FSA Analytical Notes

July 2024 vol.1

Introduction

As financial institutions' business environments and profit structures change, it is important to understand economic and market trends based on data, and to accurately grasp the business conditions of individual financial institutions and also the resilience and vulnerabilities of the financial system as a whole. From this perspective, the Financial Services Agency (FSA) has been focusing on the utilization of granular data, such as transaction-level bank loan data and financial data on individual corporations. Some case examples of data analyses using such granular data are published as a series of reports titled "FSA Analytical Notes."

The "FSA Analytical Notes (2024.7) vol.1" compiles the following three data analyses. The first case, "Analysis of trends of real estate loans by regional banks and study on credit ratings using machine learning" has utilized the transaction-level loan data of regional banks, which is obtained from a novel framework for data collection and management (Common Data Platform)[1] that has been recently launched in a phased manner by the FSA and the Bank of Japan.

While data analysis can provide quantitative and clear results, such results are subject to models and assumptions of the analyses. The data analyses presented here are at the early stage of testing various methods by utilizing currently available granular data while the FSA is still in the process of collecting and accumulating granular data. Therefore, it should be noted that it is necessary to understand the data and model limitations as well as the underlying assumptions and premises and not to draw conclusive conclusions at an early stage when interpreting the results of these analyses. In view of the full-scale data collection via the Common Data Platform from March 2025, the FSA and the Bank of Japan are working to expand the coverage of financial institutions subject to data submission and also examining the accuracy of data. The FSA will continue its efforts to progress

[1] Progress in Common Data Platform and Next Steps
https://www.fsa.go.jp/en/news/2024/20240701/20240701.html

aforementioned initiatives with the cooperation of relevant financial institutions and to consider analytical methods to make effective use of such granular data.

Enhancing the use of data in financial supervision and policy-making is a medium- to long-term agenda. The FSA will continue to build its data analysis capabilities and data infrastructure.

* Unless otherwise noted, the figures and tables in this report were prepared by the FSA.

【Contact】
Macro-financial Stability and Data Strategy Office (datastrategyoffice@fsa.go.jp)

Analysis of Trends of Real Estate Loans by Regional Banks and Study on Credit Ratings using Machine Learning

(Summary)

This paper analyzed real estate loans by regional banks and their credit ratings by using loan-by-loan level data of 62 member banks of the Regional Banks Association of Japan. The utilization of granular data provided better insight into the precise picture of real estate loans from various angles, such as by location of borrowers. In addition, the application of machine learning methods suggests that there is a certain relationship between future credit ratings, and current financial conditions and real estate market conditions. The FSA will continue to carefully monitor the trend of real estate loans. The FSA will also improve its analytical methods and modeling approaches in order to enhance its monitoring capabilities.

I. Introduction

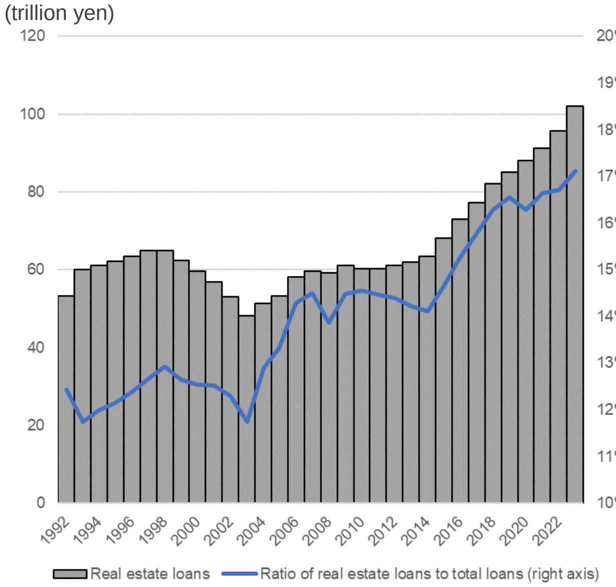
Amid the continuous accommodative financial environment, Japanese banks' real estate loans have been increasing, and the share of real estate loans in total loans has reached a record high[1] (Figure 1). In addition, Japanese real estate prices have been on an upward trend, and some reports have pointed out that the valuation of Japanese real estate prices appears to be relatively high[2] (Figure 2). The Japanese financial system has remained sound overall recently. However, banks' potential risks arising from the real estate sector should be closely monitored given the history of financial crises triggered by real estate bubbles, including the bursting of the bubble economy in the 1990s and the collapse of Lehman Brothers in 2008 stemming from the subprime mortgage crisis. A growing presence of real estate in the Japanese financial system and a sense of caution over the real estate market conditions also justify the need for close monitoring. In order to identify risks in a forward-

[1] As shown in Figure 1, loans to the real estate industry accounted for approximately 17% of total loans at all domestic banks as of end-2023. The corresponding share for regional banks, which are analyzed in the following sections, is also approximately 17%.
[2] For example, Financial System Report published by the Bank of Japan (April 2024).

looking manner, it would be useful to analyze how the current real estate market conditions will affect future bank credit risks.

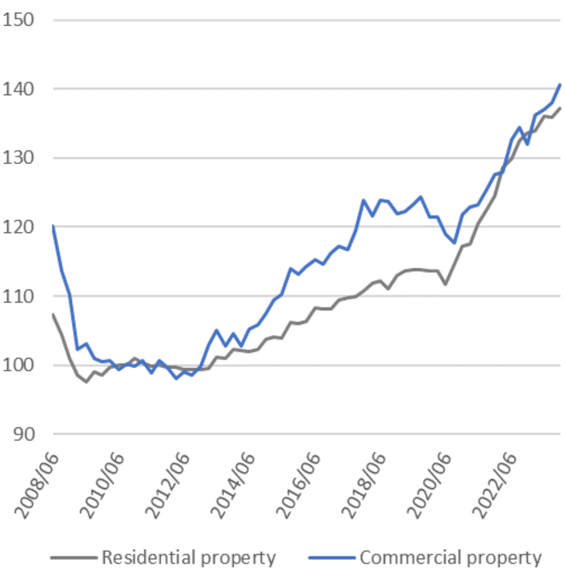
In this paper, the data used are the data from 62 member banks of the Regional Banks Association of Japan (hereinafter, “regional banks”), including loan-by-loan level data obtained through a new data collection and management framework (Common Data Platform) that the FSA and the Bank of Japan have recently launched in phases[3]. The analysis consists of two main parts. Firstly, the characteristics of real estate loans by regional banks are described by aggregating the granular data by type of industry and region of borrowers. Secondly, the relationship between borrowers' financial conditions and real estate market trends and credit ratings are analyzed by using machine learning and other techniques. The purpose of this paper is to deepen our understanding of banks' real estate risks and to consider effective ways to utilize granular data to enhance our monitoring.[4]

Figure 1: Real estate loans of all Japanese banks[5]



Source: Bank of Japan

Figure 2: Real Estate Price Index



Source: MLIT

[3] This analysis focuses on regional banks, as regional banks’ data submission through the Common Data Platform precedes other banks. However, the fact that loans to the real estate industry are increasing is also true for other types of banks, such as major banks and member banks of the Second Association of Regional Banks.

[4] Unless otherwise noted, the latest data and figures used in this report are as of the end of September 2023.

[5] The definition of real estate in Figure 1 is "real estate business (industry code 50)" used in the Bank of Japan statistics, which includes " House and room lending by households” (industry code 96) and "Special purpose companies for real estate” (industry code 89).

II. Current Overview of Real Estate Loans by Regional Banks

The two datasets mainly used in this analysis are: borrower-by-borrower level loan data, in which the names of borrowers and lender banks are anonymized (Dataset 1) and loan-by-loan level data, which has been newly acquired through the Common Data Platform (Dataset 2). The former has been accumulated quarterly since the period ending in March 2004, enabling a long-term time series analysis.[6] The latter has only been available as of end September 2023, i.e., no historical data has been accumulated, however, its high granularity enables analysis from various angles, such as by detailed industry classification and by borrower region.[7]

In this section, firstly, the historical trend of loans to the real estate transaction sector and to the real estate leasing and management sector are examined by using Dataset 1 (sub-section 1). Secondly, more detailed analyses are made by using Dataset 2, which also covers loans to individuals and loans to SPCs (non-recourse loans)[8] which have been increasing in recent years (sub-section 2). Finally, the impact of interest rates hike on the interest payment capacity of borrowers belonging to real estate industry is estimated by using both datasets (sub-section 3).

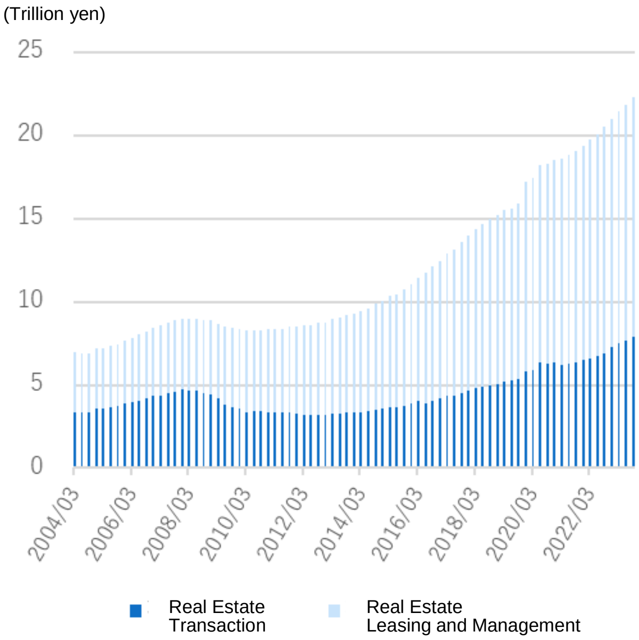
1. Historical Trend

Figures 3 and 4 show the trend in outstanding amount of real estate loans[9] by regional banks over the past 20 years. The outstanding amount of real estate loans as of the end of September 2023 was approximately 8 trillion yen for the real estate transaction sector and approximately 14 trillion yen for the real estate leasing and management sector. The growth rate of the loans to real estate transaction sector was negative immediately after the global financial crisis (GFC), but has been positive since March 2013. On the other hand, the growth rate of the loans to real estate leasing and management sector has been positive since June 2005.

---

[6] Dataset 1 is extracted and processed from a database possessed by the Regional Banks Association of Japan. The database covers mainly loans to Japanese companies, while retail (e.g. mortgage loans) and individuals (e.g. apartment loans) are not covered.
[7] It should be noted that the scope of Dataset 2 differs from that of Dataset 1, as Dataset 2 includes loans to individuals.
[8] SPCs are special purpose companies established for the purpose of acquiring real estate. Generally, in the case of a loan to an SPC, repayment is funded by cash flows generated by the underlying real estate and repayment is limited to the extent of the collateral. It is estimated that some part of loans to SPCs are also covered in the Dataset 1 as real estate transaction industry or real estate management and leasing industry, while they cannot be extracted.
[9] As explained in the footnote 6, Dataset 1 only covers loans to corporates and does not cover individuals. Therefore, it should be noted that the outstanding amounts of loans shown in Figure 1 and Figure 3 are not equivalent.

Figure 3[10]: Outstanding amount of loans by sector



Source: Dataset 1

Figure 4: Growth rate of loans by sector (YoY)

Source: Dataset 1

Figures 5-8 show the financial developments in each real estate sector (median). ROA[11] has been solid for both sectors compared to all industries, remaining at around 2% even during the GFC and the COVID-19 crisis. Interest expense[12] has declined over time, reflecting a decline in market interest rates. Debt ratio[13] has remained high compared to all industries due to the nature of the real estate business in that real estate companies normally raise funds through bank loans for the acquisition of real estate for sale or lease. By sector, the real estate transaction sector has seen fluctuations in ROA in response to economic fluctuations, as is the case for all industries. On the other hand, the real estate leasing and management sector has seen little fluctuation in ROA and a low cash and deposits ratio.[14] This reflects the nature of the sector in which large fluctuations in rent and administration fee income are unlikely to occur (i.e., stable income can be expected for the real estate leasing and management sector).

[10] It should be noted that some data discontinuities are observed due to bank mergers. For example, data for Kansai Urban bank was added in 2019 (Kansai Urban bank was a member of the Second Association of Regional Banks so it was not covered in Dataset 1 before 2019).
[11] ROA = (Operating income + Interest and dividends income) / Total assets
[12] Interest expense = Interest and discount expenses / (Short-term loans payable + Long-term loans payable)
[13] Debt ratio = (Short-term loans payable + Long-term loans payable) / Total assets
[14] Cash and deposit ratio = Cash and deposit / Total assets

Figure 5: ROA

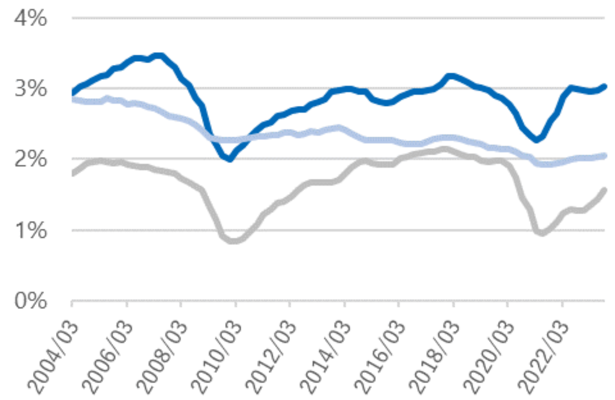


Figure 6: Interest expense

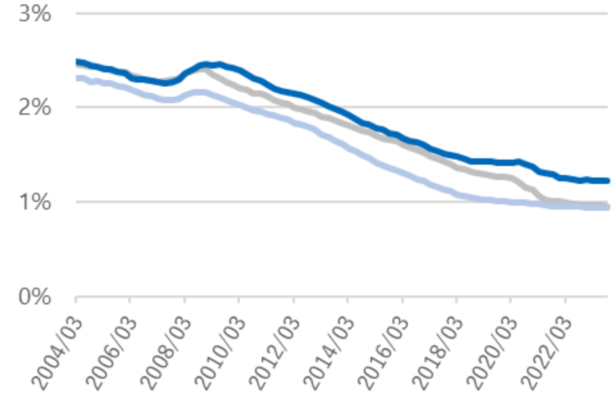


Figure 7: Debt Ratio

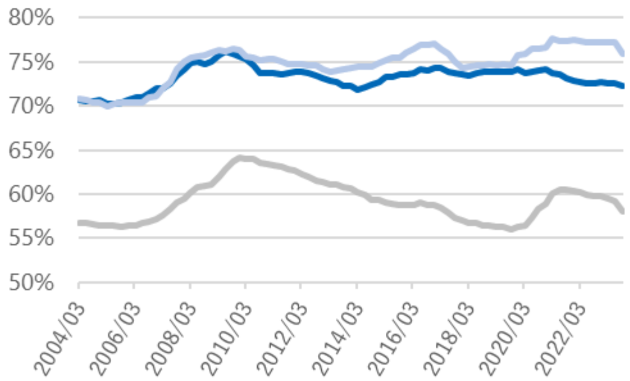
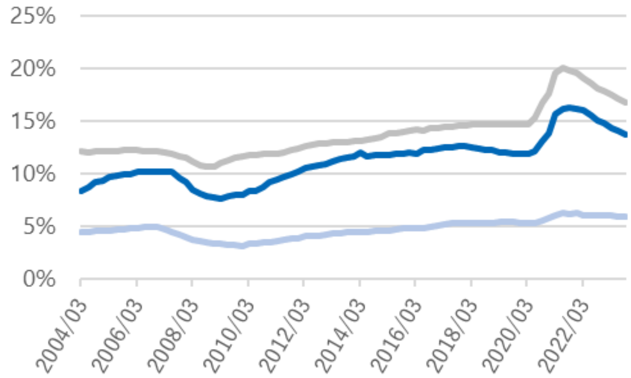


Figure 8: Cash and Deposit Ratio



Source: Dataset 1

Figures 9 and 10 show the trend in the NPL ratio and the conservation ratio.[15] At the time of the GFC, the outstanding NPL amount rapidly increased (worsened) in the real estate transaction sector, but there were no significant changes in the real estate leasing and management sector. Since then, the NPL ratio has been on a declining trend for both sectors, and no significant deterioration has been observed even during the COVID-19 crisis. The conservation ratio has been on a declining trend in all industries except for after the GFC and during the COVID-19 crisis. This may be partly due to increased efforts of providing loans that do not excessively depend on collateral and guarantees. On the other hand, the conservation ratio in the real estate industry has remained unchanged or increased during this period. This may be driven by a rise in real estate prices as real estate loans generally use

[15] The conservation ratio is the ratio of the amount of coverage provided through collateral and guarantees to the total loan balance, and does not take into account the coverage provided through provisions.

real estate as collateral (see Box 1).

From the above analysis, while the outstanding amount of real estate loans has been increasing, there has been no significant change in the credit risk of borrowers as a whole recently.

