1 & \quad \hat{y}_t = b + a_2(t - t_2) + a_1(t - t_1) \\ t_1 \leq \text{業歴 } t \leq t_2 & \quad \hat{y}_t = b + a_2(t - t_2) \\ t_2 \leq \text{業歴 } t \leq 60 & \quad \hat{y}_t = b \end{aligned}$$

分析の結果を以下に示す。表 6 のとおり、 $t_1$  年は 4 年が、 $t_2$  年は 25 年のパターンの修正  $R^2$  が 0.929 と最も高くなり、ワイブル分布の密度関数よりもやや低いものの、こちらも法人企業

表 6. 区分線形関数の  $t_1, t_2$  の分析結果(上位 20 位の組み合わせ).

| 順位 | 修正 $R^2$ | $t_1$ 年 | $t_2$ 年 | 順位 | 修正 $R^2$ | $t_1$ 年 | $t_2$ 年 |
|----|----------|---------|---------|----|----------|---------|---------|
| 1  | 0.9289   | 4       | 25      | 11 | 0.924    | 4       | 30      |
| 2  | 0.9288   | 4       | 24      | 12 | 0.923    | 4       | 31      |
| 3  | 0.928    | 4       | 26      | 13 | 0.922    | 4       | 19      |
| 4  | 0.928    | 4       | 23      | 14 | 0.921    | 4       | 32      |
| 5  | 0.927    | 4       | 22      | 15 | 0.920    | 4       | 33      |
| 6  | 0.927    | 4       | 27      | 16 | 0.919    | 4       | 18      |
| 7  | 0.927    | 4       | 21      | 17 | 0.918    | 4       | 34      |
| 8  | 0.926    | 4       | 28      | 18 | 0.918    | 3       | 21      |
| 9  | 0.925    | 4       | 29      | 19 | 0.918    | 3       | 20      |
| 10 | 0.925    | 4       | 20      | 20 | 0.918    | 6       | 26      |

図 10. 区分線形関数の  $t_1, t_2$  の修正  $R^2$  の分布.

の 3 次関数とほぼ同水準となった. 3 区分以上にすれば,  $R^2$  がさらに上昇する可能性もあるが, 多項式関数の結果を踏まえるとオーバーフィッティングの可能性があるので, これ以上の検証は行わない. 上位 20 位の組み合わせをみると,  $t_1$  年は 4 年が上位を占め安定している.  $t_2$  年は 25 年の前後が多いが, 図 10 のとおり, 修正  $R^2$  が 0.9 以上の部分(黒色部分)の  $t_2$  年は 15 年ぐらいから 40 年ぐらいまで幅がある.

#### 5.4 業種別業歴関数の当てはめ

業歴別デフォルト率は業種別に異なる. そこで, ワイブル分布の密度関数と区分線形関数について業種別に定式化し, 修正  $R^2$  の比較を行う.



表 7. 業種別ワイブル分布の密度関数の分析結果.  $b$  は  $10^6$  を上限値としてパラメータ推定している.

|          | 製造     | 建設      | 卸小売     | 飲食宿泊   | 医療福祉      | サービス    | 不動産     | 運輸     | その他    |
|----------|--------|---------|---------|--------|-----------|---------|---------|--------|--------|
| $n$      | 44,305 | 111,876 | 145,022 | 85,801 | 43,979    | 123,725 | 47,148  | 14,054 | 35,527 |
| $a$      | 0.7819 | 0.4791  | 0.4533  | 0.4360 | 0.2137    | 0.2996  | 13.0272 | 3.3426 | 0.1423 |
| $b$      | 55.50  | 75.28   | 88.58   | 19.40  | 1,000,000 | 3,082   | 81.41   | 90.44  | 786.7  |
| $c$      | 0.0088 | 0.0116  | 0.0114  | 0.0101 | 0.0059    | 0.0071  | 0.0036  | 0.0126 | 0.0078 |
| 修正 $R^2$ | 55.7%  | 82.4%   | 84.8%   | 86.9%  | 5.3%      | 65.0%   | 6.1%    | 3.1%   | 60.2%  |

#### 5.4.1 ワイブル分布の密度関数(業種別)

まずワイブル分布の密度関数を業種ごとに当てはめた結果を表 7 に示す. 「医療・福祉業」「不動産業」「運輸業」については, 修正  $R^2$  の値が低くなっている. 図 11 の業種別業歴別デフォルト率をみると, この 3 業種は「全体」に比べて終期に上昇しており, グラフの形状が異なっているが, 業歴別デフォルト率がバスタブ曲線になると考えれば, その他の業種については初期で減少, 中期で一定の 2 段階目までを示し, 終期の上昇は何らかの理由で確認できていないと考えることもできる.

ただ, 業種別になると, この 3 業種の業歴 50 年以降の業歴別企業数は件数が 100 件程度と少なくなり, デフォルト率の変動が大きくなることから評価が難しい. さらに, 「医療・福祉業」および「不動産業」のデフォルト率は, ほぼ 1.0% 以下の低位水準での変動であり, 「医療・福祉業」の大宗が医者であること, 「不動産業」には不動産賃貸業が多く含まれていることを考えると, 他の業種に比べれば, おおむね横ばいで安定推移していると見ることができる. なお, 3 業種以外の業種についてもデータ数が少ないため, 55.7%~86.9% と「全体」に比べて低い値となったが, 指数関数的に単調減少するというグラフの基本的な形状は「全体」の形状と類似している.

#### 5.4.2 区分線形関数(業種別)

次に, 区分線形関数について業種別に値を求めた結果を表 8 に示す. 5.4.1 項のワイブル分布の密度関数と同様に, 「医療・福祉業」「不動産業」「運輸業」については, 修正  $R^2$  の値が低くなっている. それ以外の業種についてもワイブル分布の密度関数と同水準となった. 「全体」に比べてサンプル数が少ないため, 57.5%~86.7% と低い値となったが, 図 12 をみると,  $t_1$  年にかけて急激に低下し,  $t_2$  年にかけて緩やかに低下したあと, 横ばいになるという基本的な形状は類似している.

次に, 「医療・福祉業」「不動産業」「運輸業」を除いた業種の  $t_1$  の値をみると, 製造業が 8 年である以外は 4 年前後が多い一方で,  $t_2$  の値は, 20 年~47 年とやや幅がある. 「全体」の  $t_2$  を 15 年くらいから 40 年くらいまでどの値にしても修正  $R^2$  値に差がないのと同様に,  $t_1$  年から  $t_2$  年にかけては低下が緩やかなため, どこから横ばいになるのかは微妙なところがあり, 時系列的にやや不安定となる可能性がある.

