

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 10 日現在

機関番号：21401

研究種目：基盤研究(C) (一般)

研究期間：2012～2015

課題番号：24530355

研究課題名(和文) 中小零細企業の欠測を考慮した信用リスクモデルと企業情報の有用性に関する実証研究

研究課題名(英文) Analyzing Effects of Missing Data, Financial Data, and Corporate Information in Credit Risk Assessment of Japanese Small and Medium Enterprises

研究代表者

宮本 道子 (Miyamoto, Michiko)

秋田県立大学・システム科学技術学部・教授

研究者番号：30469598

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：1) 中小企業の財務データに対して多重代入法を用いて欠測を補完し多くの情報を活かせるようなデータベースを構築した。2) 特定の財務変数は極端に大きな値、負の値、ゼロ近辺でのクラスターなどが混在し理想的なガウジアン分布とは程遠く、共分散を歪める可能性がある。正規分布への変換方法として、正の値のみならず負の値まで考慮できるneglog変換を行うことで、元々の尺度では隠れていた特徴を際立たせ、変数の分布を改善させた。3) 信用リスク計測に必要な指標を、財務情報に加えて、企業の独自情報を使って推測した。4) 中小企業財務データには外れ値が含まれるが、外れ値を除外せずに、影響を小さくするロバスト推定を行った。

研究成果の概要(英文)：Missing data is found in nearly every use of financial data. In the study of credit risk assessment of small banks, many prior researches have led to information loss by deleting cases with missing values, are left with complete data for all subjects, which do not take advantage of the original dataset. Financial variables may also be non-Gaussian and skewed. The neglog transformation has been introduced to overcome this difficulty. This study analyzes the nature and effects of missingness and skewness of data in credit risk modeling. This study also investigates indicators needed for credit risk measurement for Japanese SMEs, using financial information, as well as corporate information of a large SMEs database and those of a small bank. Lastly, a robust logistic regression extends the conventional logistic model by taking outlier into account, to implement forecast of defaulted firms.

研究分野：財政学・金融論

キーワード：中小零細企業 信用リスク ロジスティック回帰分析 多重代入法 neglog変換 ロバストロジスティック回帰分析

## 1. 研究開始当初の背景

### 中小・零細企業に対するデフォルト確率推計モデルの必要性

わが国の金融機関ではデフォルト確率を正確に予測し、信用リスクに見合うだけの収益を確保するという考え方が基本となっている。一方、新BIS規制の導入とともに適切な内部格付制度の確立とその検証が大きな課題となっている。都市銀行や地方銀行の大手行は、クイックローンやシンジケートローンなどの商品を通じて中小企業向けの貸出拡大を図っている。また地域金融機関同士の競争も厳しい。中小企業向けの担保・保証に過度に依存しない融資の推進等については、スコアリングモデルを活用し短期間で審査を行う融資が多く見られる。それらのスコアリングモデルの構築には、実際の財務データを元に作られた、より安定的で精度の高い計量モデルが必要となる。

このように、中小企業や小規模・零細企業の融資は銀行経営にとって重要分野となってきているが、まだ歴史が浅く、その実態はあまり知られていない。その理由として、中小企業や小規模・零細企業は財務諸表が揃っていない場合が多く、たとえ揃っていても、十分な監査が行われておらず、その信頼性は監査を受けている大企業と比べると劣っているとわざと得ない。また、それら企業は株式を上場しておらず、市場価格や公表されている会社レーティングなど、第三者的に企業経営を判断できる材料が少ない。特に小規模・零細企業の場合は、会社とオーナーである個人との間で資産・損益の混同が発生しやすい問題が指摘されている。

中小・零細企業に関する情報は不透明である一方、資本市場へのアクセスがないため、資金調達は金融機関に頼らざるを得ないというのがアメリカや、わが国の中小企業庁の調査（2005年版中小企業白書参照）

で明らかである。そこで、金融機関の多くは融資決定の際、中小・零細企業の信用リスクの不透明さを払うために、会社から提供された財務情報に加えて、様々な信用調査会社が提供する企業情報を参考にしている。