Figure 9: NPL ratio

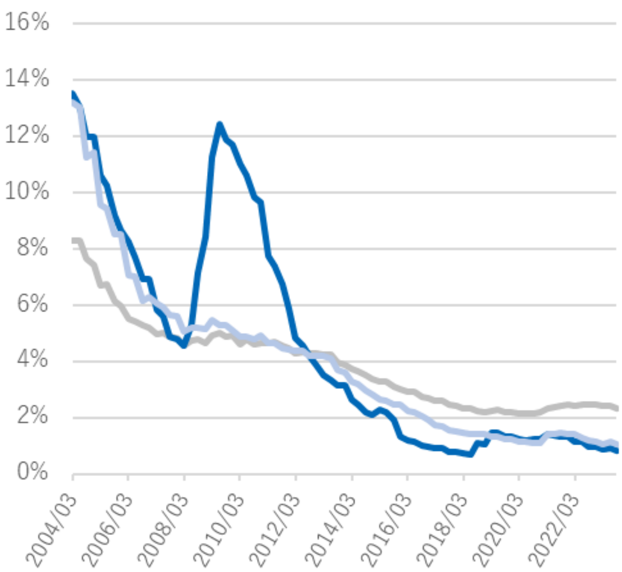
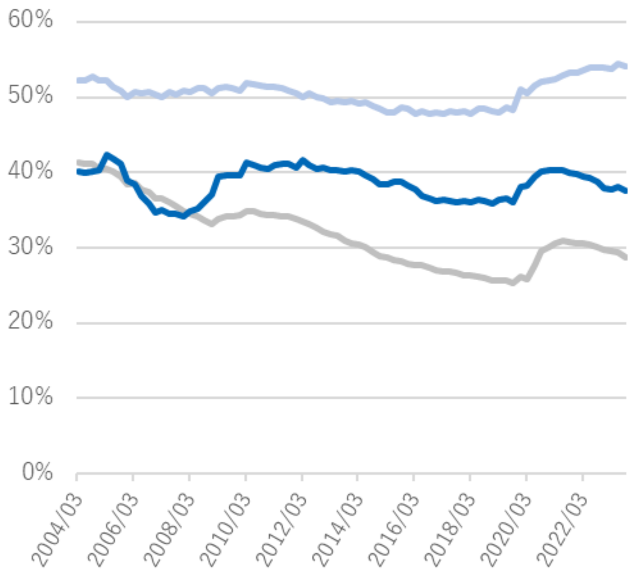


Figure 10: Conservation rate



All industries
Real estate transaction sector
Real estate leasing and management sector

Source: Dataset 1

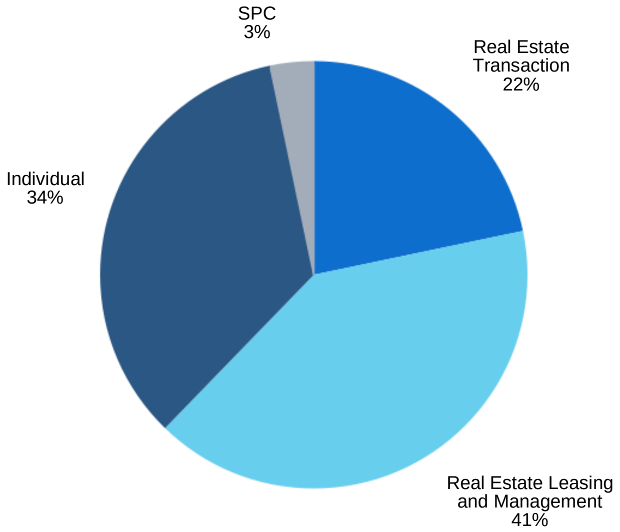
2. Current Trend

In this sub-section, more detailed analysis on current trend is conducted by using Dataset 2, which has been obtained through the Common Data Platform. Real estate loans are classified into corporate loans (real estate trading sector and real estate leasing and management sector), loans to individuals (e.g., apartment leasing business), and loans to SPCs (non-recourse loans).[16] At regional banks, loans to individuals account for approximately 34% of all real estate loans, and loans to SPCs account for approximately 3%. However, loans to SPCs have increased rapidly in recent years due to the maturation of the real estate securitization market and the solid real estate market condition

[16] A loan is classified as a loan to an individual or a SPC if: (a) the loan is defined as “House and room lending by households” (sector No. 96) or "Special purpose companies for real estate” (sector No. 89) under the Bank of Japan statistics classification or (b) the loan could be presumed to be a loan to an individual or a SPC based on the name of the borrower or the type of borrower.

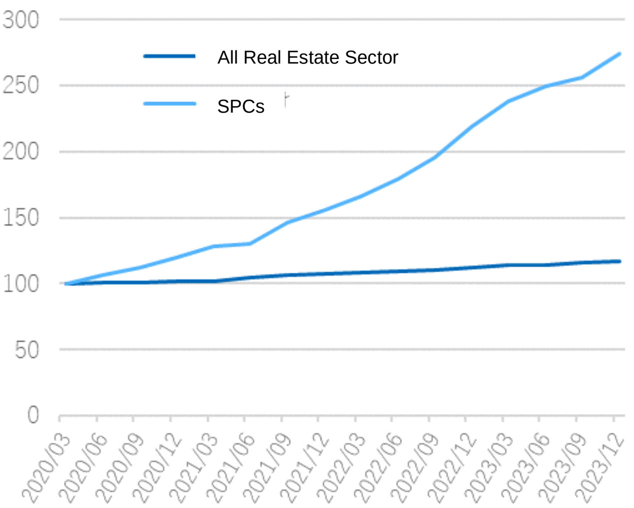
(Figures 11 and 12).

Figure 11: Composition of loans by borrower category (Outstanding amount basis; as of the end September 2023)



Source: Dataset 2

Figure 12[17]: Trend of loans to SPCs (March 2020 = 100)



Source: FSA

Figure 13 shows the composition of loans outstanding by type of borrower for each region in which banks' head offices are located.[18] In all regions, loans to the real estate leasing and management sector or individuals has a relatively large proportion, while loans to SPCs tend to be scarce. Banks in the Chubu and Kyushu/Okinawa regions have relatively large amounts of loans to SPCs compared to the other regions, however, when looking at individual bank data, this is due to the influence of some specific banks with large amounts of loans to SPCs. Therefore, at present, regional banks do not seem to be actively extending loans to SPCs. However, given the recent trend, it is possible that regional banks with small amounts of loans to SPCs at this moment will step up their lending to SPCs in the future.

While loans to SPCs could meet diverse funding needs and thus contribute to financial intermediation, they tend to be more directly affected by real estate market conditions than corporate loans. For example, repayment sources are limited to cash flows generated from the underlying real

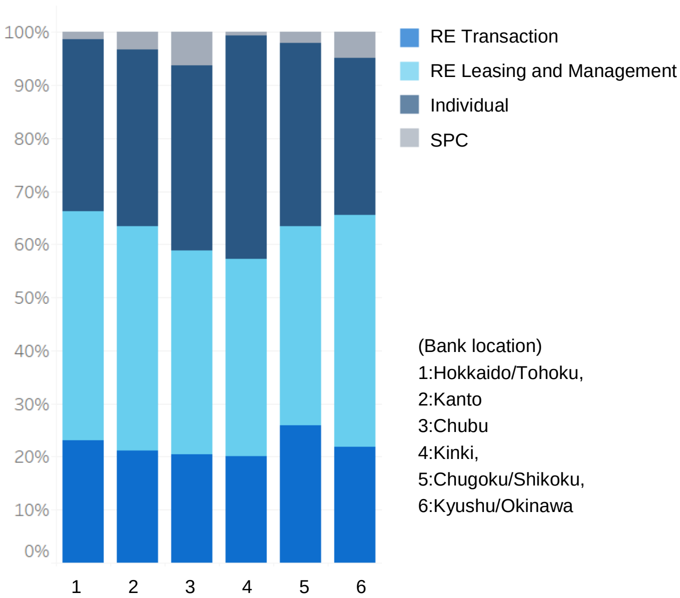
[17] Figure 12 uses a dataset that is different from Dataset 1 and Dataset 2. It should be noted that the coverage of the real estate loans and/or SPCs loans could be different from Dataset 1 and Dataset 2.

[18] In Figures 13 and 14, the geographical regions in which the head offices of banks are located are as follows: Hokkaido/Tohoku, Kanto, Chubu, Kinki, Chugoku and Shikoku, and Kyushu/Okinawa.

estate. Therefore, it is necessary for banks to develop a risk management framework that is different from the regular corporate credit risk management, such as the ability to assess the value of the real estate itself and the outlook for the real estate market.

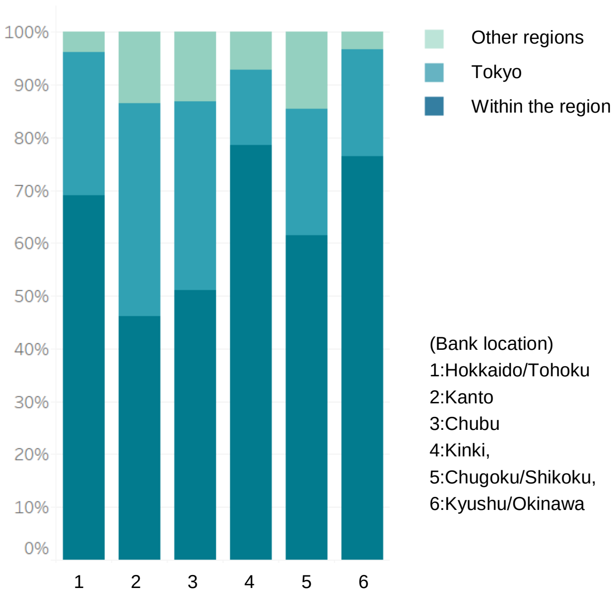
Figure 14 shows the proportion of loans outstanding by location of borrowers[19] for each region in which banks' head offices are located. While intra-regional loans account for the majority of loans, loans extended to Tokyo account for 14-36% at banks located outside the Kanto region (in which Tokyo is located).

Figure 13: Proportion of loans by borrower category



Source: Dataset 2

Figure 14: Proportion of loans by location of borrowers



Source: Dataset 2

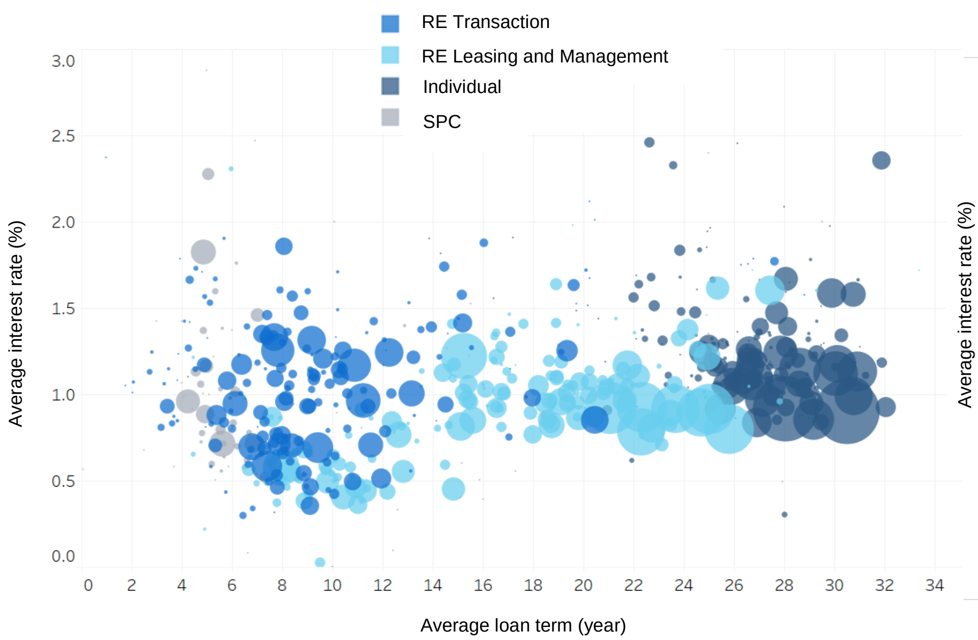
Figure 15 shows the distribution of loan terms (contract period) and loan interest rates. It is observed that the loan terms tend to be longer in the order of individuals, real estate leasing and management sector, real estate transaction sector, and SPCs. The fact that the loan terms of individuals and real estate leasing and management are longer is consistent with their business model in which the initial costs of property acquisition are recovered over the long term through rental income. As for loan interest rates, it is generally considered that the longer the contract period is, the higher the loan interest rate becomes. However, no remarkable proportional relationship is confirmed from

[19] In Figure 14, the geographical region of borrowers are classified into 3 categories: (1) the same region as that of banks' head offices ("within the region"), (2) Tokyo, and (3) other regions. Loans to Tokyo by banks headquartered in Tokyo are classified into "Tokyo."

the distribution in this study. This is because loan interest rates are also affected by factors other than the contract period, such as the interest rate environment at the time of contracting[20].

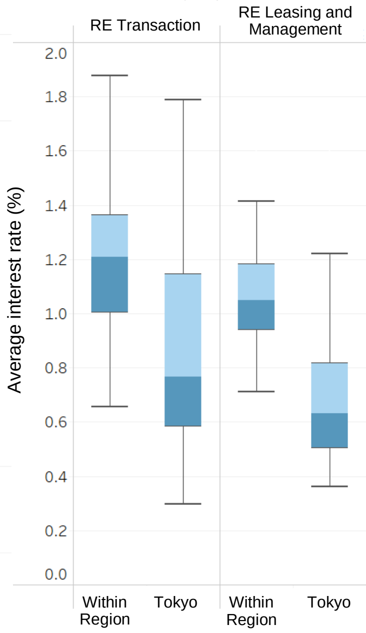
On the other hand, when looking at interest rates by region, it is observed that interest rates on loans to the Tokyo-located real estate transaction sector and real estate leasing and management sector tend to be lower than those to within region[21] (Figure 16). This may be due to the intense competition and the high number of borrowers with high creditworthiness in the Tokyo area. Under the accommodative financial environment, banks seem to have expanded their loans outside of their home regions, such as Tokyo where there is demand for funds.[22] This suggests that regional banks located outside Tokyo area are also linked to the real estate market condition in the Tokyo metropolitan area, whose property prices have been pointed out to be relatively expensive by some research in recent years.[23]

Figure 15: Distribution of average loan term and average interest rate



Source: Dataset 2

Figure 16: Distribution of average Loan interest rates (by Industry and by Region)



Source: Dataset 2

---

[20] Another factor behind this could be the difference in the creditworthiness of borrowers (generally, the higher the creditworthiness, the easier it is to borrow over a long period of time).

[21] Looking at the Interest rates trend closely by taking into account the contract period, loans to borrowers located in Tokyo still tend to be lower than those to within region.

[22] According to Dataset 2, while the proportion of loans to the Tokyo metropolitan area (loans to institutions located in Tokyo) in the total loans of regional banks is approximately 23% for all industries, it is high at approximately 46% for real estate transactions and approximately 32% for real estate leasing and administration (excluding loans to individuals).

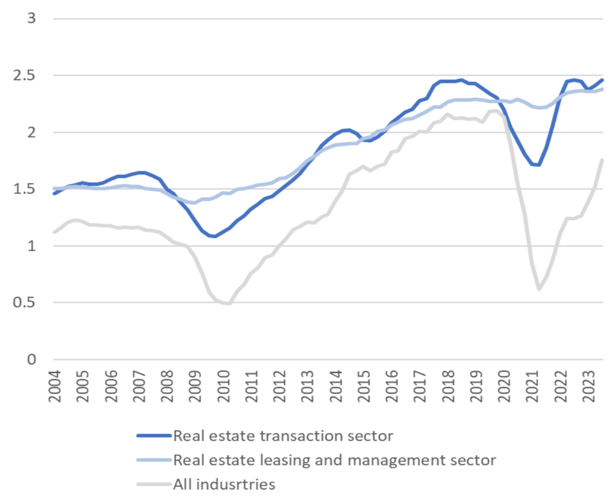
[23] A large amount of loans to SPCs are made to Tokyo. It should be noted, however, that even if the registered location of an SPC is in Tokyo, the location of the property to be acquired may be elsewhere.