そこで,  $t_1$ ,  $t_2$  の値を 4 年と 25 年で固定し, 業種ごとに係数を変えることにした. 再び表 8 で修正  $R^2$  を比較した結果をみると, 業種別に  $t_1$ ,  $t_2$  の年数を 4 年と 25 年で固定しても, 修正  $R^2$  は大きく低下しないことがわかる. ワイブル分布の密度関数との比較では, やや低下するものの, 頑健性を考慮すれば,  $t_1$ ,  $t_2$  の年数は業種別にしない方が好ましいといえる. そこで,  $t_1$ ,  $t_2$  の値は 4 年と 25 年で固定し, 業種ごとに係数を変えた業歴関数を採用することにした.

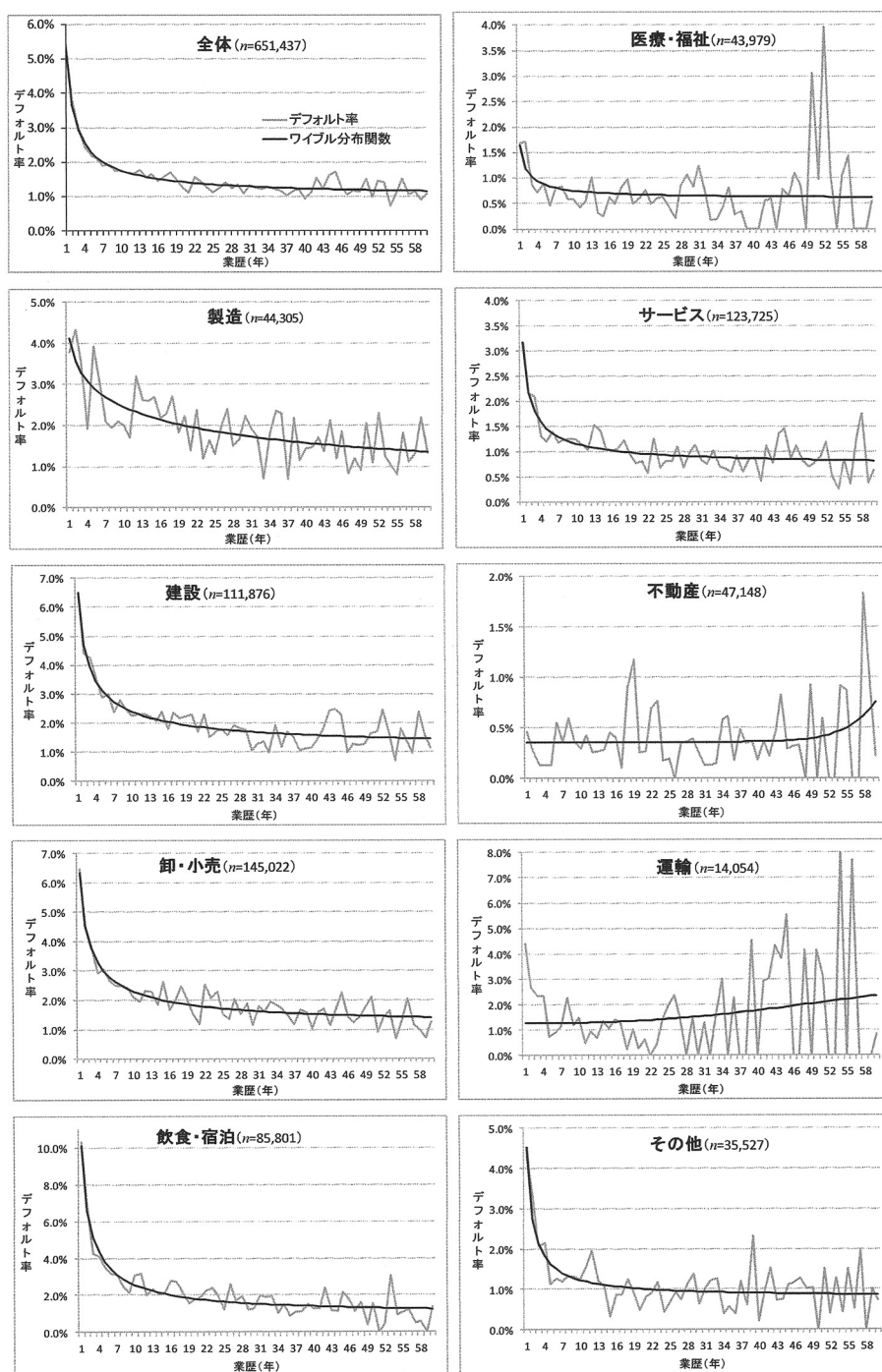


図 11. 業種別業歴別デフォルト率のワイブル分布の密度関数.