## 2. 研究の目的

### 中小企業財務大規模データベースに対する欠測値に対する配慮

日本では、1960年代から、企業の財務指標などの情報から信用リスクを計測する研究が数多くなされてきたが、中小企業の財務データは欠損値 (missing data、missing value) を多く含むという特徴がある。より多くの情報が蓄積された大規模データベースに基づく分析は、金融を初め、あらゆる分野においてリスク要因を分析するために重要である。しかし、多くの従来研究では、欠測が多い項目は削除されることによって情報損失を招いており、大規模データの利点を活かしていない研究が多く見られる。また欠測値への考慮をした研究でも、欠測値補間を行うのではなく、欠測の有無や欠測フィールド数を説明変数とすることによって、欠測情報がその企業のデフォルトを予測する上で有益であるとしたものがあつた。通常の欠測値補完の方法は、ランダムな欠測(MCAR or MAR)の場合を前提としている。

中小企業が提出した財務データの集積である大規模データベースではノンランダムな欠測が想定される。そこで、まずは通常のランダムな欠測の補完方法を適用してみて、さらに、このデータベースに合うような方法を考えていかななくてはならない。完全データが整ったあとに中小企業に対するデフォルト推計モデルの構築を行う。

### 財務データ以外の企業情報の利用

金融機関の多くは融資決定の際、中小・

零細企業の信用リスクの不透明さを払うために、会社から提供された財務情報に加えて、自己の調査、また様々な信用調査会社が提供する企業情報を参考にしている。銀行は時間をかけて貸出先との緊密な関係を築きながら、貸出先をモニターし、スクリーニングを掛け、銀行が通常、貸出先の属性や行動など、観察できない情報の非対称性を解決している。特に情報が不透明な小規模ビジネスに対して融資を行うには、そのビジネス内容、会社オーナー、現地市場に関して細部まで詳しい知識が必要となり、金融機関はそれら企業と長年に渡ってビジネス関係を築き上げることが必要となる。これまでこのような企業情報を信用リスクモデルに積極的に取り入れた研究はなかったが、本研究では最終的には、そういった企業情報も組み入れたモデル構築を目指すものである。

### 3. 研究の方法

欠測値に対する従来の処理方法は解析対象とされる複数の変数のどれか一つでも欠測値を持つケースを計算から除外するリストワイズ除去 (list-wise case deletion) 法である。この方法では、欠測値が多いデータは解析に使用できる有効ケース数が少なくなるという欠点がある。また相関係数行列を求める場合などに 2 変数のいずれか (あるいは両方) が欠測値を持つ場合は当該相関係数の計算にのみ用いないペアワイズ除去 (pair-wise case deletion) 法がある。この手法では、可能な限り多くのデータを計算に使用することができるが、相関係数ごとにその計算に関与するケース数が異なるため解釈が困難になる場合がある。

本研究は、欠測値を含んだ中小企業の財務データを用いて信用リスクを推定するためにロジスティック回帰分析を行う際、使用データに欠測値が観測される場合、前後期補完、多重代入法などの統計的な欠測値

処理法の適用について検討した。財務データ特有の構造に配慮しながら、欠測を含む中での変数選択法について考察した。

次の課題として、特定の財務変数は極端に大きな値、負の値、ゼロ近辺でのクラスターなどが混在し理想的なガウジアン分布とは程遠く、共分散を歪める可能性がある。正規分布への変換方法として、Box-Cox(1964)変換モデルがある。また、財務諸表の「利益」のような変数は負の値を取ることがあり、Whittaker et al. (2005)は、負の対数変換、neglog 変換を使う手法を紹介している。neglog 変換を行うことで、もともとの尺度では隠れていた特徴を際立たせ、変数の分布を改善させることができるとしている。そこで、本研究では、財務諸表におけるデータの分布を考慮するためのこれらの手法についても検討し、欠測処理におけるその影響について考察した。

企業情報を信用リスクモデルに取り入れた研究では、財務データと名義尺度または順序尺度で測定されるカテゴリカルデータを含む企業データを分析するため多項ロジスティック回帰分析を行った。

最後に、中小企業財務データには外れ値が含まれるが、外れ値を除外せずに影響を小さくするロバスト推定も行った。誤差分布が裾の広い非対称な分布であってもパラメータの推定値の分散が十分小さくなることが望ましく、適当な初期値からのずれの大きさに応じて重みを小さくすることで実現を目指すロバスト推定を取り入れる。信用リスクで使われるロジスティック推定に対するロバスト・ロジスティック推定分野はまだ新しいが、そこで Bianco and Yohai (BY) 推定量を用いた分析を行った。