3. Impact of Rise in Borrowing Rates on Interest Payment Capacity

With market interest rates on an upward trend as a result of the Bank of Japan's change in monetary policy, attention is paid to the impact on corporate finances due to an increase in borrowing rates. In this sub-section, a simplified calculation is made focusing on the interest coverage ratio (ICR) [24], which is an indicator of firms' interest payment capacity, in order to obtain an indication of the impact of a rise in borrowing rates on corporate finances.[25]

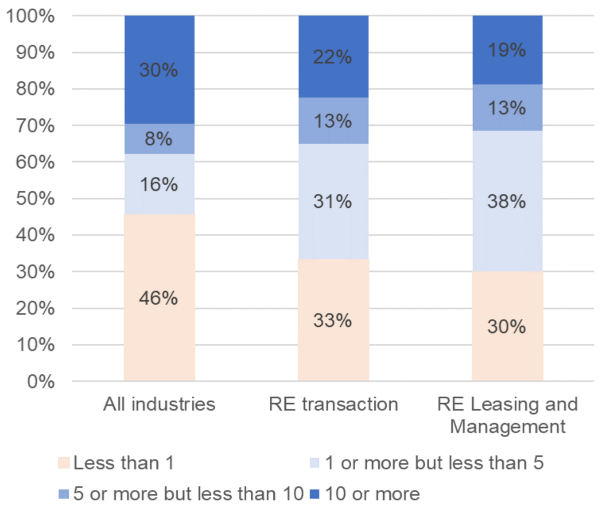
The ICR is negatively correlated with corporate defaults, and in particular, the proportion of firms that default tends to increase when the ICR drops below zero. The median ICR for the real estate sector has been on an improvement trend over the long term, exceeding all industries, although some drops were seen after the GFC and the COVID-19 pandemic (Figure 17). The distribution of the ICR as of the end of September 2023 shows that the proportion of firms with ICRs of 1 or above (i.e., with sufficient interest payment capacity) is high in both the real estate transaction and real estate leasing and management sectors, compared to all industries. However, it should be noted that the proportion of firms with ICRs of 1-5 is high at 30-40%, meaning that they might be vulnerable in the sense that ICRs may fall below 1 due to interest rate hikes and other factors (Figure 18).[26]

Chart 17: ICR (median)



Source: Dataset 1

Figure 18: ICR Distribution



Source: Dataset 1

---

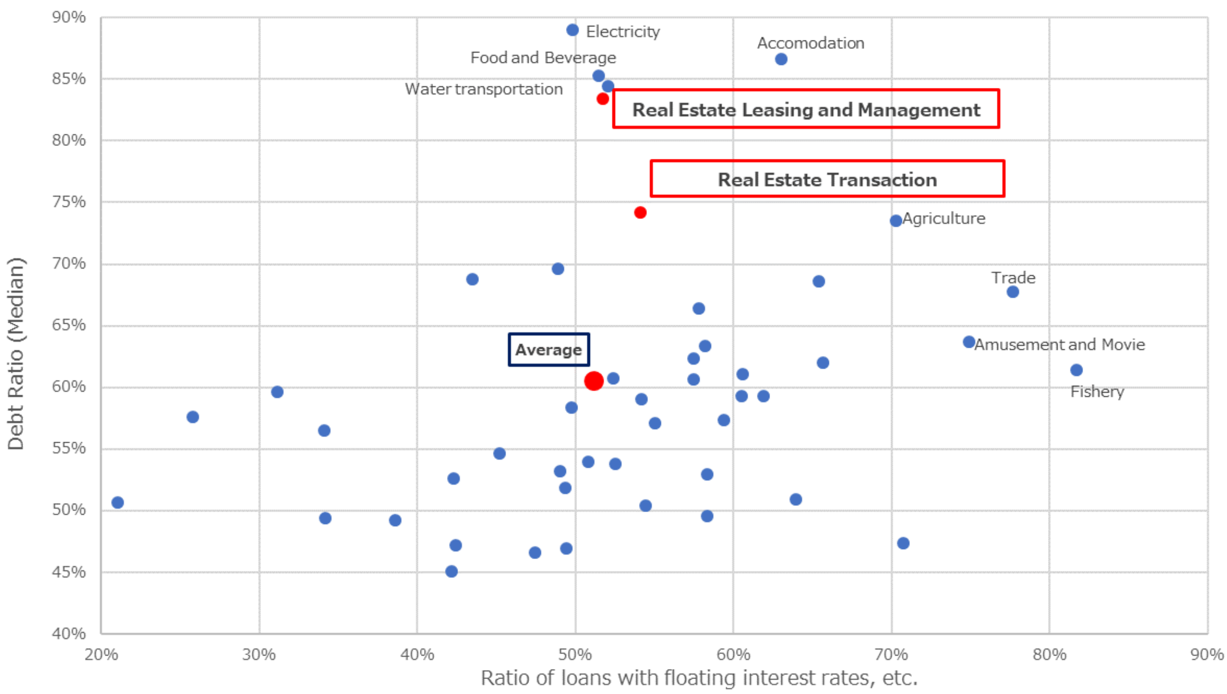
[24] ICR = (Operating income + Interest and dividend income) / Interest and discount expenses.
[25] Due to data limitation, the classification “individuals" and "SPCs" which were added in the previous section is not extracted in this section. It is estimated that some of the SPCs confirmed in the previous sub-section are included in the real estate transaction sector or real estate leasing and management sector in this sub-section.
[26] See FSA Analytical Notes (2023.6), Analysis of Credit Risk in Bank Loans.

Loans with floating interest rates and loans with fixed interest rates but with short residual maturities are more likely to be affected by interest rate increases in short periods. As such, looking at the proportion of loans with floating interest rates and loans with fixed interest rates with residual maturities of one year or less (hereafter, "floating or short-maturity loans") to total loans outstanding, it is observed that the proportion of floating or short-maturity loans in both the real estate transaction sector and real estate leasing and management sector is 50-60%, which is not very different from all industries, but the debt ratio (median) is as high as 70% or more (Figure 19).

In the real estate sector, the ratio of companies with an ICR of less than 1 to the total number of companies (hereinafter defined as "companies with an excess interest payment ratio") remains at a relatively low level. On the other hand, under the assumption that interest rates on floating or short-maturity loans uniformly increase by a certain amount (parallel shift), the increase in the ratio is larger in the real estate sector, particularly in the real estate leasing and management sector than in all industries (Figure 20).

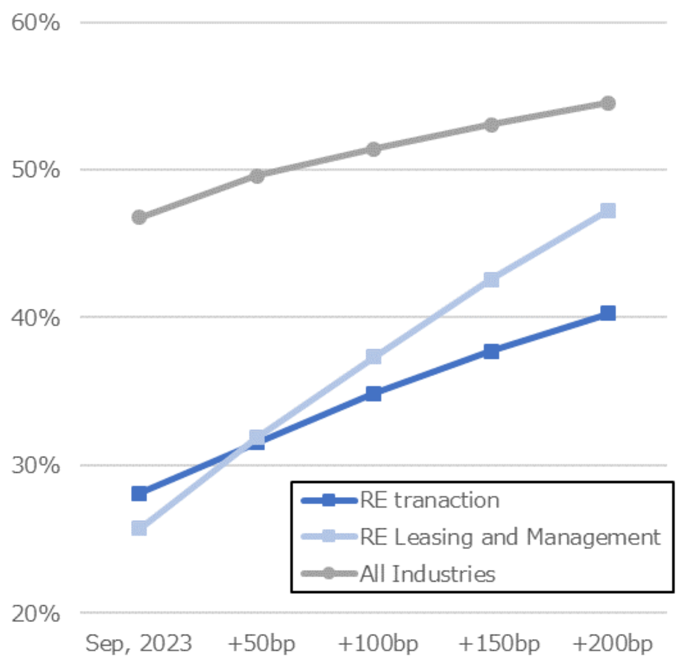
Based on the above, the real estate sector is considered to be relatively susceptible to interest rate increases.

Figure 19: Debt Ratio and Ratio of Floating or Short-maturity Loans



Source: Dataset 2

Figure 20: Changes in companies with an excess interest payment ratio under a certain rise in borrowing rates



Source: Calculated based on Dataset 2

In this analysis, only borrowing rates are assumed to rise instantly, while other variables are assumed to remain unchanged. Therefore, it should be noted that this estimate may not necessarily reflect the real macroeconomic environment, for example, business performance may deteriorate along with the borrowing rate rise, or otherwise, business performance is expected to improve over time along with the rate rise (e.g., when rental income increases in line with a rise in interest rates). In particular, as shown in Figure 17, the ICR for the real estate leasing and management sector has remained stable even during the past interest rate rise period and economic downturns. This is because real estate demand itself is unlikely to decline much despite changes in the economic environment.

In any case, the real estate sector is considered to be more vulnerable to interest rate increases than other sectors from a financial perspective, and even a small increase in interest rates could lead to a significant deterioration in the financial condition of the sector depending on the macroeconomic environment. Therefore, the FSA will continue to closely monitor developments in the real estate sector.

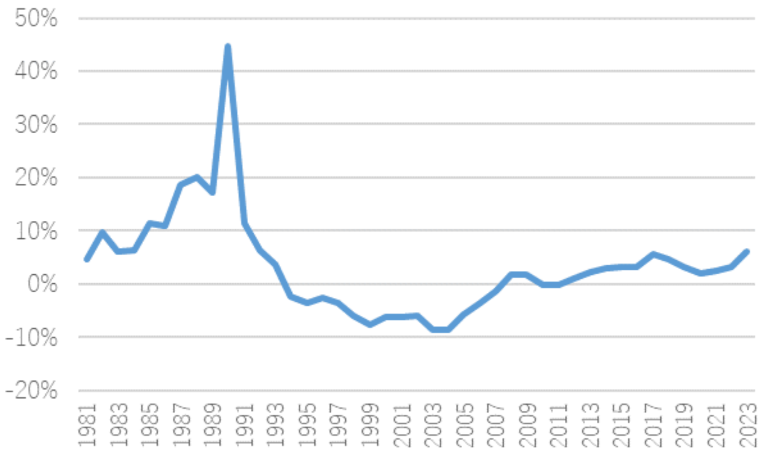
Box 1: Use of real estate as collateral

As real estate prices have risen and the outstanding loan amounts to the real estate sector have increased, loans backed by real estate have also increased in recent years (Figure 21). To understand the big picture of the use of real estate collateral, Figure 22 shows the use of real estate collateral by size of borrower and by borrower sector, using detailed loan-by-loan level data obtained from Common Data Platform (Dataset 2). The proportion of borrowers using real estate collateral was highest at medium-sized enterprises, followed by small enterprises, and then large enterprises. The proportion of large enterprises using real estate collateral was 20% or less for sectors except real estate. The reason behind this could be that large enterprises have high credit ratings so they do not need to rely on real estate collateral when borrowing, and that small enterprises possess less real estate available for collateral.

Figure 23 shows the proportion of the amount covered by real estate collateral against total outstanding loans for borrowers who use real estate collateral. The coverage ratio also tended to be higher at medium-sized enterprises than small enterprises or large enterprises. Approximately 60% of the real estate sector of medium-sized enterprises and small enterprises are covered by real estate collateral, whereas the coverage ratio for other sizes and sectors is less than 30%.

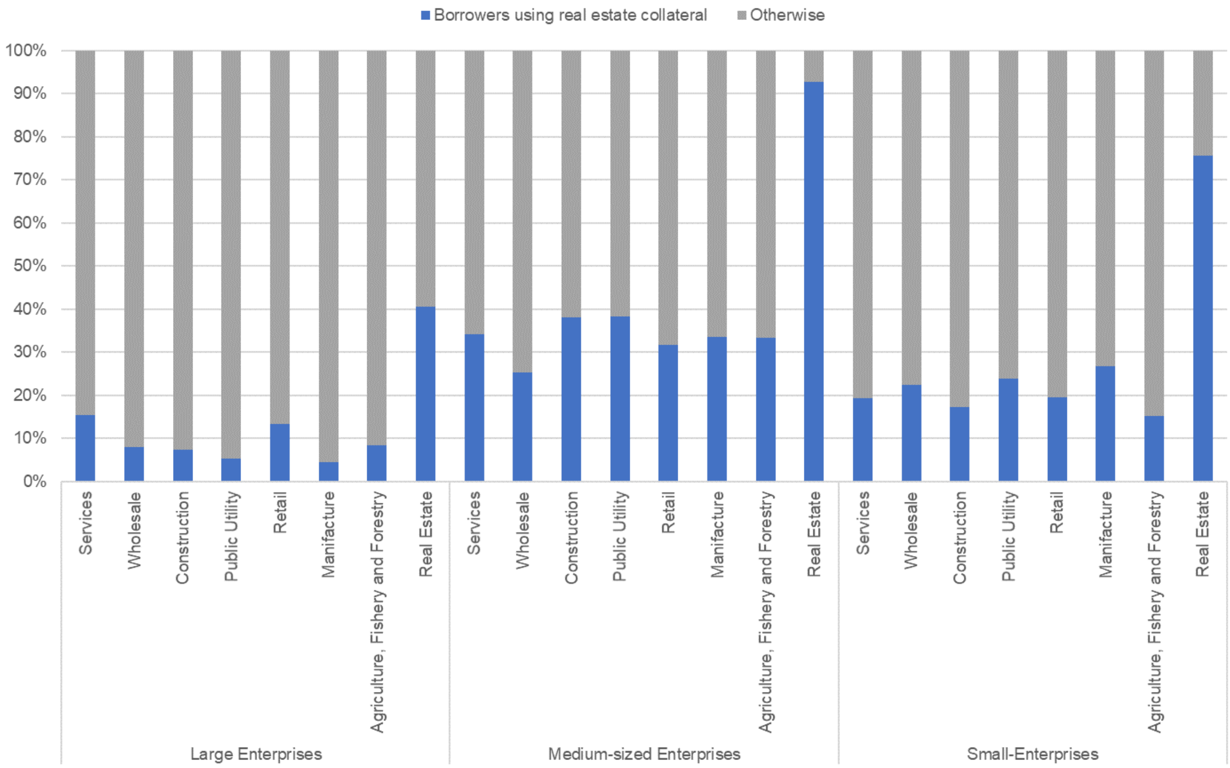
This indicates that many borrowers in the real estate sector use real estate collateral and cover a large part of the credit exposure, which is consistent with the fact that real estate companies generally use real estate collateral when borrowing for the acquisition of real estate for sale or lease. On the other hand, although it can be inferred that other sectors are not overly dependent on real estate collateral as seen in the past bubble period, a deterioration in the real estate market could have an impact on companies in other sectors through a transmission channel of a decline in the value of real estate collateral. Therefore, the FSA will continue deep-diving to further understand the use of real estate collateral by utilizing detailed data.

Figure 21. Growth in loans using real estate collateral (YoY)



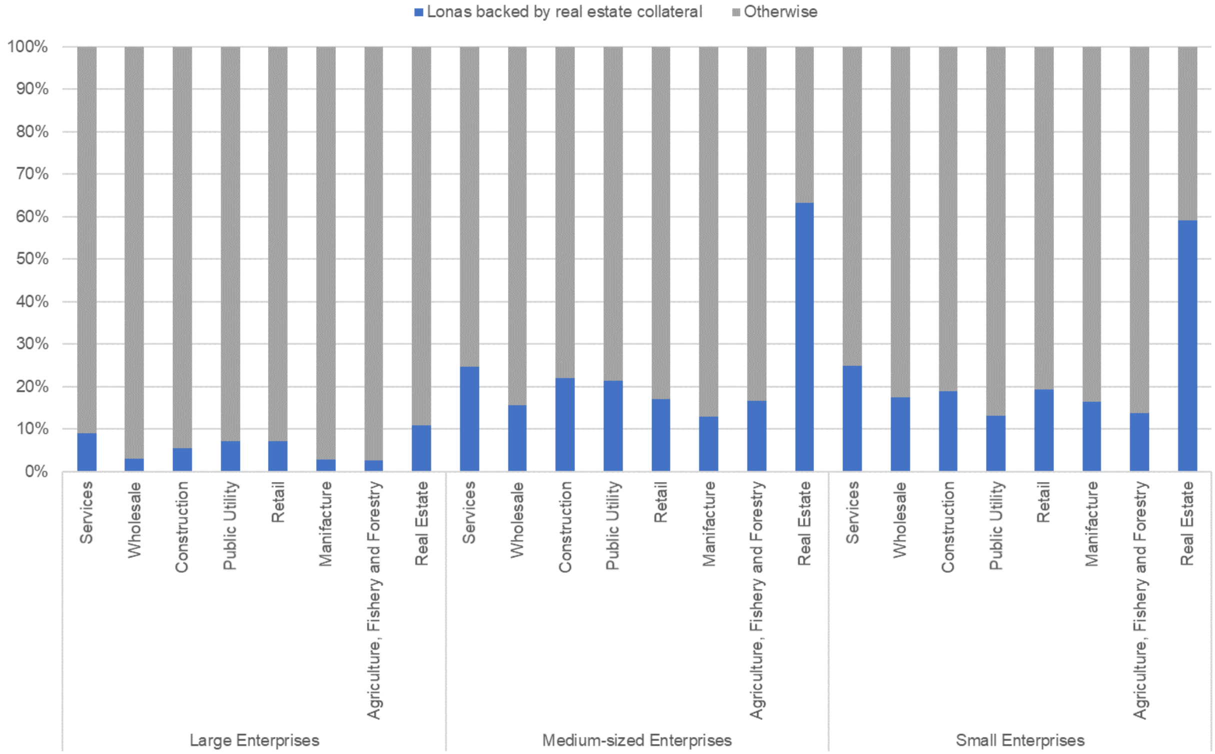
Source: Bank of Japan

Figure 22: Breakdown of borrowers using real estate collateral



Source: Dataset 2

Figure 23. Real estate collateral coverage ratio



Source: Dataset 2

III. Relationship between Changes in Credit Ratings, Financial Indicators, and Real Estate Market Conditions

As confirmed in the previous chapter, the NPL ratio of real estate loans has remained at a low level recently. On the other hand, the credit risk of real estate loans is likely to be affected by the real estate market and the financial conditions of real estate firms reflecting such market conditions, as seen in the substantial increase in the NPL ratio in the real estate transaction sector during the GFC period. However, the real estate market has multiple layers, for example, demand for office buildings and hotels stagnated while the housing market was booming due to the spread of remote working during the COVID-19 crisis. As such, it is not a straightforward matter to find a clear relationship between the real estate market, financial conditions, and credit risk of real estate loans. To deepen the understanding of real estate loans, the following step-by-step analysis is conducted in this chapter: first, construct a model that predicts the credit risk (credit rating of a borrower) of real estate loans using a machine learning technique (Section 1), and then attempt to interpret the model to make it human-readable (Section 2). Machine learning can build a model with a complex structure and incorporate non-linearity between variables, making it possible to consider relationships that are difficult for humans to detect.