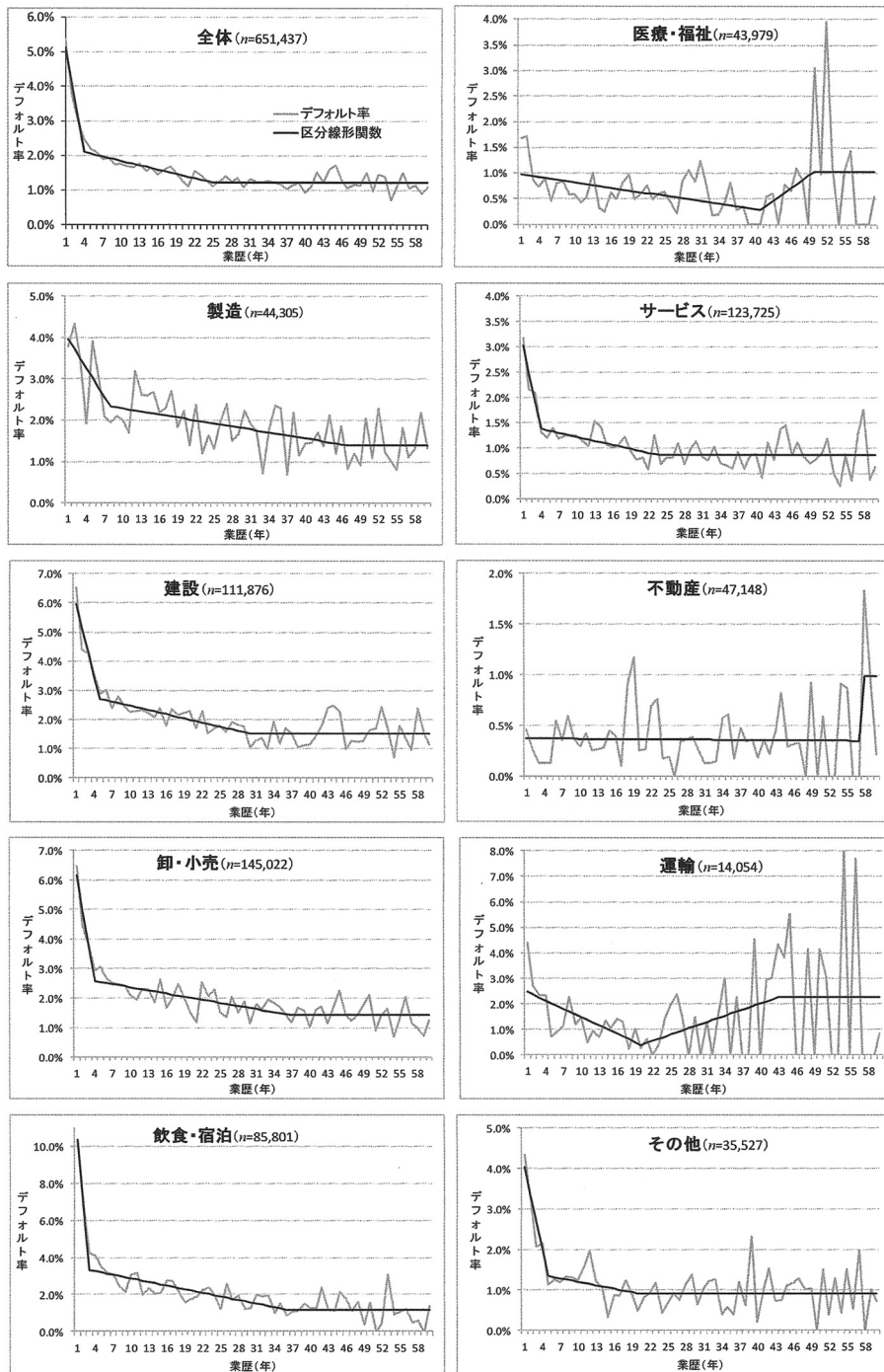


図 12. 業種別業歴別デフォルト率の区分線形関数。

表 8. 業種別区分線形関数の  $t_1, t_2$  の分析結果.

|                             | 製造     | 建設      | 卸小売     | 飲食宿泊   | 医療福祉   | サービス    | 不動産    | 運輸     | その他    |
|-----------------------------|--------|---------|---------|--------|--------|---------|--------|--------|--------|
| $n$                         | 44,305 | 111,876 | 145,022 | 85,801 | 43,979 | 123,725 | 47,148 | 14,054 | 35,527 |
| $t_1$                       | 8      | 5       | 4       | 3      | 41     | 4       | 57     | 20     | 5      |
| $t_2$                       | 47     | 30      | 37      | 36     | 50     | 24      | 58     | 43     | 20     |
| $R^2$                       | 57.5%  | 82.4%   | 83.4%   | 86.7%  | 12.7%  | 66.6%   | 17.2%  | 11.6%  | 62.0%  |
| $t_1=4, t_2=25$<br>修正 $R^2$ | 54.9%  | 81.3%   | 82.3%   | 84.5%  | -      | 66.6%   | -      | -      | 60.8%  |
| 差異                          | -2.6%  | -1.1%   | -1.1%   | -2.2%  | -      | 0.0%    | -      | -      | -1.2%  |

## 6. 個人企業向けモデルにおける業歴関数の有効性

財務モデルに業歴関数を変数に加えた業歴モデルを構築し、財務モデルと業歴モデルのパフォーマンスを AR 値により比較する。業歴関数については、ワイブル分布の密度関数と区分線形関数の修正  $R^2$  の値がほぼ同水準であったことを踏まえて、両者を用いた業歴モデルを構築し、そのパフォーマンスを比較する。なお、「医療・福祉業」「不動産業」「運輸業」については、修正  $R^2$  が低水準にあることから、本研究においては業歴関数を不採用とした。

### 6.1 個人企業向け財務モデルの AR 値

2007 年度～2014 年度の 8 年間に融資した 651,437 件のデータを使用して財務モデルを構築する。モデル構築に必要な財務データ等が揃っていない企業は除外しているため、図 5 の企業数と一致しない。モデルは (4.1) 式の第 3 項を除いたものである。

分析の結果、財務モデルの AR 値は 46.7% と、法人企業向けに比べてインサンプルで 11.9% ポイント高くなった。業種別の結果を表 9 に示す。ここで、「飲食・宿泊業」「医療・福祉業」「サービス業」については、年商規模によって特徴が異なることから規模の大小によって区分している。業種別の AR 値をみると、「サービス」「不動産」が高く、「飲食・宿泊 s」「医療・福祉 s」など規模の小さな企業の AR 値が低くなっている。

### 6.2 個人企業向け業歴モデルの構築

財務モデルと業歴モデルのパフォーマンスを AR 値により比較するため、財務モデルに、第 3 項に業歴関数を加えた業歴モデルを構築する。業歴関数には、修正  $R^2$  が同水準であったワイブル分布の密度関数を使ったモデル（以下、ワイブル分布関数モデルという）と区分線形関数（以下、区分線形関数モデルという）を用いる。

#### (1) ワイブル分布関数モデル

ワイブル分布の密度関数のパラメータも含めて、ロジットモデルの中で推定することは難しいので、業歴関数 (5.4) 式に基づき、(6.1) 式のパラメータ  $a_h, b_h, c_h$  を最小二乗法で業種  $h$  ごとに推定する。

$$(6.1) \quad p_g = \frac{a_h}{b_h} \left( \frac{g}{b_h} \right)^{a_h-1} \exp \left[ - \left( \frac{g}{b_h} \right)^{a_h} \right] + c_h, \quad (g = 1, \dots, G)$$

推定されたパラメータ  $a_h^*, b_h^*, c_h^*$  と個人企業  $i (i = 1, \dots, I)$  の業歴  $g_i$  を使用して、業歴関数によるデフォルト率の推定値  $q_i^*$  を業種  $h (h = 1, \dots, H)$  ごとに (6.2) 式で計算する。

表 9. 財務モデルの推計結果.