### 4. 研究成果

中小企業データベースの2010年における完全データ (Complete Data)、neglog変

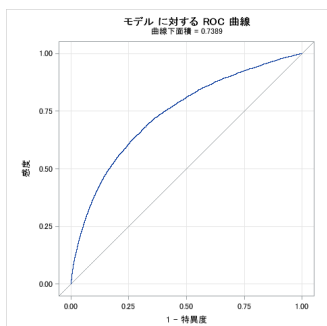
換した完全データ，前後期補完，neglog変換した前後期補完，多重代入法（MI），neglog変換した多重代入法（MI + neglog）を行ったデータを使って、それぞれデフォルト確率の推定を行った（表1参照）。多重代入法は100回行った結果の平均値を示す。多重代入法の回数に関してはCarpenter and Kenward (2013)を参考にした。

その結果，前後期補完，多重代入法，neglog変換後多重代入法とも完全データより分析に使えるデータ数が増え，有意となる変数が増えたことが分かった。多重代入法に至っては、完全データと比べてデータ数は、2倍以上となった。また分布の歪みを修正したMIにより、統計的にも有意な多くのデータの使用が可能となることが示唆された。

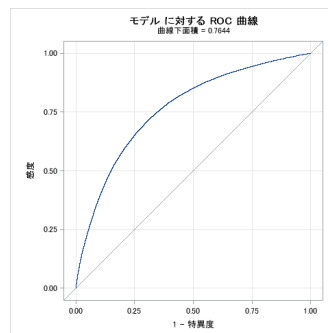
モデルの評価方法として多く使われているAR値は完全データが73.9%，前後期補完が76.4%，多重代入法が75.2%，neglog変換した場合は完全データが73.2%，前後期補完が75.4%，多重代入法が75.2%となった（図1参照）。

表1 分析結果（2010年度）

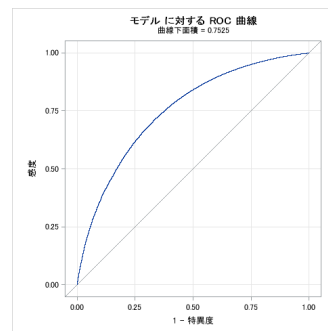
	Complete Data n=417,593	Complete Data +neglog n=417,593	前後期補完 n=419,824	前後期補完 +neglog n=309,391	多重代入法 100回平均 n=917,967	多重代入法 +neglog 100回平均 n=917,967
Intercept	-2.4451***	-8.2808***	-0.4384***	-8.4312***	-6.2895***	-6.6979***
売上高-営業収益	-0.0129	0.5618***	-0.0628***	1.2011***	1.5807***	1.7757***
ROA	0.00867	0.1819***	-0.2568***	-0.00601	0.29***	0.2345***
総資産利益率	-8.26E-08**	-0.1479***	6.39E-07***	-0.2103***	1.2725***	1.5261***
現預金総資産比率	-4.6672***	-2.2387***	-5.269***	-2.9314***	-3.3354***	-3.3478***
売上高利益率	0.000273	0.0335	-0.00011	0.2893**	-1.9873***	-2.4668***
インクレスト-カバレッジ	0.000018***	0.0773***	0.000015***	0.0745***	-0.0738***	0.1118***
負債比率	0.00901***	4.8173***	0.0178***	6.3083***	4.1924***	4.2616***
当座比率	-0.815***	-0.1423***	-0.7879***	-0.1998***	-0.2226***	-0.1939***
流動比率	-0.1863***	-0.294***	-0.0732***	0.3361***	0.5057***	0.504***
株主資本営業利益率	-0.00901	0.000771	0.1274**	-0.2027***	-0.1986***	-0.2127***
売上総利益営業利益率	0.000045	0.2767***	-0.00009**	0.3351***	0.287***	0.2951***
固定長期適合率	-0.8646***	0.1869	-1.315***	-0.6646***	-0.966***	-0.9545***
AR	0.739	0.732	0.764	0.754	0.753	0.752



(1)完全データ



(2)前後期補完



(3)多重代入法

図1 ROC曲線（2010年度）

非財務データを取り入れた小規模な銀行のデータを使った分析では、財務データのほかに、滞納回数、銀行借入回数が統計的に正で有意となった一方、不動産融資の数は小規模企業にとっては負で有意となったが、個人事業主では正で有意となった。中小企業データベースでは、「不動産の所有なし」と「後継者なし」がデフォルトに対して統計的に正で有意となった。このように財務データ以外でも企業データが信用リスク推定には有用な情報となることがわかった。

BY 推定量を使った結果と通常のロジスティック回帰分析の結果を比較したところ、外れ値に配慮した分析が推定精度を改善させる可能性を示した。

【主な参考文献】

G. E. P. Box and D. R. Cox, "An analysis of transformation," J. R. Stat. Soc. B, 26, pp.211-252, 1964.  
 J.R. Carpenter and M.G. Kenward, "Multiple Imputation and its Application," 1st Edition, John Wiley & Son. 2013.  
 J. Whittaker, J. Whitehead and M. Somers, "The neglog transformation

and quantile regression for the analysis of a large credit scoring database,” Applied Statistics 54: pp.863-878. 2005.