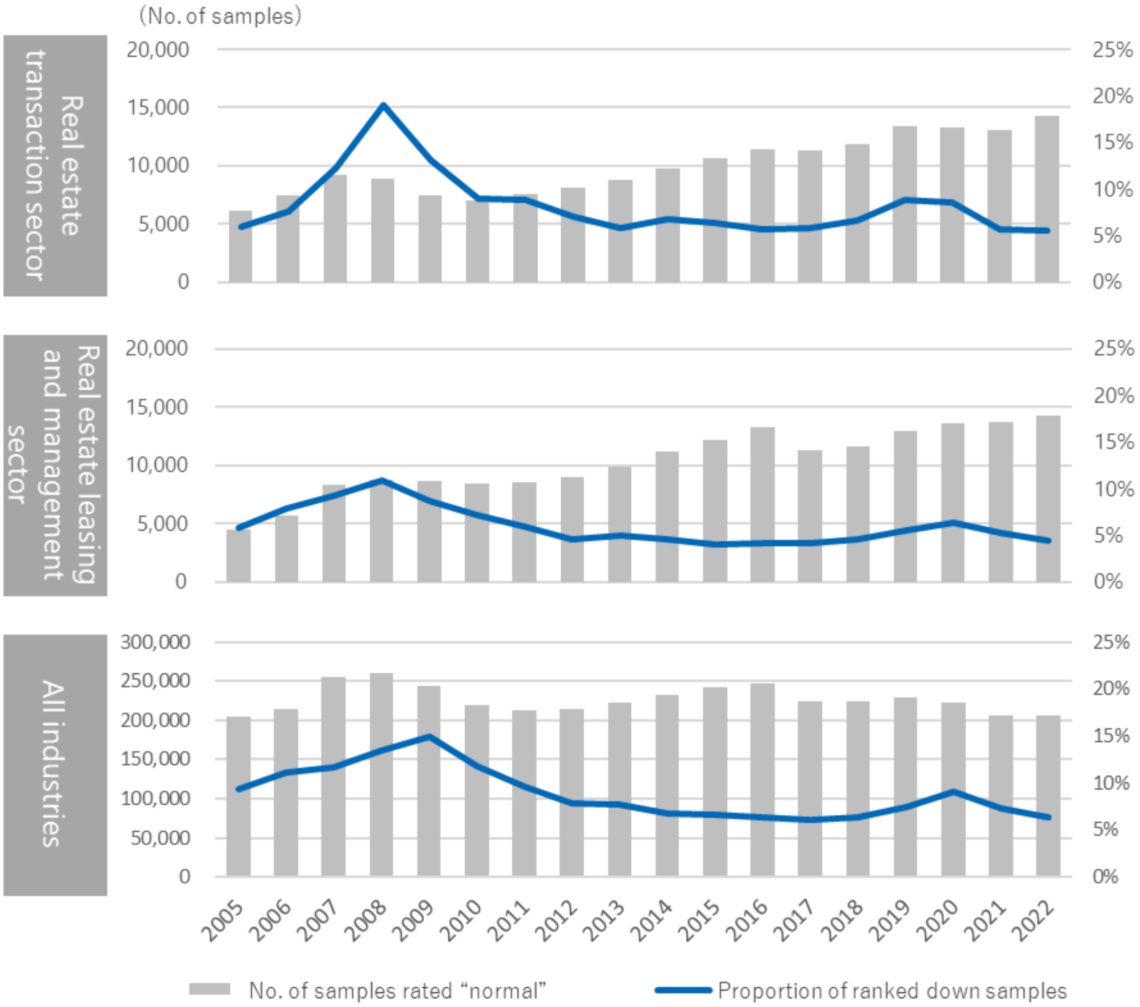
1. Overview of Machine Learning Models and Predictive Results

The machine learning model used in this analysis is XGBoost (eXtreme Gradient Boosting), which improves accuracy for prediction by combining multiple low-accuracy decision trees. XGBoost has been accepted in numerous surveys and research in the field of machine learning and is known for its high performance. The list of an objective variable and explanatory variables input into the model are shown in Figure 24. Of all available data, 75% is allocated as training data and the remaining 25% as test data. The model estimates the probability that the credit rating of a borrower will deteriorate from "normal" to "needs attention" or lower (hereinafter referred to as the "probability of downgrading") within one year from the record date by using data including Dataset 1 as well as data on corporate financial indicators, corporate characteristics and real estate market conditions. Figure 25 shows the sample size of borrowers rated "normal" and the proportion of borrowers who actually were downgraded in each fiscal year from 2005 to 2022.

Figure 24: List of Objective Variable and Explanatory Variables

Objective variable	• Probability of rank down	Probability that a credit rating of a borrower will deteriorate from "normal" to "needs attention" or within one year from the record date	
Explanatory variables	• ROA	(Operating income + Interest and dividends income)/Total asset	Historical loan data obtained from CRITS (dataset 1)
	• Interest Expense	Interest and discount expenses / (Short-term loans payable + Long-term loans payable)	
	• Debt ratio	(Short-term loans payable + Long-term loans payable) / Total asset	
	• Cash and Deposit ratio	Cash and deposit / Total asset	
	• Size	Size of the borrower (large, medium, small)	
	• Land price	Rate of change in land price (nationwide, all land usage categories, YoY)	MLIT
	• Inventory DI	Inventory DI for real estate sector	BoJ "Tankan"
	• No. of housing starts	Logarithm transformation of total number of housing starts	MLIT
	• Vacancy rate	Average office vacancy rate (Tokyo)	Miki Shoji Co.,ltd
	• No. of transaction	Rate of change in number of real estate transactions (YoY)	Japan Real Estate Institute
	• Floor area	Logarithm transformation of floor area for total number of building starts	MLIT
	• Stock price	Rate of change in Nikkei 225 stock price index (QoQ)	Bloomberg

Figure 25: Number of samples rated “normal” and proportion of downgraded borrowers



Source: Dataset 1

The performance of the learned model is shown in Figure 26. Since the precisions exceeded 70% in the test data, this model is considered to have accomplished a certain level of performance in terms of predicting future downgrading; however, high recalls could not be achieved. In general, precision and recall have a trade-off relationship, i.e., if the performance of one is pursued, then the performance of the other deteriorates. In this analysis, precision is prioritized for model construction assuming that there is a need to detect a borrower with a high probability of downgrading. On the other hand, depending on the purpose of the analysis, there may be some cases in which recall should be prioritized, for example, performance for recall should be pursued if the purpose is to detect a wide range of borrowers with signs of downgrading for early warning.[27]

[27] As indicators for measuring the predictive accuracy of a model in a binary classification problem, accuracy rate, precision, and recall are used. The calculation formulas for each indicator are as follows, where TP (True Positive) is the number of correct predictions for samples to be downgraded, TN (True Negative) is the number of correct predictions for samples not to be downgraded, FP (False Positive) is the number of false predictions for samples to be downgraded, and FN (False Negative) is the number of false predictions for samples not to

Figure 26: Performance of the model

		Accuracy rate	Precision	Recall
Real estate transaction sector	Training data (134,928)	0.923	0.954	0.031
	Test data (44,977)	0.919	0.703	0.017
Real estate leasing and management sector	Training data (139,545)	0.943	0.981	0.013
	Test data (46,515)	0.943	0.739	0.006
All industries	Training data (3,062,654)	0.911	0.923	0.006
	Test data (1,020,885)	0.910	0.704	0.003

* ( ) indicates number of samples.

2． Interpreting machine learning models

While machine learning enables predictions with a certain level of performance, whether humans can understand the logic of the model (explainability) has been an issue for practical application. In response to such issues, research on a method to increase the explainability of machine learning (eXplainable AI: XAI) has rapidly advanced in recent years. In this sub-section, one of the XAI method, SHapley Additive exPlanations (SHAP) value, is used to gain better insight into the model constructed.

The SHAP value quantitatively expresses which explanatory variable contributes to each predictive value output by the machine learning model. Figure 27 shows the SHAP values of a sample predicted to "downgrade" by the model (upper part of Figure 27) and a sample predicted "not to downgrade" (lower part of Figure 27). The more positive (negative) the SHAP value, the larger the contribution in the direction of raising (lowering) the downgrading probability. For example, the sample with the real estate transaction sector predicted to downgrade (upper left of Figure 27) shows a prediction result with a high probability of 0.693 while the average of the predictive values of all samples is 0.077. The variables that contributed to the difference were land prices, company size, the number of macro-real estate transactions, and borrowers' interest expense. The strength of the SHAP value is that it makes it possible to estimate which variables contributed to the prediction for each sample. Therefore, it can be used in practice, for example, to check the results for a borrower with large exposure or another borrower who needs intensive risk management.

be downgraded. Since an output of the models used is a probability value, in this analysis, the threshold was set at 0.6, i.e., samples with predictive values of 0.6 or more were classified as downgraded, while samples with predictive values of less than 0.6 were classified as not downgraded. Changing the threshold will change the results of each indicator.

Accuracy = (TP+TN) / (TP+FN+TN+FP), Precision = TP / (TP+FP), Recall = TP / (TP+FN)

Figure 27: SHAP values of four samples in test data (Top 4 variables)
(Top: samples predicted to downgrade, Bottom: samples predicted not to downgrade)

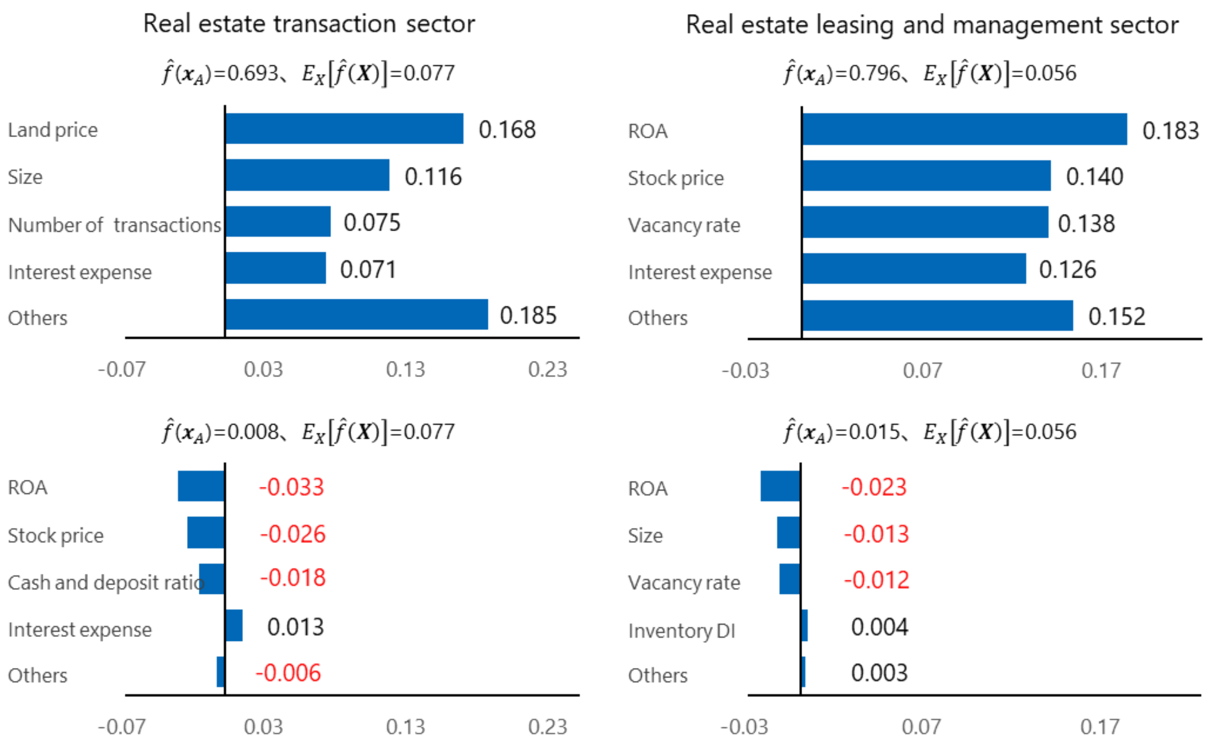
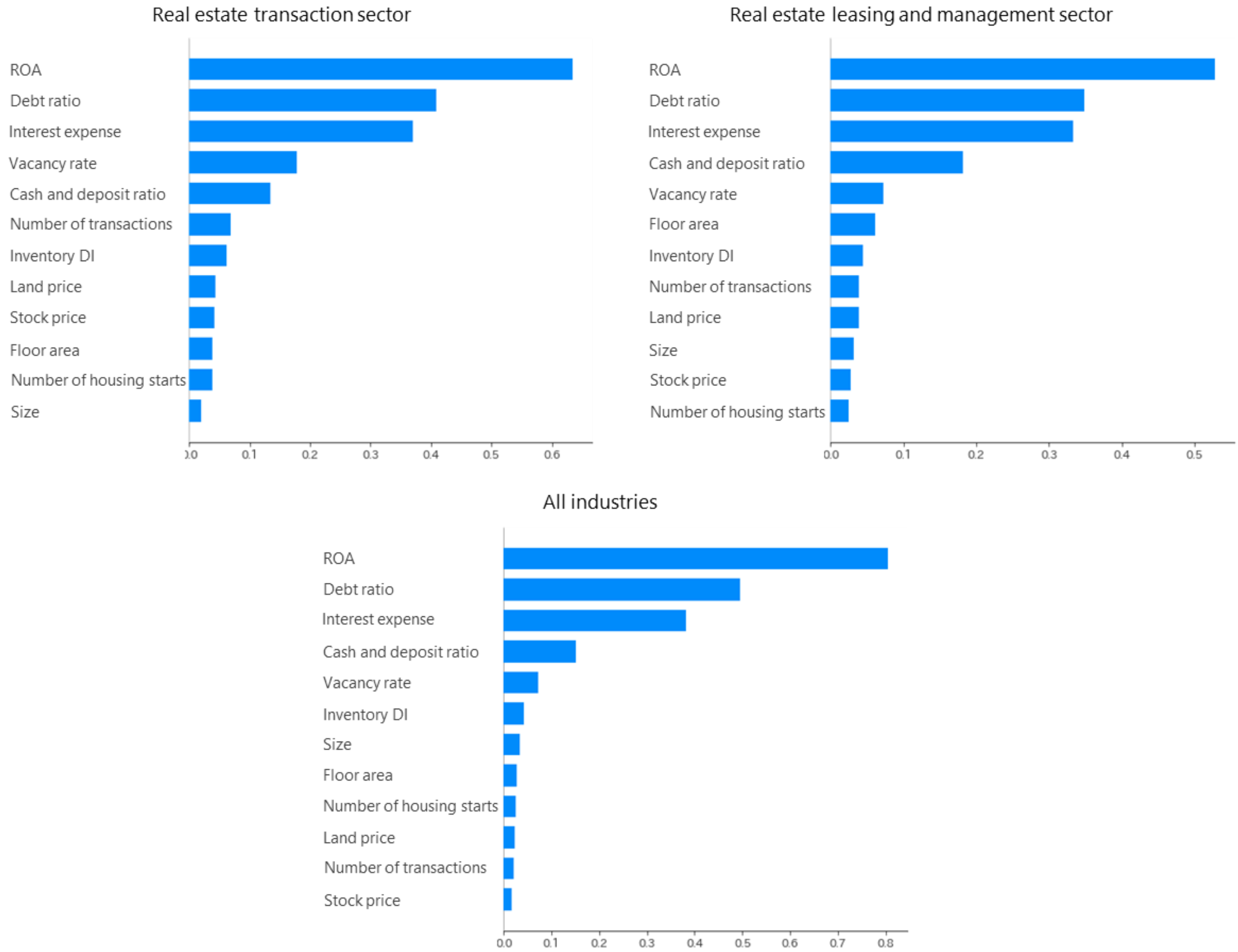


Figure 28 shows the average of the absolute value of SHAP values for all test data, which makes it easier to understand the overall trend of this model. When comparing the real estate sector with all industries, the absolute value of SHAP values of variables related to the real estate market, such as the vacancy rate and the real-estate inventory DI, tend to be somewhat larger for the real estate sector. Therefore, the real estate sector seems to be relatively more affected by the real estate market than other industries. On the other hand, the ROA, the debt ratio, the interest expense, and the cash and deposits ratio rank high in the real estate sector as well, suggesting that financial indicators of individual borrowers have a larger impact on the prediction of a downgrading than variables related to the real estate market.

However, as is the case for the sample in Figure 27, which has the number of macro real-estate transactions as a large contribution to the forecast, the degree of contribution of each sample does not necessarily match the overall trend. In the case of loans to SPCs, which are considered to have a closer relationship with the real estate market conditions, variables related to real estate market conditions are expected to have more contributions to the prediction. As such, the contribution of each variable will vary depending on the characteristics of the borrowers. Since the dataset used in this machine learning model is anonymized, it is difficult to verify at the individual sample level at this stage.

However, when detailed data are accumulated in the future, a more detailed analysis may provide a deeper understanding of the relationship between various variables and the future forecast by verifying the performance of the model while checking individual factors specific to each sample, such as the borrower's business model and location.