| 業種    | $n$     | 定数項<br>( $\alpha_0$ ) | 申告形態<br>( $\alpha_1$ ) | 流動性指標<br>( $\alpha_2$ ) | 安全性指標<br>( $\alpha_3$ ) | 効率性指標<br>( $\alpha_4$ ) | AR値  |
|-------|---------|-----------------------|------------------------|-------------------------|-------------------------|-------------------------|------|
| 製造    | 44,597  | 3.68<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.06<br>(0.028)         | -0.12<br>( $<0.001$ )   | 0.39 |
| 建設    | 112,653 | 3.74<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.74<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.36<br>( $<0.001$ )   | 0.37 |
| 卸小売   | 145,141 | 3.64<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.07<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.40 |
| 飲食宿泊s | 47,911  | 3.24<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.36 |
| 飲食宿泊  | 37,510  | 3.67<br>( $<0.001$ )  | -0.89<br>( $<0.001$ )  | 0.89<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.39 |
| 医療福祉s | 16,087  | 4.02<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.36 |
| 医療福祉  | 27,759  | 5.28<br>( $<0.001$ )  | -2.17<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.43 |
| サービスs | 85,221  | 4.16<br>(0.003)       | -0.68<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.06<br>(0.002)         | -0.37<br>( $<0.001$ )   | 0.41 |
| サービス  | 39,786  | 4.48<br>(0.003)       | -0.68<br>( $<0.001$ )  | 0.94<br>(0.002)         | 0.06<br>( $<0.001$ )    | -0.37<br>( $<0.001$ )   | 0.50 |
| 不動産   | 48,925  | 4.63<br>( $<0.001$ )  | -0.81<br>(0.045)       | 0.49<br>( $<0.001$ )    | 0.07<br>(0.041)         | -0.64<br>( $<0.001$ )   | 0.51 |
| 運輸    | 14,073  | 4.02<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.59<br>( $<0.001$ )   | 0.44 |
| その他   | 31,774  | 4.02<br>( $<0.001$ )  | -0.47<br>( $<0.001$ )  | 0.49<br>( $<0.001$ )    | 0.05<br>( $<0.001$ )    | -0.12<br>( $<0.001$ )   | 0.37 |

注 1：業種の後部に s が付いているものは、年商規模が「小」のものである。

注 2：その他業種は、全体モデルの係数を使用している。

注 3：カッコ内は Wald カイ 2 乗検定に対する  $p$  値。業種モデルの係数は、全体モデルの係数に  $p$  値 5% 水準で有意になった業種ダミー（定数項ダミー、係数ダミー）の値を加えたものである。したがって、業種別モデルの係数の  $p$  値は係数ダミーの値に対する結果である。

注 4：実際に使用しているモデルであり、具体的な財務指標名については、実務での影響を考慮し、秘匿している点をご理解いただきたい。

$$(6.2) \quad q_i^* = \frac{a_h^*}{b_h^*} \left( \frac{g_i}{b_h^*} \right)^{a_h^*-1} \exp \left[ - \left( \frac{g_i}{b_h^*} \right)^{a_h^*} \right] + c_h^*, \quad (i \in F_h; h = 1, \dots, H)$$

ここで、 $F_h$  を業種  $h$  に属する企業の集合とする。ワイブル分布関数モデルを (6.3) 式に示す。個人企業  $i (i = 1, \dots, I)$  について、財務変数  $f_{i,j} (j = 1, \dots, J)$  と業歴関数によって計算されたデフォルト率  $q_i^*$  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  $a_h, \beta_{j,h}, w_h$  を業種  $h$  ごとに推定する。

$$(6.3) \quad p_i = \frac{1}{1 + e^{z_i}},$$

$$z_i = \ln \left( \frac{1 - p_i}{p_i} \right) = \alpha_h + \sum_{j=1}^J \beta_{j,h} f_{i,j} + w_h q_i^*, \quad (i \in F_h; h = 1, \dots, H)$$

#### (2) 区分線形関数モデル

区分線形関数モデルを (6.2) 式に示す。個人企業  $i (i = 1, \dots, I)$  について、財務変数  $f_{i,j} (j = 1, \dots, J)$  と業歴  $g_i$  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  $a_h, \beta_{j,h}, \gamma_{h,1}, \gamma_{h,2}$  を推定する.

$$\begin{aligned}
 (6.4) \quad p_i &= \frac{1}{1 + e^{z_i}}, \\
 z_i &= \ln \left( \frac{1 - p_i}{p_i} \right) \\
 &= \alpha_h + \sum_{j=1}^J \beta_{j,h} f_{i,j} + \gamma_{h,1} \times \min(g_i - 4, 0) + \gamma_{h,2} \times \min(g_i - 25, 0), \\
 &\quad (i \in F_h; h = 1, \dots, H)
 \end{aligned}$$

### 6.3 個人企業向け業歴モデルの AR 値

業歴モデルの推計結果を表 10 に示す. 業歴関数を投入すると全体の AR 値はワイブル分布関数モデルが 49.0%, 区分線形関数モデルが 49.3% となり, ほぼ同水準で, それぞれ法人企業モデルに比べて 2.3% ポイント, 2.6% ポイント上昇する. 法人企業に比べて AR 値の水準は高いが, 業歴関数を加えたことによる改善幅はいずれも小さい. 仮説では, 法人企業と同様に, 業歴別デフォルト率は, 非開示となっている事業主の私的資産が多くを占める小企業のパフォーマンスを示す指標として期待したが, 本分析結果を見る限り, 開示されている事業資産に事業主の私的資産が相応に含まれている (非開示の資産が少ない), もしくは, 非開示の私的資産は事業に投入しない傾向にあるなどの理由で期待したほどの効果が得られなかった可能性が考えられる.

表 11 と表 12 で業種別に AR 値の改善幅をみると「飲食宿泊」はワイブル分布関数モデルが 9.6% ポイント, 区分線形関数モデルが 9.4% ポイント, 「その他」が同 6.4% ポイント, 同 6.5% ポイント, 「飲食宿泊 s」が同 6.1% ポイント, 同 5.9% ポイントと大きくなっている. それ以外は改善幅が小さく, サービス業はほとんど改善していない. ただ, AR 値の水準の低い業種ほど改善幅が大きい傾向にあり, 業歴関数を投入したことにより, 業種間の差が改善されている.

二つの業歴関数を比較すると, 水準および改善幅に大きな違いはないが, ワイブル分布関数モデルでは, 「製造」「飲食宿泊」「サービス s」「サービス」「その他」の業種のいくつかの財務変数の  $p$  値が 5% 水準で有意にならなかった.

### 6.4 モデルの頑健性の検証

個人企業向け業歴モデルの頑健性を確認するために, アウトオブサンプルテストを行った. アウトオブサンプルは, 2015 年度に融資を行った債権 70,383 件で, 2007 年度～2014 年度に融資した 651,437 件のインサンプルに対して 10.8% のサンプルサイズである. 結果を表 13 に示す. 全体の AR 値は, いずれのモデルでもほぼ横ばいで良好な結果となっている. ワイブル分布関数モデルと区分線形関数モデルの差はほとんど確認できない.