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 18 件)

Michiko Miyamoto, Predicting Default for Japanese SMEs with Robust Logistic Regression, International Journal of Economics, Commerce and Research (IJEER), 6(3), 27-34.(2016)

Michiko Miyamoto, Predicting Default for a Small Bank with Robust Logistic Regression, 2016 Seoul International Conference on Sciences and Management, Proceedings, pp.163-171.

Michiko Miyamoto, Event Study of Credit Rating Announcement in the Tokyo Stock Market, Journal of Economics, Business and Management, 4(2), pp.138-143. (Feb. 2016)

Michiko Miyamoto, The Value of Financial and non-Financial Information in Japanese SMEs Risk Assessment, The 3rd Asia-Pacific Conference on Management and Business (APCMB 2015) Proceedings, pp.364-371. (2015)

Michiko Miyamoto, Location Choices of Regional Financial Institutions in the Tohoku Region of Japan, International Journal of Finance and Accounting, 4(1) :71-78 (2015) doi:10.5923/j.ijfa.20150401.07

Michiko Miyamoto, The Value of Non-Financial Information in Japanese SMEs Risk Assessment for a Small Bank, Asia-Pacific Social Science Conference (APSSC), Proceedings, pp.122-128. (2014)

Michiko Miyamoto, Credit Risk Assessment for a Small Bank by Using a Multinomial Logistic Regression Model, International Journal of Finance and Accounting, Volume 3, Number 5 (2014), DOI: 10.5923/j.ijfa.20140305.07

Michiko Miyamoto, Small Business Lending by a Small Bank; An Empirical Study in Japan, 2013 International Symposium on Business and Management (ISBM 2013), Vol. 57, pp.86-91. Proceedings pp.436-447. (2013)

Michiko Miyamoto, Analyzing Effects of Missing Data in Credit Risk Assessment of a Small Bank, 2013 International Conference on Business and Social Sciences (ICBASS), Proceedings, pp.522-528. (2013.6).

Michiko Miyamoto, Investigating Effects of Missingness and Skewness of Data in Credit Risk Assessment of a Small Bank, The Global Symposium on Social Sciences (IBSSS2013), Proceedings, pp.267-277. (2013.11).

〔学会発表〕(計 5 件)

宮本 道子、安藤 雅和、山下 智志 「欠測を考慮したロバストな一般化線形モデルを用いた信用リスクの予測について 中小企業大規模財務データベースにおける考察」2015年度 統計関連学会連合大会, 2015年9月, 於岡山大学

宮本 道子、安藤 雅和、逸見 昌之、山下 智志、高橋 淳一 「欠測値を含む中小企業大規模財務・非財務データを用いた信用リスク評価」2014年度 統計関連学会連合大会2014年9月 於東京大学

宮本 道子、逸見 昌之、安藤 雅和、山下 智志、高橋 淳一 「中小企業大規模財務データベースの分布を考慮した欠測処理について」2013年度 統計関連学会連合大会2013年9月於大阪大学

宮本 道子 「信用格付のイベント・スタディによる研究」日本金融学会 2013 年度春季大会 2013 年 5 月於一橋大学

宮本 道子、逸見 昌之、安藤 雅和  
山下 智志、高橋 淳一 「中小企業大規模データベースに基づく欠測を考慮した信用リスク評価について」2012 年度 統計関連学会連合大会 2012 年 9 月於北海道大学

〔図書〕(計 2 件)

宮本道子、「法と統計学」『法廷のための統計リテラシー-合理的討論の基盤として-』第 4 章, 近代科学社 ISM 統計数理シリーズ, 123-168 頁 2014 年 8 月刊行.

宮本道子、第 7 章 中国・クリーンエネルギーへの挑戦, 「日本のエネルギー政策を考える」, 創成社, 147-172 頁, 2012 年.

〔産業財産権〕

出願状況(計 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

取得状況(計 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

宮本 道子 (MICHIKO MIYAMOTO)

秋田県立大学・経営システム工学科・教授

研究者番号：30469598