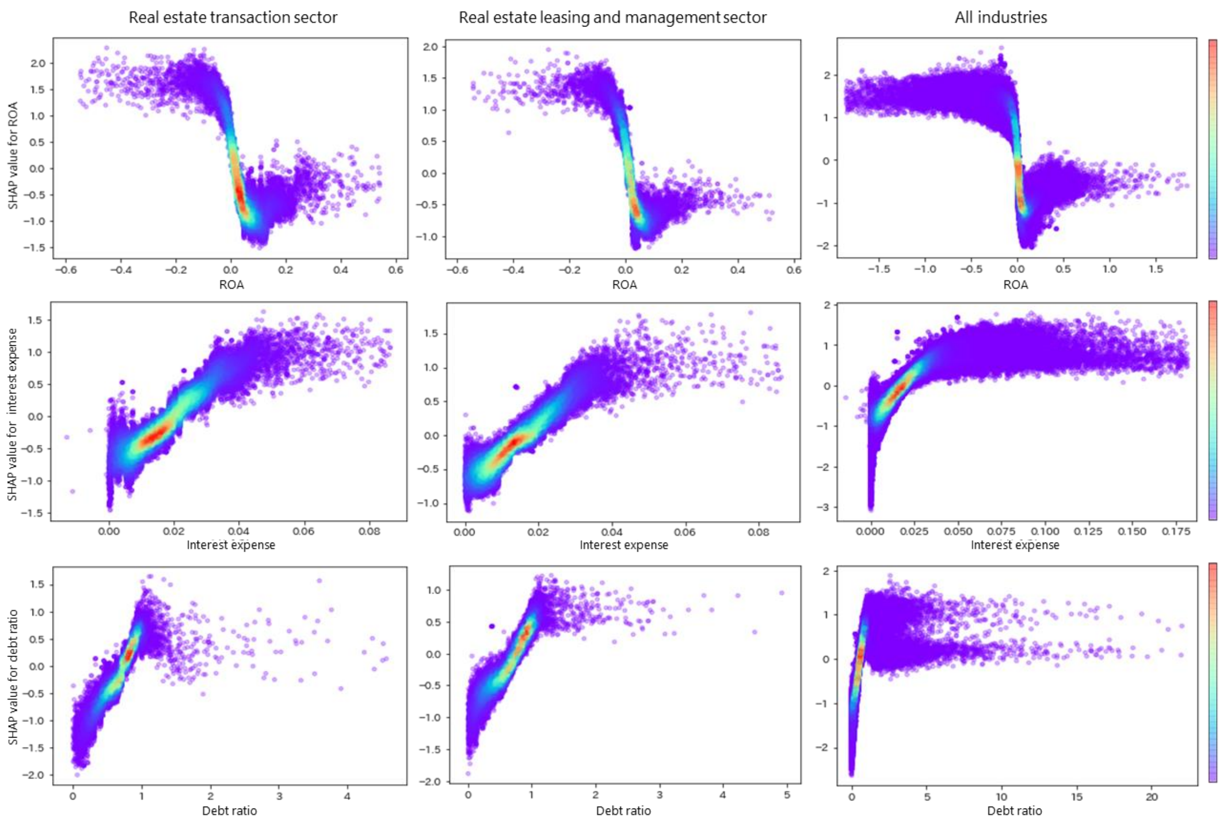
Figure 28: Interpretation of overall test data based on SHAP values
(explanatory variables that largely contribute to predicted results)



Finally, the relationship between the overall trend and the SHAP values of individual samples is visualized as a scatter plot in Figure 29, which shows the density of plotted data as a heat map. It can be seen from the upper part of Figure 29 that the impact of ROA, which is considered to have the highest contribution to the overall trend, changes depending on whether ROA is positive or negative. When ROA is positive, the SHAP value is negative, i.e., the ROA contributes to a decrease in the probability of a downgrading. However, when ROA exceeds +0.1, the number of samples decreases and the distribution of SHAP values becomes almost flat, and a similar trend is observed in the reverse direction (the SHAP values are distributed around -0.5 or + 1.5). In other words, when the absolute value of ROA exceeds a certain level, the impact of ROA on the forecast decreases.

As for the debt ratio, the model is in line with intuition that when the ratio is in the range of 0 to 1, the debt ratio and the SHAP value have a positive proportional relationship, i.e., the higher the debt ratio, the higher the probability of a downgrading. However, once the debt ratio exceeds 1, the proportional relationship collapses (the lower part of Figure 29). Given the fact that borrowers with a debt ratio over 1 are classified as "normal" at the beginning of the fiscal year, they may have unique financial and business models. It is in general undesirable for these special samples to have an excessive impact on the model construction, however, in this model, the impact of the special samples seems to be limited given the distribution of SHAP values shown in Figure 29.

Figure 29: Scatter plot of SHAP values



3． Conclusions and Discussion

As described above, the construction and interpretation of prediction models using machine learning confirmed the following: (1) the impact of real estate market conditions on the prediction of downgrading probability may be larger in the real estate sector than in other sectors. (2) However, financial indicators have larger contributions to predictions than real estate market conditions. (3) When looking at individual cases, there are samples in which real estate market conditions make a large contribution to predictions, indicating that focusing on individuality is also important.

However, it should be noted that there are likely to be many factors that cannot be captured by this dataset, for example, trends in the real estate sector and real estate market conditions are usually affected by structural and macro factors, such as population dynamics and trends in foreign investors, as well as by micro factors, such as the decline in demand in certain areas due to redevelopment in an adjacent area. In addition, since the learning was made based on the data for the past 20 years, the performance of predictions may deteriorate when the situation changes from the past, such as the situation under the interest rate hike period.

Box 2: Explanation of the relationship between the downgrading and financial indicators based on the logistics regression

A logistic regression analysis is a typical analytical method for a binary problem of estimating whether or not an event occurs. It is often used in the field of credit risk as a tool to estimate downgrading and defaults, since the assumption that an objective variable and explanatory variables have a linear relationship enables the interpretation of the model by checking the sign and significance level of the estimated coefficient. In the "FSA Analytical Notes (2023.6)," a logistic regression model is also used to estimate the probability of borrowers' default by using financial indicators and macro variables as explanatory variables.

Given that the machine learning model in this analysis estimated that the SHAP values of various financial indicators were relatively large, in order to deepen the understanding of the relationship between financial indicators and the probability of a downgrading, a logistic regression model is developed by using the same financial indicators from the same dataset. To avoid the problem of collinearity or spurious correlation, which often occurs when developing regression models with an excessive number of explanatory variables, four explanatory variables that had a large average of absolute SHAP values (ROA, interest payment, debt ratio, and cash and deposits ratio, as shown in Figure 28) are selected as the explanatory variables (and are standardized to adjust the scale of variables), and other variables related to the macro economy and market conditions are treated as dummy variables as time-fixed effects. The regression formula is as follows:

$$log\frac{p_i}{1-p_i} = \beta_0 + \sum_{k=1}^{4} \beta_k \cdot Zaimu_{k,i} + \sum_{t=2006}^{2022} \beta_t \cdot Year_{t,i} + \varepsilon$$

The results of the estimation are shown in Figure 30. All four financial indicators were significant by 0.1% in both the real estate transaction sector and the real estate leasing and management sector, indicating that there is a significant correlation with the probability of a downgrading. The absolute value of the regression coefficient for standardized-ROA was the largest in both the real estate transaction sector and the real estate leasing and management sector. This is the same suggestion as that provided by the machine learning model in this analysis.

In the logistic regression analysis, the "direction" in which each explanatory variable affects the objective can also be easily confirmed by looking at the sign of the estimated result. The results show that a decrease in ROA, an increase in interest expense, an increase in debt ratio, and a decrease in the cash and deposits ratio contribute to an increase in the probability of a downgrading.

On the other hand, as shown in Figure 29, this dataset includes a certain number of samples with extreme financial indicators, such as debt ratios far exceeding 1. Since the logistic regression model assumes a linear combination of variables, there is a concern that, in case of the forecast using such an extreme sample as an input, the result might be significantly influenced by the extreme variable and become implausible.

As described above, logistic regression analysis is simple in structure and provides clear and easy-to-interpret results, and thus it can be applied to a practical area, for example, sensitivity analysis and stress testing. However, it should be carefully examined whether the structure is consistent with the characteristics of the dataset. It is important to select models and variables according to the actual situation and the way to use the model, while taking into account the characteristics of such models.

Figure 30: Estimation results of regression coefficients in logistics regression models

Explanatory variables	Real estate transaction sector		Real estate leasing and management sector		All industries	
ROA	-0.481	***	-0.504	***	-0.545	***
Interest expense	0.333	***	0.356	***	0.248	***
Debt ratio	0.411	***	0.387	***	0.308	***
Cash and deposit ratio	-0.248	***	-0.266	***	-0.212	***
pseudo-R2	0.099		0.069		0.081	

"***" indicates that the significance level is met at 0.1%.

No. of samples in total	179,905	186,060	4,083,539
No. of samples ranked down	14,438	10,730	366,204

IV. Wrap-up

In this paper, real estate loans by regional banks have been analyzed using detailed loan data obtained from the newly established Common Data Platform. Although there has been no significant change in the credit risk of real estate loans, the regional banks' loan portfolios are more likely to be affected by real estate market conditions in the Tokyo metropolitan area because they extend loans to areas other than their home regions, such as Tokyo. In addition, the real estate sector is more likely than other sectors to be adversely affected by a rise in borrowing interest rates.

In addition, the relationship between financial indicators, real estate market conditions, and borrowers' credit ratings has also been analyzed using machine learning and other techniques. Although there is still room to improve the performance of the model, certain observations are found, for example, (1) the real estate sector may be more susceptible to the real estate market conditions than other sectors, and (2) financial indicators have a large contribution to the deterioration of credit ratings for all sectors.

The FSA will continue to gather high-quality datasets to gain a detailed understanding of real estate loans and conduct monitoring based on such data. The FSA will also improve its analytical models and techniques to deepen the understanding of the relationship between real estate loans and real estate market conditions, and enhance the monitoring capabilities using granular data.

Analysis of Corporate Transaction Network

(Summary)

This paper presents the FSA's recent initiatives to visualize the network structure among financial institutions and enterprises. Focusing on the corporate transaction network, this paper also considers indicators that can be used to identify significant enterprises that may spread or accumulate the impact of the default of an enterprise on the network. The computed indicators suggest that some enterprises are more important than others. The FSA will continue its research on network analysis methods to deepen its understanding of the financial system and the real economy and enhance its monitoring capabilities.

I. Introduction

From the macroprudential perspective, this paper emphasizes the interconnectedness (network structure) between economic agents and how it may affect the financial system, focusing on, for example, how financial institutions conduct transactions with one another and how they are interconnected with financial markets and the real economy. In the case of bank lending, it is necessary to assess risks not only on a borrower firm basis but also in light of the network structure to which the borrower firm belongs, such as supply chains.

It is important for the FSA to understand the network structure surrounding the Japanese financial system and its characteristics to ensure macroprudence. Network analysis techniques, which express networks as nodes (points) and links (lines) and then quantify and examine the characteristics, are applied in various fields, including communications and transportation, human friendships, and relationships between words in sentences. Economic activities, such as business-to-business transactions, are no exception, and a variety of previous studies are available, covering various industries and regions in scope. Research on analytical methods for efficiently handling complex networks using large-scale data is also developing.[1]

[1] For more information, see HAYASHI Yuki: "Python and Complex Network Analysis" (Kindai Kagaku Sha, Impress Group Publication

This paper presents the following two initiatives conducted by the FSA in relation to network structure: (1) visualizing the network structure between financial institutions and enterprises using bank loan-by-loan level data (Section II); and (2) identifying significant enterprises within the corporate transaction network that may spread or accumulate the impact of a corporate default in the network (Section III).

II. Visualization of the network structure between financial institutions and enterprises

The FSA has developed an internal tool to visualize a network comprised of corporate transactions, equity investment, and bank lending. The tool is based on bank loan data collected by the Common Data Platform,[2] a novel framework for data collection and management that has been launched in phases jointly with the Bank of Japan, and domestic business-to-business transaction data provided by a third-party vendor.[3] (Figure 1)

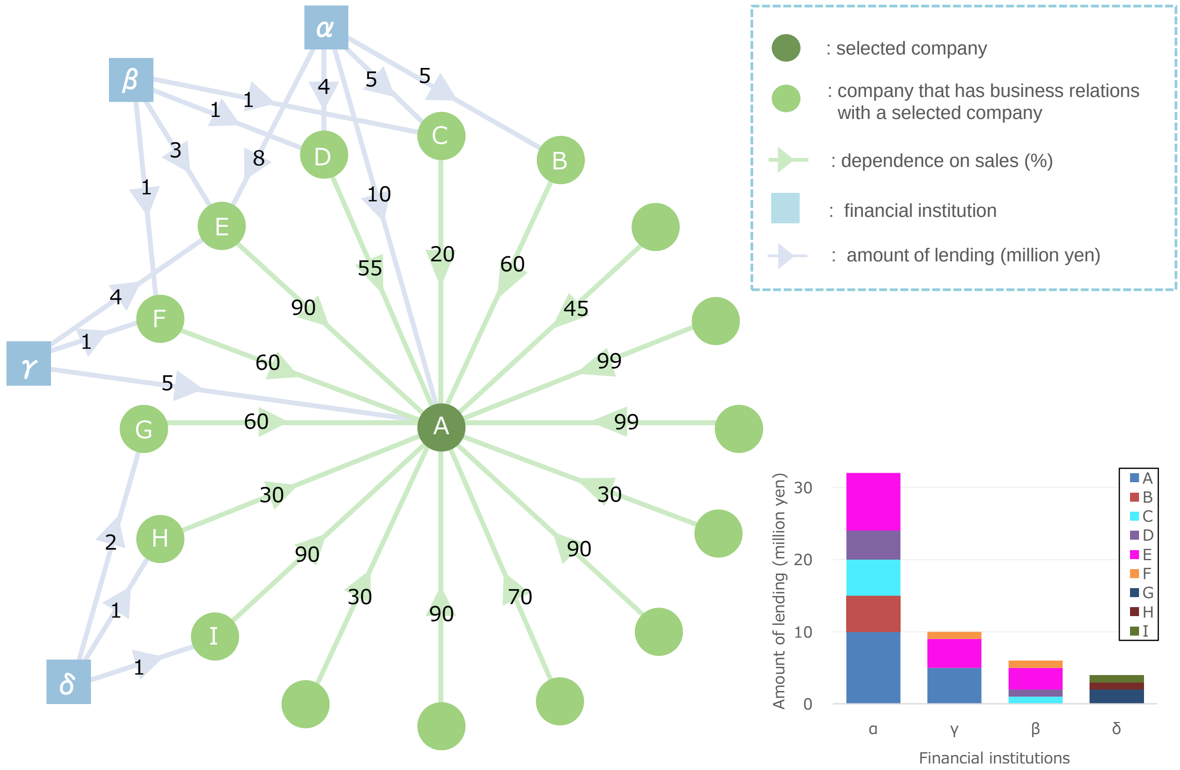
The tool renders a network that shows business relationships, such as lending and investment, as directional links with companies and financial institutions as nodes, and display relevant information, such as transaction amounts, all at once. Such visualization makes it possible to grasp the impact of a credit event, such as corporate bankruptcy, to a certain extent by explicitly showing the companies trading densely with the bankrupted firm and financial institutions with large amounts of loans to the firm within the scope of the data held.

Information, 2019) and SUZUKI Tsutomu: "Network Analysis" (Kyoritsu Shuppan, 2017). Hayashi's research includes an analysis of trade networks that takes into account the industrial diversity of countries, as well as research on methods for extracting "communities" in networks in which nodes are more closely connected.

[2] FSA, “Progress in Common Data Platform and Next Steps” <https://www.fsa.go.jp/en/news/2024/20240701/20240701.html>

[3] Information obtained from Teikoku Databank, Ltd.

Figure 1: Visualization of a network structure between financial institutions and enterprises



A network shown in Figure 1 only illustrates the network centered on Company A to ensure visibility. However, the actual transactions of the illustrated companies could be more complex given that Company B and others usually have business partners other than Company A. It is not easy to visualize the network including these secondary and tertiary customers in a form that can be understood intuitively. Therefore, in order to accurately understand the impact of the whole network on the financial system, it would be effective to introduce some indicators that cover the network features not shown in the visualization.

Other initiatives within the FSA to visualize the network structure are undertaken in the field of the derivatives market using transaction data on over-the-counter derivatives[4] and securities companies' funding network using data on repo markets[5]. In order to further promote such analyses using granular transaction data, it is necessary to continue efforts to ensure the quality of the collected data.[6]

[4] KAWAI Daisuke, HASEGAWA Masaki, and YAGI Risa, "An analysis of the transaction network in the Japanese OTC derivatives markets," 2021, FSA Staff Reports and Columns < https://www.fsa.go.jp/frtc/english/seika/srhonbun/20210707_SR_Derivative_ArticleEN.pdf>

[5] See p56 in column of "The JFSA Strategic Priorities July 2023-June 2024."

[6] For example, the data collected under the over-the-counter derivatives reporting system has a problem of duplicate reporting as both parties are required to report under the original system. The system is being improved.

III. Network analysis of corporate transactions

Various indicators have been proposed for measuring important nodes in network analysis. Typical indicators include degree centrality, which uses the number of links to other nodes, and closeness centrality, which focuses on the shortest distance to other nodes. Such indicators are useful in that they enable quantitative comparisons among nodes and network structures, even if these structures are complex. On the other hand, in order to utilize the indicators for systemic risk analysis, it is desirable for indicators to reflect the characteristics of nodes, such as the financial condition of companies.

Since the financial crisis in 2008, the assessment of interconnectedness has become an important part of global systemic analysis. In the Financial Sector Assessment Program (FSAP)[7] for Japan conducted by the International Monetary Fund (IMF) from 2023 to 2024, interconnectedness was reviewed as a part of the systemic risk assessment for the Japanese financial system. In the FSAP, mutual exposure data, including deposits, securities holdings of Japanese banks, insurance companies, and securities companies, is used to understand the network structure of the Japanese financial system. In addition, the following two assessments were conducted, taking into account not only the network structure but also the characteristics of nodes, such as the soundness of each financial institution: (i) the degree of impact of the failure of a certain financial institution on other financial institutions, and (ii) the degree of impact of the failure of other financial institutions on a certain financial institution.