表 10. 業歴モデルの AR 値 (インサンプル). 法人企業の数値は, 尾木 他 (2016) の値.

|                      | 個人企業(a)       | 法人企業(b)       | (a) - (b)     |
|----------------------|---------------|---------------|---------------|
| 財務モデル                | 46.7          | 34.8          | 11.9          |
| ワイブル分布関数モデル<br>(改善幅) | 49.0<br>(2.3) | 42.4<br>(7.6) | 6.6<br>(-5.3) |
| 区分線形関数モデル<br>(改善幅)   | 49.3<br>(2.6) | 42.4<br>(7.6) | 6.9<br>(-5.0) |

表 11. 区分線形関数業歴モデルの推計結果((6.3)式)。

| 業種    | 定数項<br>( $\alpha$ )  | 申告形態<br>( $\beta 1$ ) | 流動性指標<br>( $\beta 2$ ) | 安全性指標<br>( $\beta 3$ ) | 効率性指標<br>( $\beta 4$ ) | $\min(ki-4,0)$<br>( $\gamma 1$ ) | $\min(ki-25,0)$<br>( $\gamma 2$ ) | AR値   | 改善幅   |
|-------|----------------------|-----------------------|------------------------|------------------------|------------------------|----------------------------------|-----------------------------------|-------|-------|
| 製造    | 3.76<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.06<br>(0.004)        | -0.16<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.01<br>( $<0.001$ )              | 40.2% | 1.6%  |
| 建設    | 3.93<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.72<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.32<br>( $<0.004$ )  | 0.24<br>( $<0.001$ )             | 0.02<br>(0.001)                   | 40.5% | 3.5%  |
| 卸小売   | 3.87<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.06<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.02<br>(0.002)                   | 42.6% | 2.7%  |
| 飲食宿泊s | 3.63<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.03<br>( $<0.001$ )              | 42.0% | 5.9%  |
| 飲食宿泊  | 4.18<br>( $<0.001$ ) | -0.69<br>( $<0.001$ ) | 0.88<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | 0.35<br>(0.004)                  | 0.03<br>( $<0.001$ )              | 48.6% | 9.4%  |
| 医療福祉s | 4.08<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | -<br>-                           | -<br>-                            | 36.2% | -     |
| 医療福祉  | 5.31<br>( $<0.001$ ) | -2.17<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | -<br>-                           | -<br>-                            | 42.7% | -     |
| サービスs | 4.29<br>(0.008)      | -0.63<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.06<br>( $<0.001$ )   | -0.36<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.01<br>( $<0.001$ )              | 41.8% | 0.4%  |
| サービス  | 4.58<br>( $<0.001$ ) | -0.63<br>( $<0.001$ ) | 0.93<br>(0.003)        | 0.06<br>( $<0.001$ )   | -0.36<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.01<br>( $<0.001$ )              | 50.0% | -0.1% |
| 不動産   | 4.62<br>( $<0.001$ ) | -0.81<br>(0.010)      | 0.50<br>( $<0.001$ )   | 0.07<br>( $<0.001$ )   | -0.63<br>(0.010)       | -<br>-                           | -<br>-                            | 50.8% | -     |
| 運輸    | 4.08<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.69<br>( $<0.001$ )  | -<br>-                           | -<br>-                            | 43.7% | -     |
| その他   | 4.08<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.50<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.16<br>( $<0.001$ )  | 0.24<br>( $<0.001$ )             | 0.01<br>( $<0.001$ )              | 43.3% | 6.5%  |

注 1：その他業種は、全体モデルの係数を使用している。

注 2：カッコ内は Wald カイ 2 乗検定に対する  $p$  値。業種モデルの係数は、全体モデルの係数に  $p$  値 5% 水準で有意になった業種ダミー(定数項ダミー、係数ダミー)の値を加えたものである。したがって、業種別モデルの係数の  $p$  値は係数ダミーの値に対する結果である。

業種別にみると、「サービス業」が 10% ポイント以上の上昇と変動幅が大きくなっている。サービス業は年商規模がやや大きな「サービス業」と年商規模がやや小さな「サービス業s」の二つに分類している。「サービス業」は、①サンプル数が少ない、②「学術研究、専門・技術サービス業」「生活関連サービス業、娯楽業」「複合サービス事業」「サービス業(他に分類されないもの)」など大分類ベースでも 4 つに分類されるなど多種多様である、③年商規模の上限がなく規模のバラつきが大きい、といった複数の要因が影響していると考えられる。

また、「製造業」「医療・福祉業s」「医療・福祉業」「運輸業」とサンプル数の少ない業種でやや変動幅が大きくなる傾向にあるほか、景気変動の影響を受けやすい「建設業」が上昇している。ワイブル分布関数モデルと区分線形関数モデルは、アウトオブサンプルテストの結果においても大きな差はみられない。

以上のように、業種別にみると、サンプル数や景気変動、業種特性などの影響を受けるため、AR 値の変動幅はやや大きくなる傾向があるものの、今回のアウトオブサンプルテストの結果を見る限り、実務での使用には大きな問題はなさそうである。ただ、通常、アウトオブサンプルの AR 値はインサンプルより低い傾向にあるが、本検証では上回っているケースもあり、データ数が少なく不安定な結果となっている点是否めない。頑健性については、データの蓄積を待って、引き続き検証する必要があると考えている。

## 7. まとめと今後の課題

本研究では、個人企業向けスコアリングモデルの精度を改善するために、業歴別デフォルト

表 12. ワイブル分布密度関数業歴モデルの推計結果((6.4)式).

| 業種    | 定数項<br>( $\alpha$ )  | 申告形態<br>( $\beta_1$ ) | 流動性指標<br>( $\beta_2$ ) | 安全性指標<br>( $\beta_3$ ) | 効率性指標<br>( $\beta_4$ ) | ワイブル分布<br>( $wg$ )     | AR値   | 改善幅  |
|-------|----------------------|-----------------------|------------------------|------------------------|------------------------|------------------------|-------|------|
| 製造    | 4.21<br>( $<0.001$ ) | -0.27<br>(0.379)      | 0.55<br>(0.030)        | 0.06<br>( $<0.001$ )   | -0.15<br>(0.529)       | -27.94<br>(0.081)      | 39.3% | 0.6% |
| 建設    | 4.29<br>( $<0.001$ ) | -0.41<br>( $<0.001$ ) | 0.71<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.32<br>(0.038)       | -24.78<br>(0.001)      | 40.6% | 3.6% |
| 卸小売   | 4.21<br>( $<0.001$ ) | -0.39<br>(0.007)      | 0.57<br>( $<0.001$ )   | 0.06<br>( $<0.001$ )   | -0.18<br>(0.007)       | -25.78<br>(0.002)      | 42.8% | 2.9% |
| 飲食宿泊s | 3.74<br>( $<0.001$ ) | -0.39<br>(0.020)      | 0.46<br>( $<0.001$ )   | 0.05<br>(0.021)        | 0.00<br>( $<0.001$ )   | -17.19<br>( $<0.001$ ) | 42.3% | 6.1% |
| 飲食宿泊  | 4.39<br>( $<0.001$ ) | -0.67<br>(0.023)      | 0.95<br>(0.002)        | 0.04<br>(0.006)        | -0.41<br>(0.436)       | -21.57<br>( $<0.001$ ) | 48.7% | 9.6% |
| 医療福祉s | 4.22<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.58<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.20<br>( $<0.001$ )  | -                      | 36.2% | -    |
| 医療福祉  | 5.28<br>( $<0.001$ ) | -2.17<br>( $<0.001$ ) | 0.58<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.20<br>( $<0.001$ )  | -                      | 42.7% | -    |
| サービスs | 4.66<br>(0.008)      | -0.62<br>(0.001)      | 0.31<br>(0.018)        | 0.07<br>( $<0.001$ )   | -0.33<br>(0.050)       | -25.64<br>(0.045)      | 41.8% | 0.4% |
| サービス  | 5.36<br>( $<0.001$ ) | -0.85<br>(0.128)      | 0.89<br>(0.051)        | 0.06<br>(0.007)        | -0.57<br>(0.210)       | -48.58<br>(0.1294)     | 50.8% | 0.6% |
| 不動産   | 4.65<br>( $<0.001$ ) | -0.84<br>(0.005)      | 0.58<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.64<br>(0.010)       | -                      | 50.8% | -    |
| 運輸    | 4.22<br>( $<0.001$ ) | -0.38<br>( $<0.001$ ) | 0.58<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.89<br>( $<0.001$ )  | -                      | 43.7% | -    |
| その他   | 4.74<br>( $<0.001$ ) | -0.29<br>(0.337)      | 0.57<br>( $<0.001$ )   | 0.05<br>( $<0.001$ )   | -0.19<br>(0.322)       | -51.31<br>( $<0.001$ ) | 43.1% | 6.4% |