In this paper, with reference to the above FSAP point of view, two indicators that measure the importance of networks are developed for the domestic corporate transaction network, i.e., (i) the impact of a failure of a certain firm on other firms (Contagion index) and (ii) the impact of a failure of other firms on a certain firm (Vulnerability index).

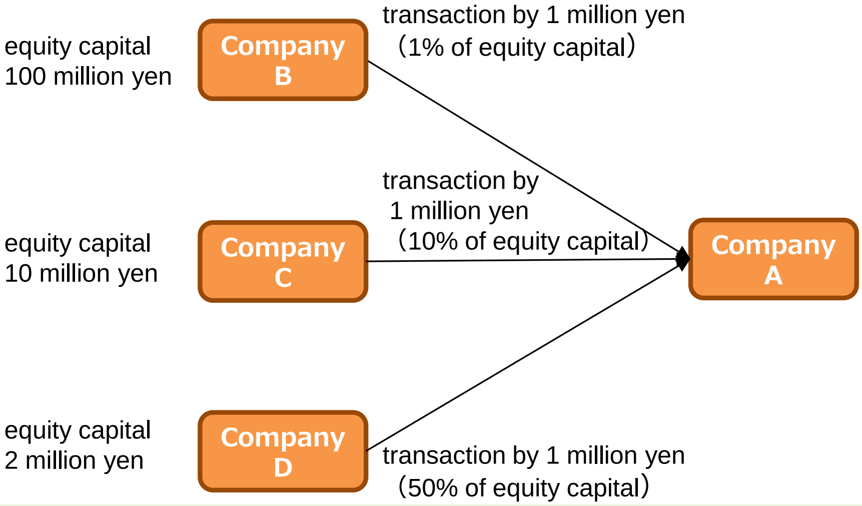
[7] The FSAP is a program by which the IMF assesses the stability of member countries' financial sectors. Major countries, including Japan, undergo a review every five years. For the results of the 2023 FSAP, see below.<https://www.fsa.go.jp/inter/etc/20240514/20240514.html>

1. Impact of failure of a certain firm on other firms (Contagion index)

As a simplified example, assume that Company A has purchased (accounts payable) 100 million yen from Companies B, C and D, respectively (Figure 2). If Company A enters bankruptcy, making it impossible for Companies B, C and D to collect their accounts receivable from Company A, Companies B, C and D will book credit losses for that amount. The impact of the loss will vary depending on the business strength of each company. To capture the impact of loss to all suppliers (in this case, Companies B, C and D) due to the counterparty failure (in this case, Company A), taking into account the business strength of the suppliers, the "cumulative capital loss ratio"[8] is calculated as follows: compute the ratio of loss to each supplier relative to equity capital when the counterparty fails, and aggregate these ratios. The indicator enables a consideration of the impact in accordance with the amount of equity capital of the company that suffers losses, in addition to the transaction amount.

In the following sections, the above "cumulative capital loss ratio" is defined as a "Contagion index" that measures the impact of one firm's failure on other firms in the network.[9]

Figure 2: Example calculation of the “cumulative capital loss ratio”



“Cumulative capital loss ratio” (by failure of Company A) equals 61%.
(1% + 10% + 50% = 61%)

[8] Instead of equity capital, the ratio to total assets or sales can also be used.
[9] It should be noted that the "Contagion index" focuses on the financial impact of the failure of a firm receiving goods or services on the firm providing those goods or services, but not vice versa.

2. Impact of failure of other firms on a certain firm (Vulnerability index)

If the impact of the failure of a certain firm in the network is likely to lead to other firms' failure, there is a concern that the impact may propagate through the network in secondary and tertiary ways, thereby increasing systemic risk. Hereinafter, whether or not a firm is susceptible to the impact of other firms is calculated by referring to the method of Freeman et al.[10]

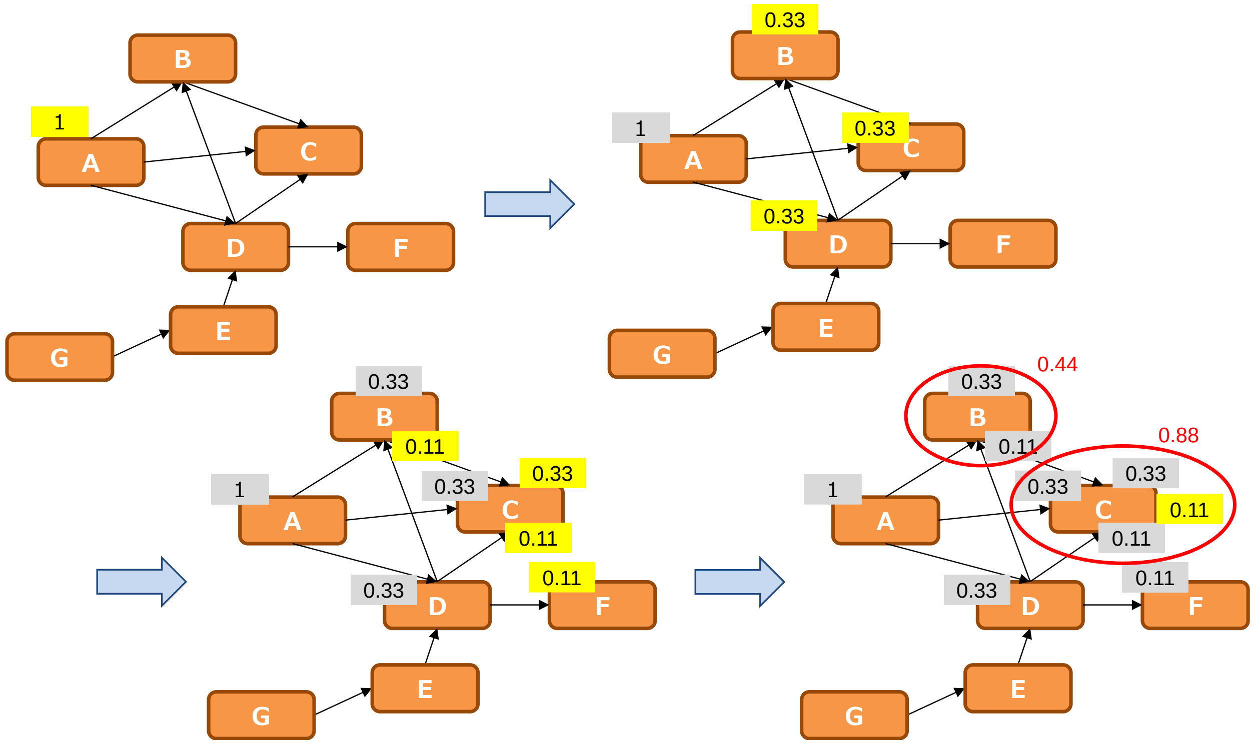
Freeman et al. assumed that certain company A in the network has a resource of 1 and distributes the resource equally among companies that have business relations with the company. A company that receives the resource also distributes the resource to each company that has business relations in the same manner, and the operation is repeated until reaching a steady state. The total of the resources that have passed through each node is defined as a "dependence index." Taking Figure. 3 as an example, it is possible to make a quantitative comparison, such as "the impact of the failure of Company A is larger for Company C than for Company B on the network."

Figure 3 shows the "dependence index" only for Company A, but the same calculation can be done for all companies on the network. Freeman et al. defined "importance index" as a value obtained by adding up, at each node, the "dependence index" for each company calculated in this way. This makes it possible to identify nodes through which resources uniformly distributed on the network frequently pass. In other words, a company with a higher "importance index" is more likely to be affected by other companies on the network.

In this paper, the "importance index" in Freeman et al. mentioned above is defined as "Vulnerability index."

[10] Freeman, Linton C., Stephen P. Borgatti, and Douglas R. White. "Centrality in valued graphs: A measure of betweenness based on network flow." Social networks 13.2, 1991, P.141-154.

Figure 3:[11] Example of the calculation of "Dependence index" (The case of Company A)



[11] The arrows in the figure represent the directions in which the services are provided. Specifically, it is assumed that Company A sells some sort of products to Companies B, C and D.

3. Results and discussion

Figures 4 and 5 show the distribution of the Vulnerability index and Contagion index calculated for a domestic corporate transaction network constructed from seclected samples[12] of business-to-business transaction data, with the maximum value of each index set at 100 for normalization.

Figure 4: Distributions of Vulnerability index and Contagion Index

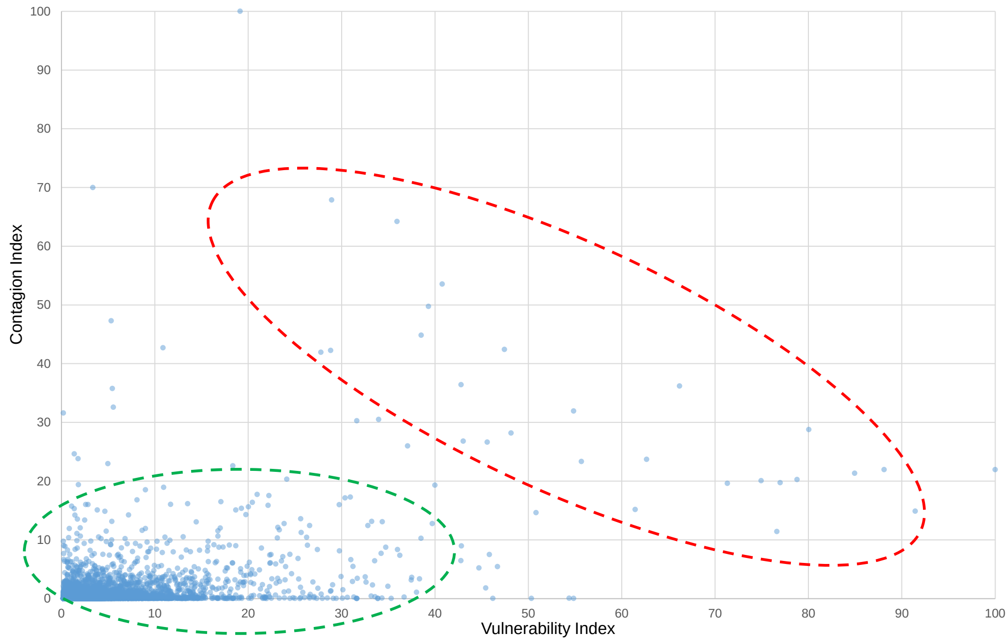
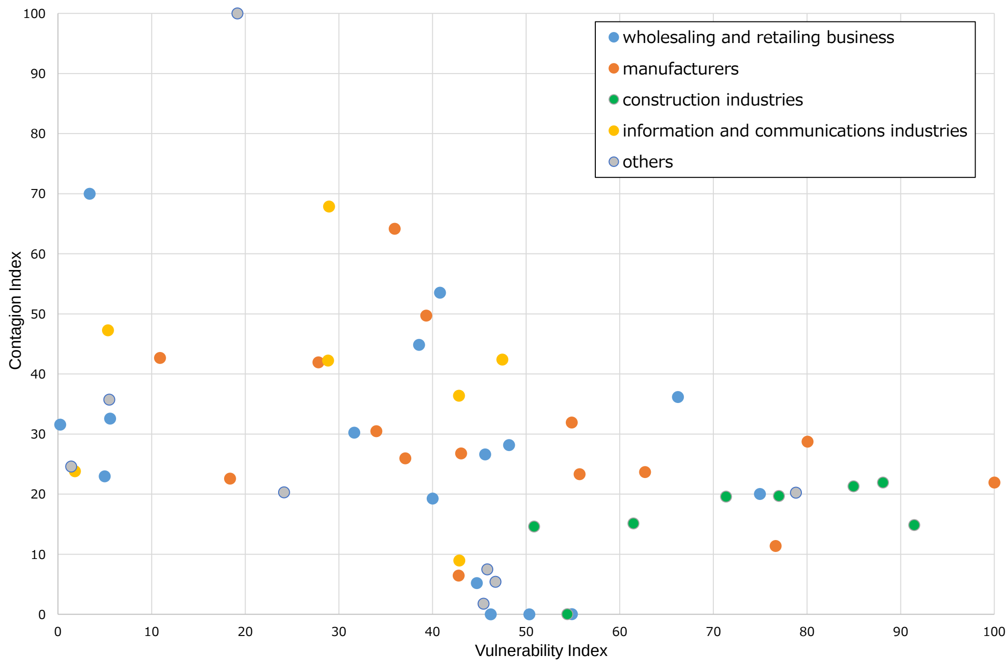


Figure 5: Distributions of Vulnerability index and Contagion Index (broken down by industry, only Vulnerability index > 40 and Contagion index > 20 are shown to ensure visibility)



[12] Firms with a transaction amount of 100 million yen or more are included in the scope. The number of samples is around 22,000.

Two trends can be observed from Figure 4. First, for both indices, companies tend to be concentrated where the index is small (circled in green), indicating that only a small number of companies have a relatively large network influence or are likely to be significantly affected.

Second, for firms where one or both indicators are above a certain level (circled in red), there is a tendency for a trade-off to occur in which a large one of the indicators leads to another indicator being small. As shown in Figure 5, looking at the industries of firms where both indicators are above a certain level, it is observed that manufacturers (automobile manufacturers) and information and communications industries have relatively high contagion indices, while manufacturers (heavy industries) and construction industries have relatively high vulnerability indices.

This trend is likely due to differences in the natures of firms that have an influence on other firms and firms that are affected by other firms. Both indicators, by definition, have a common feature in that they increase as the number of firms (suppliers) that have sales to the firm increases. While the Contagion index increases as the number of suppliers that have sales to the firm increases relative to their own capital, the Vulnerability index tends to increase as nodes on a network become densely populated, such as where the firm and its suppliers use up many resources. In a nutshell, differences in the number and nature of suppliers among firms or industries may be a factor in such distribution.

Among the samples used in this analysis, no firms had extremely high values for both the Contagion index and the Vulnerability index. However, if the same calculation is performed for a network limited to a specific region or industry, a different result may be obtained. In addition, as stated in the footnote, since the scope of the analysis is limited to firms with a transaction amount of 100 million yen or more, transactions related to SMEs and retail consumers are considered to be excluded in many cases. If the index is constructed without setting a threshold for the transaction amount, the result of industries that conduct small-scale transactions with many counterparties or industries that conduct transactions with retail consumers may change.

IV. Conclusion

In this paper, the FSA's initiatives to visualize the network structure among financial institutions and enterprises are presented, and indicators to identify enterprises that are important in the network are calculated, taking into account the characteristics of enterprises, such as their size, using the corporate transaction data obtained from a third-party vendor. The results suggest that these indicators could be utilized to grasp and identify not only the network structure but also the significant nodes in the network. However, the indicators developed in this paper form just one example. There is room for further improvement, for example, incorporating the evaluation of enterprises that are difficult to substitute in the supply chain (such as those possessing monopolistic technology).

The results of this analysis, when used in conjunction with banks' loan data, are expected to contribute to the assessment of banks' credit risk and contagion simulation in light of the network structure to which borrowers belong. Furthermore, the addition of foreign firms to the network could lead to the development of a country risk simulation. In addition, while this analysis focused on micro-level data, such as business-to-business transaction data, more macro-level data (e.g., input-output table) could also be used to deepen the analysis. The FSA will continue to advance research on network analysis methods with the aim of enhancing monitoring while deepening our understanding of the financial system and the real economy.

Analysis of Corporate Bankruptcies arising from Human Resource Shortages amid the Changing Working Environment

(Summary)

This paper attempts to understand the financial characteristics of firms that went bankrupt due to human resource shortages by using corporate financial and bankruptcy data obtained from a third-party data vendor. It revealed that firms that went bankrupt due to human resource shortages had different characteristics from other firms in terms of labor cost burden, operational efficiency, and other factors. The FSA will continue to monitor, in a forward-looking manner, the impact of changes in the macro environment on the financial system, such as whether the changing labor situation leads to an increase in credit risk for financial institutions.

I. Introduction

In Japan, the working-age population (15 to 65 years old) has been declining since its peak in 1995 due to the aging of the population and the declining birthrate.[1] In addition, the labor situation is facing major challenges, such as the so-called "2024 problem," in which labor shortages are a concern due to the application of overtime caps to the construction and other industries starting in April 2024,[2] and the so-called "2025 problem," in which the generation born between 1947 and 1949, a large group in terms of population composition, will reach the late-elderly age (over 75 years old). These changes may affect the society from various perspectives in the future. If these changes exacerbate various social and economic challenges, it may have an adverse impact on financial institutions in the form of increased credit risk and decreased lending demand and borrowers. Therefore, it is useful to analyze the impact of changing labor environment on the business conditions of Japanese firms, to ensure

[1] National Institute of Population and Social Security Research, "Population Projections for Japan: 2021 to 2070"
https://www.ipss.go.jp/pp-zenkoku/j/zenkoku2023/pp_zenkoku2023.asp
[2] Ministry of Health, Labour and Welfare
https://hatarakikatakaikaku.mhlw.go.jp/overtime.html
Due to the revision of the Labor Standards Law, the upper limits have been in effect for large companies since April 2019 and for small and medium-sized companies since April 2020. Of these, the upper limits became effective in April 2024 for the construction, transportation, and physician industries, although there was a five-year grace period before they became applicable.

the financial stability.

This report focuses on the recent growing human resources shortages and examines the characteristics of corporate bankruptcies by defining three types of bankruptcies[3]: bankrupted companies caused by human resources shortages ("HR-shortage bankrupt companies"), bankrupted companies caused by other reasons ("other bankrupt companies"), and neither ("surviving companies"). The analysis in this report is based on financial data and bankruptcy data obtained from a third-party data vendor, but it should be noted that the number of HR-shortage bankrupt companies covered in the database is small compared to the total number of bankrupt companies, and thus the sample may be biased.[4]

II. The labor situation in Japan

Figure 1 shows the change in the employment headcount determination D.I., which expresses a company's perception of an excess or shortage of manpower. A positive (negative) value for this indicator indicates that a high percentage of companies believe they have an excess (shortage) of workers. Note that a temporary increase seen from 2020 to 2021 is due to a sharp decline in labor demand following the COVID-19 pandemic and the declaration of a state of emergency.

Figure 2 shows the trend of the D.I. for production and sales equipment, which indicates companies' perception of an excess or shortage of equipment. A positive (negative) value for this indicator indicates that a high percentage of firms consider their facilities to be excessive (insufficient). After peaking in 2009, the sense of excess facilities gradually dissipated, with values for the non-manufacturing sector and the manufacturing sector becoming negative in September 2013 and September 2017, respectively. Subsequently, the sense of excess temporarily strengthened as with the employment headcount D.I., but a sense of shortage is now emerging again. The reasons for this resurgence in the percentage of firms that believe their facilities are inadequate may be due to pent-up demand for capital investment, which had been postponed due to the COVID-19 pandemic, and

[3] Bankruptcy in this report refers to a company that is recognized by Teikoku Databank, Inc. as falling under one of the following six cases: (1) suspension of bank transactions, (2) internal liquidation (when the representative acknowledges bankruptcy), (3) filing for commencement of corporate reorganization proceedings with the court, (4) filing for commencement of civil rehabilitation proceedings with the court, (5) filing for commencement of bankruptcy proceedings with the court, or (6) filing for commencement of special liquidation with the court. The term "bankruptcy data" refers to data on companies that are recognized as falling under one of the six cases listed in the six categories above.

[4] This report analyzes companies for which financial and bankruptcy data are available based on information from Teikoku Databank, Ltd. In this report, "bankrupt firms with insufficient human resources" are defined as those whose bankruptcy factors in the bankruptcy data of Teikoku Databank include "lack of human resources," which indicates difficulty in securing human resources, and "illness or death of the manager," which indicates the absence of successor managers.

expectations for an economic normalization. In addition, the increased demand for capital investment to improve operational efficiency in response to human resources shortages may also have had an impact.

Figure 3 shows the year-on-year quarterly growth rates of nominal and real wages. At present, the growth rate of nominal wages has been in positive territory due to factors such as labor shortages and inflation. However, the growth rate of real wages as of January-March 2023 is negative due to the high rate of price inflation.

In light of the above, it is possible that the worsening shortage of human resources in Japan is having a ripple effect on trends in capital investment and wages, affecting the survival of firms. Figure 4 shows the number of bankrupt companies and the percentage of HR-shortage bankrupt companies among all bankrupt companies. Unlike the trend in the number of bankrupt companies, the percentage of HR-shortage bankrupt companies is in the single-digit range but is on an upward trend.

Figure 1 :Change in the employment headcount determination D.I.[5]

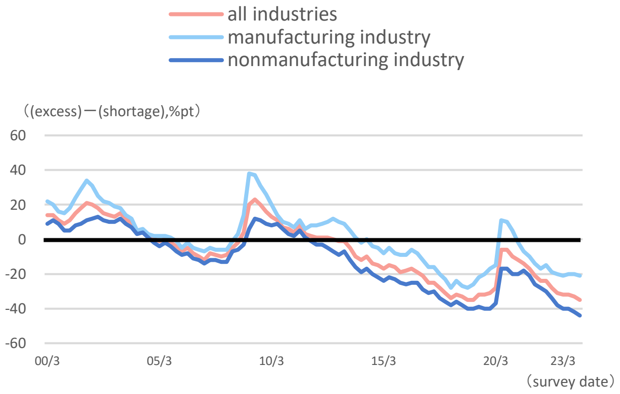
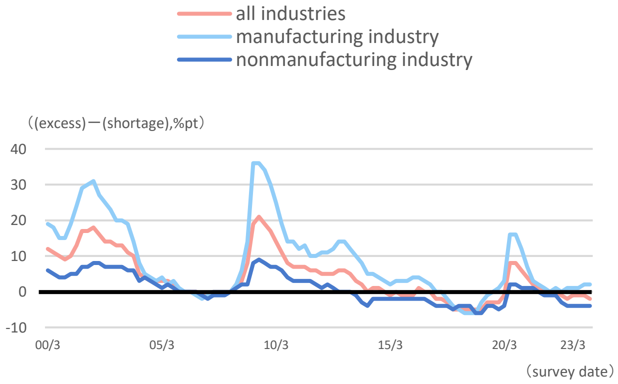


Figure 2 :Trend of the D.I. for production and sales equipment [6]



[5] Source: Bank of Japan, “National Short-Term Economic Survey of Enterprises in Japan.”
[6] Source: Bank of Japan, “National Short-Term Economic Survey of Enterprises in Japan.”

Figure 3: Year-on-year quarterly growth rates of nominal wages and real wages [7]

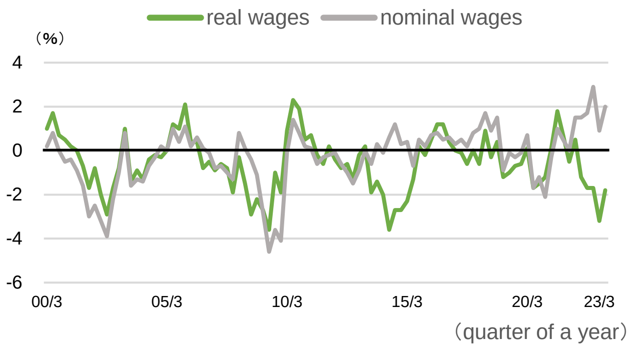
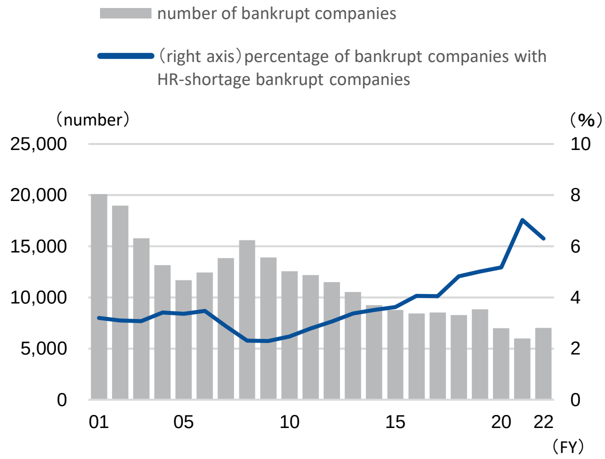


Figure 4: Number of bankrupt companies



III. Analysis of HR-shortage bankrupt companies

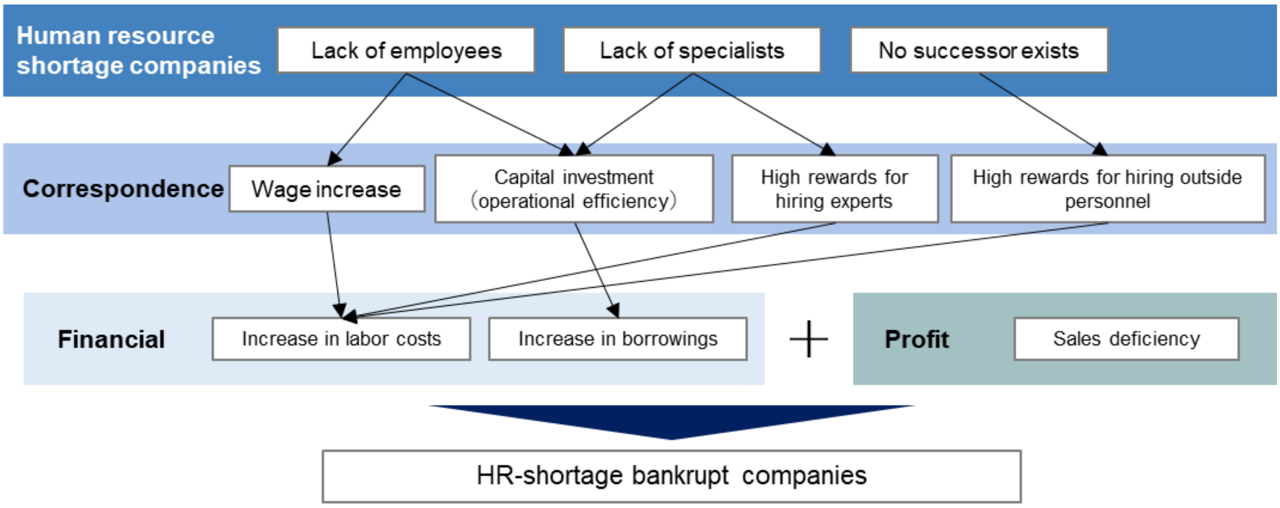
As mentioned in the previous section, the worsening of the human resource shortage and the rising trend in the ratio of companies that have gone bankrupt due to human resource shortages have been observed. In this chapter, the distribution and trend of each financial indicator up to bankruptcy are examined, in order to understand the characteristics of firms that went bankrupt due to a shortage of human resources.

1. Background to human resource shortage bankruptcies

In examining bankruptcies caused by human resource shortages, the transmission channels of bankruptcy are briefly summarized in Figure 5. While there is a wide range of measures that can be taken to address human resource shortages, this report focuses mainly on the burden of labor costs, in light of the recent trend of a wage increase.

[7] Source: Ministry of Health, Labour and Welfare, “Provisional Report of Monthly Labour Survey”

Figure 5: Background to human resource shortage bankruptcy[8]



2. Comparison of the distribution of financial indicators by bankruptcy status

First, the distribution of financial indicators are analyzed for each bankruptcy categories. i.e., "HR-shortage bankrupt companies," "other bankrupt companies," and "surviving companies." For the HR-shortage bankrupt companies and other bankrupt companies, the figures for the fiscal year of the bankruptcy are used, and for the surviving firms, the figures for the latest fiscal year available before the end of September 2023 are used. In this section, the distributions are presented without considering the industry sector in order to ensure sufficient sample size and to grasp the overall trend.

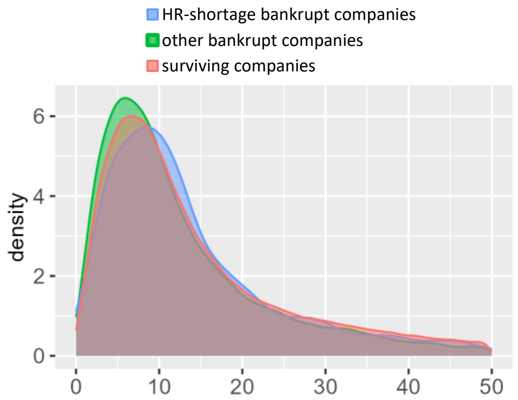
Figures 6 and 7 show the distribution of the labor cost over net sales ratio and the labor share, respectively. A comparison of the medians of the labor cost over net sales ratio and the labor share for HR-shortage bankrupt companies and other bankrupt companies revealed that the labor cost burden may be heavy for HR-shortage bankrupt companies, as both ratios are high.

Table 1: Number of samples[9]

	n	manufacturing	wholesale	construction	service	retail	transportation	real estate	other
HR-shortage bankrupt companies	1,637	17.9%	35.7%	13.9%	18.9%	6.1%	4.3%	3.0%	0.2%
other bankrupt companies	45,094	23.3%	31.3%	11.3%	18.6%	8.3%	3.4%	3.3%	0.5%
surviving companies	1,036,355	22.2%	29.7%	9.8%	22.1%	7.2%	4.0%	4.4%	0.6%

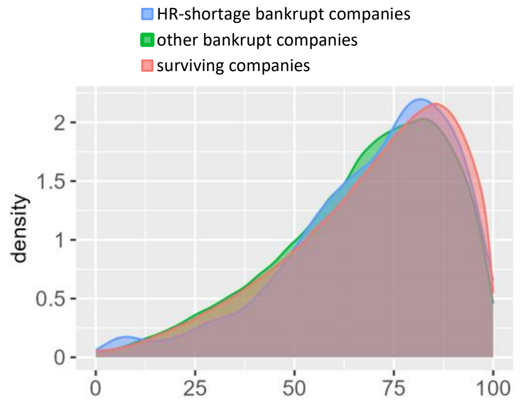
[8] Other possible measures include mergers and acquisitions.
[9] The period covered is April 2001 to September 2023.

Figure 6: Ratio of labor cost to net sales(%)[10]



	median	mean
HR-shortage bankrupt companies	10.41	14.14
other bankrupt companies	8.70	11.76
surviving companies	8.90	11.88

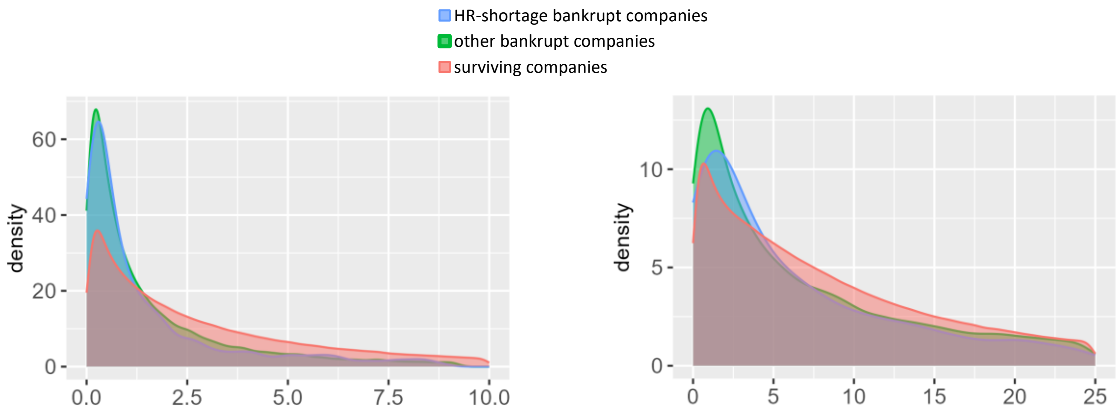
Figure 7: Labor share(%)[11]



	median	mean
HR-shortage bankrupt companies	82.15	75.54
other bankrupt companies	77.44	67.77
surviving companies	74.77	65.95

In Figures 8 through 10, the distribution of profitability, safety, and productivity indicators are shown respectively. Again, the median values indicate that the dependence on debt as the safety indicator, and turnover of tangible fixed assets as the productivity indicator, tend to be higher for HR-shortage bankrupt companies.

Figure 8: The profitability index
(Left: ratio of ordinary income to net sales(%)[12]; Right: ROE(%)[13])



	median	mean
HR-shortage bankrupt companies	0.26	-0.93
other bankrupt companies	0.38	-0.56
surviving companies	1.61	2.52

	median	mean
HR-shortage bankrupt companies	2.97	7.96
other bankrupt companies	3.77	9.45
surviving companies	6.21	9.43

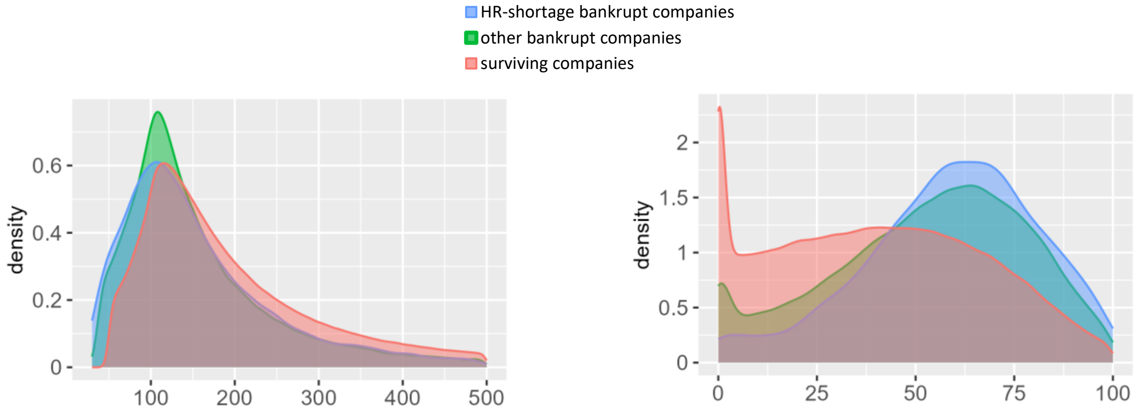
[10] labor cost ÷ sales × 100 (labor cost = salary expense + directors' compensation + bonus expense + directors' bonus + welfare expense)

[11] labor cost ÷ value added × 100 (value added = operating income + real estate rent + interest payment and discount expenses +depreciation costs + taxes and dues + labor cost)