注 1: その他業種は, 全体モデルの係数を使用している.

注 2: カッコ内は Wald カイ 2 乗検定に対する  $p$  値. 業種モデルの係数は, 全体モデルの係数に  $p$  値 5% 水準で有意になった業種ダミー(定数項ダミー, 係数ダミー)の値を加えたものである. したがって, 業種別モデルの係数の  $p$  値は係数ダミーの値に対する結果である. 網掛けは,  $p$  値 5% 水準で有意にならなかったもの.

表 13. アウトオブサンプルテストの結果(AR 値). 医療福祉, 不動産業, 運輸業は業歴関数を用いていない(シャドウ部分).

| 業種    | n      | ワイブル分布関数モデル |        |                        | 区分線形関数モデル |        |                        |
|-------|--------|-------------|--------|------------------------|-----------|--------|------------------------|
|       |        | イン(A)       | アウト(B) | 差分(A-B)                | イン(C)     | アウト(D) | 差分(C-D)                |
| 製造    | 3,629  | 39.8%       | 36.8%  | -2.4% <sup>※</sup> イント | 40.2%     | 36.3%  | -3.9% <sup>※</sup> イント |
| 建設    | 11,998 | 40.7%       | 46.1%  | 5.5% <sup>※</sup> イント  | 40.5%     | 45.9%  | 5.4% <sup>※</sup> イント  |
| 卸小売   | 13,707 | 42.7%       | 40.7%  | -2.0% <sup>※</sup> イント | 42.6%     | 40.3%  | -2.3% <sup>※</sup> イント |
| 飲食宿泊s | 5,006  | 42.2%       | 44.5%  | 2.2% <sup>※</sup> イント  | 42.0%     | 43.5%  | 1.5% <sup>※</sup> イント  |
| 飲食宿泊  | 4,200  | 48.4%       | 51.2%  | 2.5% <sup>※</sup> イント  | 48.6%     | 51.9%  | 3.3% <sup>※</sup> イント  |
| 医療福祉s | 2,251  | 36.2%       | 42.2%  | 6.0% <sup>※</sup> イント  | 36.2%     | 42.2%  | 6.0% <sup>※</sup> イント  |
| 医療福祉  | 2,832  | 42.7%       | 37.8%  | -5.0% <sup>※</sup> イント | 42.7%     | 37.8%  | -5.0% <sup>※</sup> イント |
| サービスs | 9,767  | 41.8%       | 39.5%  | -2.3% <sup>※</sup> イント | 41.8%     | 40.2%  | -1.7% <sup>※</sup> イント |
| サービス  | 4,111  | 50.7%       | 62.0%  | 11.2% <sup>※</sup> イント | 50.0%     | 63.1%  | 13.1% <sup>※</sup> イント |
| 不動産   | 7,046  | 50.8%       | 49.1%  | -1.7% <sup>※</sup> イント | 50.8%     | 49.1%  | -1.7% <sup>※</sup> イント |
| 運輸    | 1,566  | 43.7%       | 39.1%  | -4.7% <sup>※</sup> イント | 43.7%     | 39.1%  | -4.7% <sup>※</sup> イント |
| その他   | 4,270  | 43.4%       | 43.5%  | 0.3% <sup>※</sup> イント  | 43.3%     | 45.6%  | 2.3% <sup>※</sup> イント  |
| 全体    | 70,383 | 49.0%       | 50.2%  | 1.2% <sup>※</sup> イント  | 49.1%     | 49.7%  | 0.6% <sup>※</sup> イント  |

率を定式化した業歴関数を投入した. 先行研究をみると, 尾木 他 (2016) は, 小規模な法人企業向けモデルに業歴関数を用いることを提案し, 業歴別デフォルト率が重要な変数であることを示した. さらに, 業歴別デフォルト率は, 決算書の数値には表れない小企業のパフォーマンス



スを示す指標と考えられるが、主要な要素として経営者の個人資産の大きさが含まれている可能性があることも示した。

個人企業の場合、事業資産と事業主の私的資産が混在しており、それを明確に区分することは難しいが、個人企業の業歴別デフォルト率を事業主の私的資産が主要な要素とする個人企業のパフォーマンスを示す指標として利用できれば、それを定式化した業歴関数をモデルの変数に含めることによって、個人企業向けモデルの精度を改善できる可能性があると考えた。そこで、業歴別デフォルト率を定式化した業歴関数をモデルの変数に投入して、個人企業向けモデルの精度改善を図ることにした。公庫が保有する約 68 万件のデータを使用して分析した結果、以下のことがわかった。

- (1) 個人企業の業歴別デフォルト率は、業歴 4 年まで急激に低下し、その後 25 年まで緩やかに低下、25 年を過ぎるとおおむね一定になるという法人企業と異なる特徴を有していることがわかった。法人企業で有効であった 3 次関数の  $R^2$  は 0.765 となり、法人企業の 0.932 を大きく下回る結果となり、個人企業に対しては 3 次関数を利用することは適切でないことがわかった。
- (2) 個人企業の特徴を踏まえて、ワイブル分布の密度関数および業歴が「0-3 年」「4-25 年」「26 年以上」の 3 つの区分に分けた区分線形関数を用いた結果、修正  $R^2$  はいずれも 93% 程度となり、法人企業の 3 次関数と同程度の水準を確保することができた。
- (3) ワイブル分布関数の密度関数および区分線形関数をモデルの変数に投入すると、業種別では AR 値を最大でワイブルは 9.6% ポイント、区分線形で 9.4% ポイント、全体では、ワイブルで 2.3% ポイント、区分線形で 2.6% ポイント向上させることができた。
- (4) ワイブル分布関数の密度関数と区分線形関数は、ほぼ同水準のパフォーマンスを示しているが、ワイブル分布関数の密度関数を投入するといくつかの財務変数が有意にならなくなることで、実務では線形の方がオーバーフィッティングしにくい経験則を考慮すると、区分線形関数が望ましい可能性がある。

本研究の成果において、実務上の観点からは(3)の貢献が大きい。個人企業は決算書がなく、その代わりとなる税務申告書に貸借対照表(B/S)を添付していない企業も多く、B/Sに関する情報は法人企業に比べて不足している。モデルの精度を改善するには、個人企業の資産負債に関連する客観的な変数を加えることが有効であるが、個人企業は家計と事業経営の関係が緊密で、会計が混在しているため、資産や負債の把握が極めて困難かつ手間暇がかかる。

本研究によって、個人企業においても業歴別デフォルト率を定式化した業歴関数が有効であることがわかり、信用スコアリングモデルの精度を手間とコストをかけずに高めることが可能になった。ただ、改善幅は法人企業の 7.6% ポイントと比べると低い水準にとどまった。仮説では、個人企業も法人企業と同様に、業歴別デフォルト率が事業主の私的資産の大きさも含めた非開示の小企業のパフォーマンスを示す指標として有効となることを期待したが、開示されている事業資産の中に事業主の私的資産が相応に含まれている(非開示の資産が少ない)、もしくは、非開示の私的資産は事業に投入しない可能性があるため、効果が小さかったと考えられる。

本研究の結果を踏まえた今後の課題は以下のとおりである。

- (1) 個人企業向けモデルの精度をさらに改善するため、入手の手間とコストがかからない資産負債関連の変数を追加することを検討する。
- (2) 副業の収入が個人企業の資産形成に少なからぬ影響を与えていると思われるため、副業に関するファクターの導入について検討する。

- (3) ワイブル分布関数の密度関数と区分線形関数の精度比較を行うため、データの蓄積を待つて再度検証を行う。

なお、他の金融機関に対する利用可能性に関して、本分析で利用したデータは公庫が融資した企業のみであり、サンプルバイアスを持つ可能性があることには注意が必要である。現在、民間金融機関は取引先拡大のために、個人企業をはじめとした小企業への融資を積極化しており、これらのデータを含む本論文の結果は、リスクテイクを積極化する民間金融機関に多くの示唆を与えるものと信じている。本研究の成果が小企業向けの融資を行っている他の金融機関における信用リスク管理の高度化の参考になれば幸いである。

注.

- <sup>1)</sup> 個人企業とは法人格を持たない個人で事業を行う個人事業者のことであるが、個人企業の定義は明確にされておらず、個人事業主、自営業者、自己雇用者といった表現も用いられる。本論文では、いずれも表記を個人企業に統一して用いる。総務省統計局「統計 Today No.82」(総務省統計局, 2014)によれば「個人企業とは、法人としての登記をしていない個人が経営する経営体を意味します。例えば、街中の法人登記をしていない個人営業の工場、工務店、小売店、飲食店、美容室、学習塾、税理士などは個人企業ですし、医療法人となっていない医療機関も 1 つの個人企業です。また、例えば、複数の店舗を持つ美容のチェーン企業であっても法人登記していなければ個人企業になります。」と記されている。
- <sup>2)</sup> 個人資産額以外のファクターとして経営者の資質や経営ノウハウなどが考えられるが検証は困難である。
- <sup>3)</sup> AR 値 (Accuracy Ratio) はスコアリングモデルの精度を示す代表的な指標として普及している。所得の集中度や不平等度を測る「Gini 係数」の考え方を応用したもので、推計されたデフォルト確率を順序性の尺度とし、モデルの精度を評価する。0 から 1 の値をとり、1 に近いほど精度が高い。詳しくは、山下・三浦 (2011) を参照されたい。
- <sup>4)</sup> デフォルトに統一的な定義はなく、金融機関やモデルによって定義や算出方法が異なるので、比較には注意が必要である。たとえば、RDB は、過去 12 カ月以内に、3 カ月以上延滞先もしくは破綻懸念先以下の債務者区分に初めて該当した債務者をデフォルトと定義している。また、CRD 協会は①3 カ月以上延滞先、②実質破綻先、③破綻先、④信用保証協会による代位弁済先と定義している。
- <sup>5)</sup> 企業  $i$  の決算書 1 期分の  $J$  個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  $\alpha_j(j = 1, \dots, J)$  を推定する。ここで、 $p_i$  は融資した企業が翌年にデフォルトする確率、 $N$  は企業数を表す。 $z_i$  が大きいほどデフォルト確率は低くなる。

$$p_i = \frac{1}{1 + e^{z_i}}, z_i = \ln \left( \frac{1 - p_i}{p_i} \right) = \alpha_0 + \sum_{j=1}^J \alpha_j f_{ij}$$

次に、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  $CS_i$  を計算する。

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

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。これは信用スコアが  $z_i^* = Z(1\%)$  ならば  $\eta_0$  点,  $z_i^* = Z(99\%)$  ならば  $\eta_1$  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。
- 6) 法人企業も同じ融資年度で分析を行えば、観測期間の影響も受けずに比較が可能である。ただし、本研究の目的は個人企業向けのモデルを構築し、その特徴を考察することであるため、この点は今後の課題としたい。
  - 7) 日本政策金融公庫国民生活事業本部は従業者数 20 人(商業・サービス業は 5 人)未満の小企業への融資が約 8 割を占める。融資先の業種や地域構成比はわが国の事業所の構成比と比べて大きな差はない。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、融資先の特徴等については、枇々木 他 (2010) および公庫の HP を参照されたい。
  - 8) データ数で加重平均しているもので、5 年間のデフォルト率と等しい。ただし、業歴 1 年は 3 年間(1~3 年)、業歴 2 年は 4 年間(1~4 年)のデフォルト率を計算している。
  - 9) 業歴以外に、財務指標 15 変数(全業種共通 9 変数、業種別 6 変数)を使用している。モデルの概要は枇々木 他 (2010) も参照されたい。

## 参 考 文 献

- Akben-Selcuk, E. (2016). Does firm age affect profitability? Evidence from Turkey, *International Journal of Economic Sciences*, **5**(3), 1-9.
- Altman, E. I. and Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the U.S. market, *ABACUS*, **43**, 332-357.
- Cabral, L. M. B. and Mata, J. (2003). On the evolution of the firm size distribution: Facts and theory, *American Economic Review*, **93**(4), 1075-1090.
- 中小企業庁 (2002). 2002 年版中小企業白書, 中小企業庁, 東京.
- 中小企業庁 (2014). 2014 年版中小企業白書, 中小企業庁, 東京.
- Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries, *Journal of Industrial Economics*, **35**(4), 567-581.
- Gweyi, M. O. (2013). Credit risk mitigation strategies adopted by commercial banks in Kenya, *International Journal of Business and Social Science*, **4**, 71-87.
- 枇々木規雄, 尾木研三, 戸城正浩 (2010). 小企業向けスコアリングモデルにおける業歴の有効性, 『定量的信用リスク評価とその応用』, ジャファイ・ジャーナル: 金融工学と市場計量分析, 83-116, 朝倉書店, 東京.
- Kneiding, C. and Kritikos, S. A. (2011). Funding self-employment: The role of consumer credit, IZA Discussion Paper, 6093 IZA, Germany.
- 国民生活金融公庫総合研究所 (2002). 自己雇用者(≡ 自営業者)に関する実態調査～アンケート結果の概要～, 国民生活金融公庫レポート, 国民生活金融公庫総合研究所, 東京.
- Loderer, C. F. and Waelchli, U. (2010). Firm age and performance, available at SSRN: <http://ssrn.com/abstract=1342248>.
- 森平爽一郎, 岡崎貴治 (2009). 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, WIF-09-004, 早稲田大学ファイナンス総合研究所, 東京.
- 森川正之 (2008). 同族企業の生産性—日本企業のマイクロデータによる実証分析—, RIETI Discussion Paper Series, 08-j-029, 独立行政法人経済産業研究所, 東京.
- 内閣府 (2011). 平成 23 年 年次経済財政報告, 内閣府, 東京.
- 日本政策金融公庫総合研究所 (2016). 中小企業の事業承継に関するインターネット調査～アンケート結果の概要～, 日本政策金融公庫総合研究所, 東京.
- 日本政策金融公庫総合研究所 (2017). 2017 年度新規開業実態調査～アンケート結果の概要～, 日本政策

- 金融公庫総合研究所, 東京.
- 尾木研三, 戸城正浩, 枇々木規雄 (2016). 小規模企業向け信用スコアリングモデルにおける業歴の頑健性と経営者の個人資産との関係性, *Transactions of the Operations Research Society of Japan*, **59**, 134-159.
- 尾木研三, 内海裕一, 枇々木規雄 (2017). 創業企業の信用リスクモデル, 『リスク管理・保険とヘッジ』, ジャフィー・ジャーナル: 金融工学と市場計量分析, 133-156, 朝倉書店, 東京.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, **18**, 109-131.
- Sakai, K., Uesugi, I. and Watanabe, T. (2010). Firm age and the evolution of borrowing costs: Evidence from Japanese small firms, *Journal of Banking and Finance*, **34**, 1970-1981.
- Schreiner, M. (2000). Credit scoring for microfinance: Can it work?, *Journal of Microfinance*, **2**, 105-111.
- 清水龍宝 (1985). 中小企業倒産の構造的要因の分析, 三田商学研究, **28**(2), 1-24.
- 総務省統計局 (2014). 平成 26 年度統計 Today, No.82, 総務省統計局, 東京, <https://www.stat.go.jp/info/today/082.html>.
- 総務省統計局 (2016). 平成 26 年経済センサス—基礎調査, 総務省統計局, 東京.
- 高川泉, 亀田制作 (2008). わが国における個人企業の動向とその背景, 日銀レビュー, 2008-J-12, 日本銀行, 東京.
- 友田信男 (2008). 中小企業の倒産動向, 日本政策金融公庫調査月報, **2**, 4-15.
- 鶴田大輔 (2005). ノンバンク融資と中小企業のモラルハザード問題, RIETI Discussion Paper Series, 05-J-035, 独立行政法人経済産業研究所, 東京.
- 上村協子 (2006). 自営業世帯の家計資産, 信金中金月報, **397**, 55-74.
- 山下智志, 三浦翔 (2011). 『信用リスクモデルの予測精度—AR 値と評価指標—』, 朝倉書店, 東京.
- 山下智志, 川口昇, 敦賀智裕 (2003). 信用リスクモデルの評価方法に関する考察と比較, Financial Research and Training Center Discussion Paper Series, **11**, 金融庁金融研究研修センター, 東京.
- 安田武彦, 高橋徳行, 忽那憲治, 本庄裕司 (2007). 『テキストライフサイクルからみた中小企業論』, 同友館, 東京.
- 柳澤健太郎, 下田啓, 岡田絵理, 清水信宏, 野口雅之 (2007). 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, 249-263.

# Effect of Firm Age in Credit Scoring Model for Loans to the Self-employed

Kenzo Ogi<sup>1</sup>, Yuichi Utsumi<sup>1</sup> and Norio Hibiki<sup>2</sup>

<sup>1</sup>Micro Business and Individual Unit, Japan Finance Corporation

<sup>2</sup>Faculty of Science and Technology, Keio University

A number of Japanese banks have utilized credit scoring models to reduce screening costs by automating part of the screening process. Several types of credit scoring models exist. To calculate the credit scores of small firms with twenty or fewer employees, it is common to utilize the logistic regression model, linked to the correlations between financial indicators and the occurrence of default.

Small firms include not only corporate businesses, but also self-employed businesses. The accuracy of the credit scoring model for the self-employed is not as high as might be expected. The main reason for this is that we lack adequate information regarding financial accounting variables on the balance sheet, such as assets or debts, because self-employed firms are not obligated to prepare a financial statement. We are convinced that it is effective to use explanatory variables associated with assets or debts. However, it is difficult for banks to obtain accurate data, because most self-employed households confuse the owner's private assets with business finances.

Ogi et al. (2016) suggest the use of “firm age” in the credit scoring model for small corporate business, and we have found this to be a statistically significant variable. There are at least three reasons for this. First, for loans to small firms, it is effective to evaluate the private assets of the owner together with the assets of the firm, but accurate information on this subject is not easily available. Second, the default rates of small firms decrease as the firms get older because the private assets of owners are expected to accumulate year by year, and this can help with management activities. Third, in the traditional screening process of loans by experienced bankers, firm age is an important factor for assessing the credit risk to debtors. In this paper, we propose to utilize “firm age” in credit scoring models for self-employed as well as small corporate businesses, and to analyze the correlation between firm age and default rate using a dataset for the more than 680,000 self-employed in Japan from 2007 to 2014. This dataset is owned by the Japan Finance Corporation. We will evaluate the credit scoring model using the accuracy ratio, which is commonly used as a measure.

We find that the default rate can be expressed by the density function of the Weibull distribution or a piecewise-linear function of firm ages with three ranges: up to 3 years, 4–25 years, and 26 or more years. We also show that the accuracy ratios increase by about nine percent at the maximum in both approaches, and that the model can be effectively used in practice.