[12] ordinary income ÷ sales × 100

[13] net income ÷ equity×100

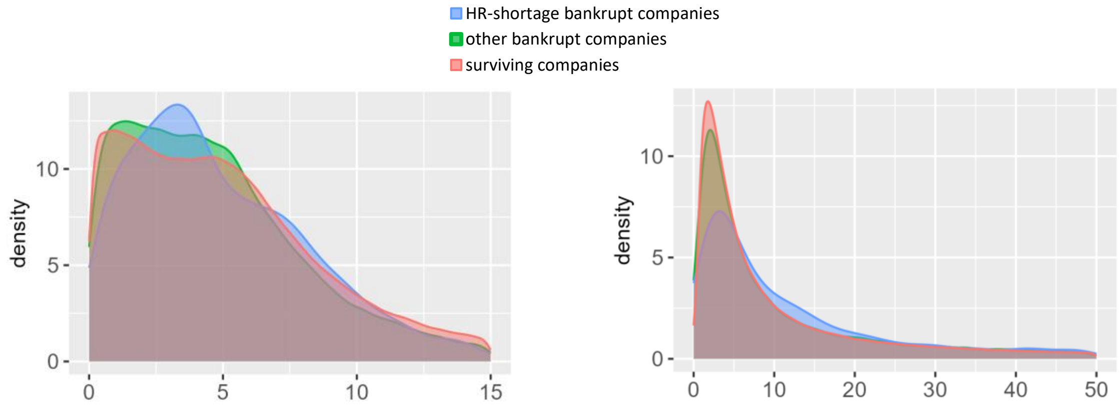
Figure 9: The security index
(Left: liquidity ratio(%)[14]; Right: degree of dependence on debt (%)[15])



	median	mean
HR-shortage bankrupt companies	64.13	66.26
other bankrupt companies	57.79	56.00
surviving companies	40.12	40.67

	median	mean
HR-shortage bankrupt companies	133.86	158.59
other bankrupt companies	131.76	157.26
surviving companies	168.13	210.75

Figure 10: The productivity index
(Left: labor productivity (millions of yen)[16];
Right: tangible fixed assets turnover (number of turnings)[17])



	median	mean
HR-shortage bankrupt companies	3.93	4.56
other bankrupt companies	3.90	4.50
surviving companies	4.50	5.44

	median	mean
HR-shortage bankrupt companies	13.31	61.00
other bankrupt companies	8.71	38.53
surviving companies	8.51	80.76

3. Comparison of trends in financial Indicators by bankruptcy status

Next, the trends in financial indicators up to bankruptcy are analyzed for HR-shortage bankrupt companies and other bankrupt companies. This section used the changes in the past five fiscal years

[14] current assets ÷ current liabilities × 100
[15] (short-term loans + long-term loans + interest expense - employee deposit) ÷ (total assets + amount notes receivable discounted + amount notes receivable endorsed) × 100
[16] value added/number of employees
[17] sales/tangible fixed assets

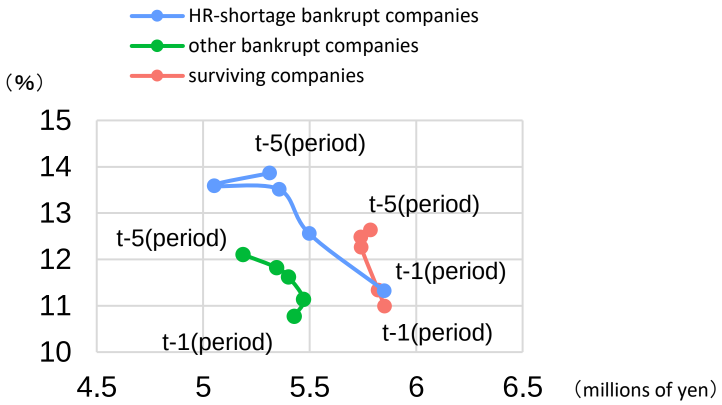
up to the previous period of bankruptcy (from t-5 to t-1) when the year of bankruptcy is set as period "t". For reference, the figures for surviving companies are also included by defining the latest financial year available as the "t-1 period." The number of verified samples by industry breakdown is shown in Table 2, although industry is not taken into account in this section.

Table 2: The number of samples [18]

	n	manufacturing	wholesale	construction	service	retall	transportation	real estate	other
HR-shortage bankrupt companies	406	15.8%	29.3%	17.7%	18.5%	8.9%	4.7%	4.9%	0.2%
other bankrupt companies	11,557	19.7%	28.0%	15.3%	20.6%	8.5%	3.5%	3.8%	0.5%
surviving companies	182,214	16.5%	24.0%	13.7%	25.8%	7.9%	4.2%	7.1%	0.9%

Figure 11 shows labor productivity on the horizontal axis and the labor cost over net sales ratio on the vertical axis. Compared to other bankrupt companies, the labor productivity of HR-shortage bankrupt companies in the t-5 period was not low, and the labor cost over net sales ratio was not high. On the other hand, the decline in labor productivity and the increase in labor cost over net sales ratio may be larger than those of the other bankrupt companies when observed over the span up to the t-1 period.

Figure 11:Trend of productivity and labor cost
(Median; X axis: labor productivity; Y axis: labor cost over net sales ratio)



Focusing on labor productivity (value added divided by number of employees), which has shown a larger declining trend than other bankrupt companies, Figures 12 and 13 show the trend of labor productivity by decomposing it by sales and property, plant and equipment. Although it is necessary to pay attention to the large standard error in sales per worker for HR-shortage bankrupt companies, it was confirmed that the decline in sales per worker may be larger and the labor equipment ratio is

[18] The period covered is April 2001 to March 2023.

lower than that of other bankrupt companies.

Figure 12: Decomposition by sales (median）
(Left: sales per worker (millions of yen)[19]; Right: value added ratio (%)[20])

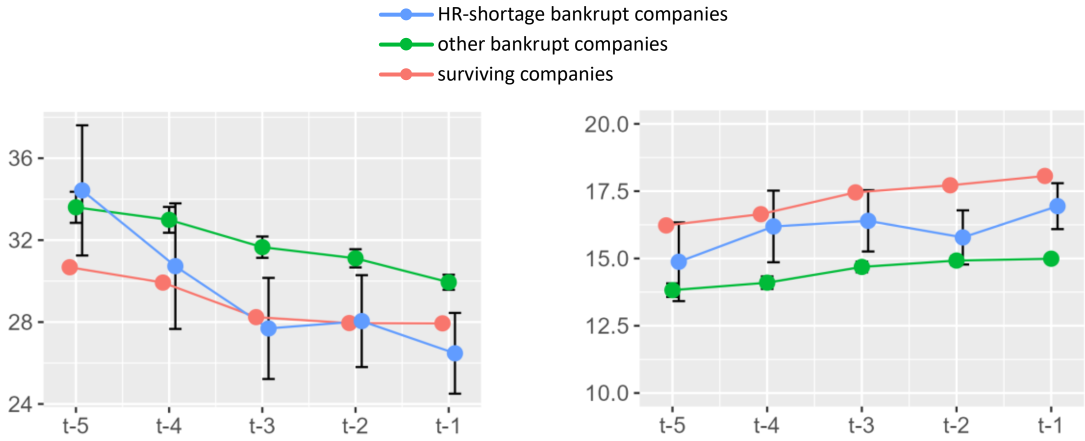
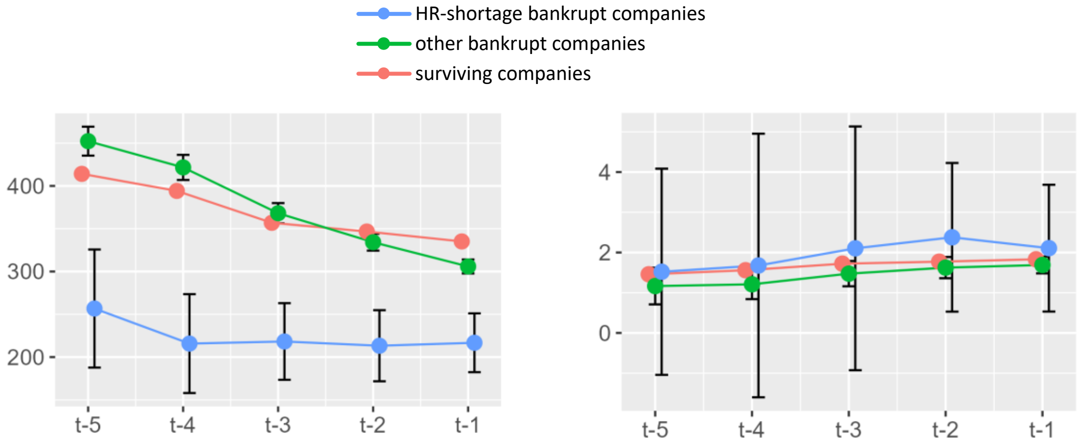


Figure 13: Decomposition by property, plant and equipment (median)
(Left: labor equipment ratio (%)[21]; Right: equipment productivity (millions of yen)[22])



[19] sales ÷ number of employees
[20] value added ÷ sales × 100
[21] property, plant and equipment ÷ number of employees × 100
[22] value added ÷ property, plant and equipment

Figure 14: Degree of dependence on borrowings (%; median)

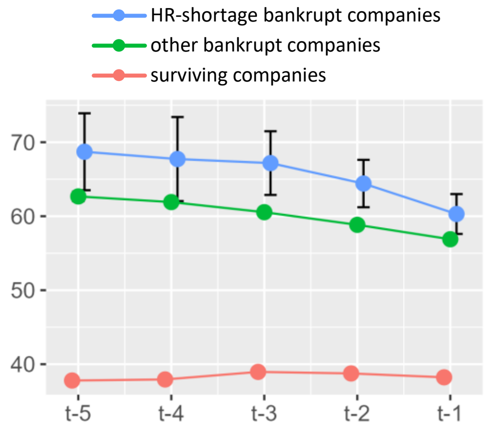
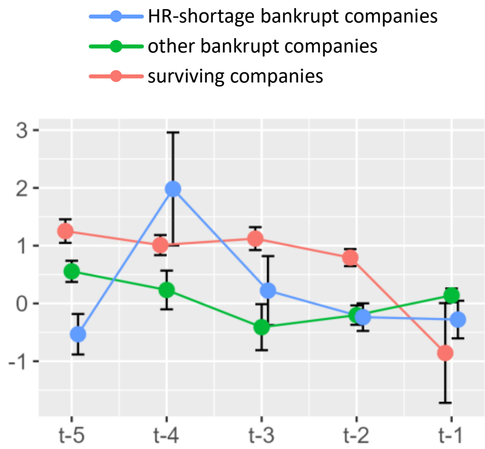


Figure 15: Change in the number of employees (mean)



(X axis: period, Error bars: standard error)

Figure 14 shows the degree of dependence on borrowings and Figure 15 shows the change in the number of employees from the previous quarter. The degree of dependence on borrowings by HR-shortage bankrupt companies remained high. No trend of increase or decrease in the number of employees from the previous period could be observed for the HR-shortage bankrupt companies.

4. Estimation of financial characteristics of HR-shortage bankruptcies

It was suggested that HR-shortage bankrupt companies are subject to progressively heavier labor cost burdens, are more dependent on borrowing, and may have inadequate operational efficiencies. To examine these trends, the following multinomial logistic regression is conducted.

$$\begin{cases} y_{B,i} = ln\left(\frac{P_{B,i}}{P_{A,i}}\right) = \alpha_B + \sum_{m=1}^{4} \beta_{B,m} explanatory\ \ variable_{m,i} + \sum_{n=1}^{4} \beta_{B,n} control_{n,i} + \varepsilon_{B,i} \\ y_{C,i} = ln\left(\frac{P_{C,i}}{P_{A,i}}\right) = \alpha_C + \sum_{m=1}^{4} \beta_{C,m} explanatory\ \ variable_{m,i} + \sum_{n=1}^{4} \beta_{C,n} control_{n,i} + \varepsilon_{C,i} \end{cases}$$

$$\left(\begin{array}{l} A: human\ \ resource\ \ shortage\ \ bankrupt \\ B: other\ \ bankrupt \\ C: surviving \\ i: company \end{array}\right)$$

Table 3: Explanatory variables and controls

<Explanatory variables>

regression 1 (v1)

(1) ⊿labor productivity(※a)

(2) ⊿labor cost over net sales ratio(※a)

(3) labor equipment ratio

(4) degree of dependence on borrowings

(※a) difference from the previous period

regression 3 (v3)

(1) labor productivity

(2) labor cost over net sales ratio

(3) labor equipment ratio

(4) degree of dependence on borrowings

regression 2 (v2)

(1) ⊿sales per worker(※b)

(2) ⊿labor cost per worker(※b)

(3) labor equipment ratio

(4) degree of dependence on borrowings

(※b) difference from the previous period of the natural logarithm

regression 4 (v4)

(1) sales per worker(※c)

(2) labor cost per worker(※c)

(3) labor equipment ratio

(4) degree of dependence on borrowings

(※c) natural logarithm

<Controls>

(1) ratio of ordinary income to net sales

(2) size (the natural logarithm of the current period capital)

(3) year dummies (the period covered is March 2002 to March 2023)

(4) industry dummies (manufacturing, wholesale, construction, service, retail, transportation and communications, real estate, others)

The estimation equations used consist of a regression equation ($y_{B,i}$) that examines the characteristics of other bankrupt companies relative to HR-shortage bankrupt companies and a regression equation ($y_{C,i}$) that examines the characteristics of surviving companies relative to HR-shortage bankrupt companies. The list of variables for each regression equation is shown in Table 3, and four regressions v1 through v4 were conducted for each variable. The samples used in the estimation are the same as those shown in Table 2.

For the explanatory variables $(explanatory\ \ variable_{m,i})$, in v1, the difference in labor productivity from the previous period ($\varDelta labor\ \ productivity$) is used to confirm the transition in productivity, and the difference in the labor cost over the net sales ratio from the previous period ($\varDelta labor\ \ cost\ \ over\ \ net\ \ sales\ \ ratio$) is used to confirm the transition in the labor cost burden. For the explanatory variables in v2, the differences from the previous period in the natural logarithm transformations of sales per worker and labor cost per worker ($\varDelta sales\ per\ worker$, $\varDelta labor\ \ cost\ \ per\ \ worker$) are used for the same purpose. For v 3 and v4, the characteristics of the levels are verified by including the figures themselves for the relevant period instead of the difference from the previous period, respectively. In addition, in each estimation equation, the labor equipment ratio and the degree of dependence on borrowings, which have shown distinctive features in the previous analyses, are also included as explanatory variables. Furthermore, return on sales, size (the natural logarithm of the current period capital), year dummy and industry dummy are used as control variables ($control_{n,i}$), where $\varepsilon_{B,i}$ and $\varepsilon_{C,i}$ represent error terms.

5. Estimation results

Table 4 shows the estimation results of the multinomial logistic regression. Since the comparison is done against HR-shortage bankrupt companies, a positive (negative) coefficient for an explanatory variable means that HR-shortage bankrupt companies are negatively (positively) correlated with the corresponding explanatory variable relative to other bankrupt companies or surviving companies.

The results of v1 and v3, which focus on productivity and labor cost burden, show that the positive correlation in $\varDelta labor\ productivity$ and the negative correlation in $\varDelta labor\ \ cost\ \ over\ \ net\ \ sales\ \ ratio$ for other bankrupt companies and surviving companies were significant. It indicates that HR-shortage bankrupt companies had worsened productivity and faced with higher labor cost burdens than other bankrupt companies. The negative correlation was also significant for the labor costs over net sales ratio, indicating that the ratio was higher for HR-shortage bankrupt companies. In addition, a significant positive correlation was found for labor productivity for the other bankrupt companies, but no significant difference was found for the surviving companies. This may be due to the fact that surviving companies include firms that have been in business for a short period of time, which is generally considered to have low labor productivity.

The results of v2 and v4 focus on the breakdown of the labor cost to sales ratio. The positive correlation for $\varDelta sales\ \ per\ \ worker$ and the negative correlation for $\varDelta labor\ \ cost\ \ per\ \ worker$ for both other bankrupt companies and surviving companies were significant, suggesting that the increase in the labor cost over net sales ratio for HR-shortage bankrupt companies was caused by both a decrease in sales per employee and an increase in labor cost per employee. In addition, the signs of the coefficients indicate a low labor equipment ratio and a high degree of dependence on borrowers for HR-shortage bankrupt companies, which may be due to the immature automation, the burden of debt repayment, and the negative impact on cash flow.

Table 4: Results of logit estimation

	v1				v3			
	other bankrupt companies		surviving companies		other bankrupt companies		surviving companies	
	coefficient	Std.Error	coefficient	Std.Error	coefficient	Std.Error	coefficient	Std.Error
Constant	2.692	0.75 ***	5.898	0.74 ***	3.198	0.75 ***	6.347	0.74 ***
⊿labor productivity	0.170	0.03 ***	0.237	0.03 ***				
⊿labor cost over net sales ratio	-10.715	1.68 ***	-8.338	1.67 ***				
labor productivity					0.049	0.02 **	-0.012	0.02
labor cost over net sales ratio					-3.465	0.33 ***	-1.937	0.32 ***
labor equipment ratio	0.037	0.01 ***	0.033	0.01 ***	0.027	0.01 ***	0.030	0.01 ***
degree of dependence on borrowings	-0.004	0.00 ***	-0.030	0.00 ***	-0.003	0.00 ***	-0.030	0.00 ***
pseudo-R2	0.35				0.35			

	v2				v4			
	other bankrupt companies		surviving companies		other bankrupt companies		surviving companies	
	coefficient	Std.Error	coefficient	Std.Error	coefficient	Std.Error	coefficient	Std.Error
Constant	2.751	0.75 ***	5.910	0.74 ***	1.438	0.78 †	5.336	0.76 ***
⊿sales per worker	1.558	0.20 ***	1.850	0.20 ***				
⊿labor cost per worker	-1.144	0.19 ***	-0.600	0.19 **				
sales per worker					0.479	0.08 ***	0.192	0.08 *
labor cost per worker					-0.150	0.08 †	0.047	0.08
labor equipment ratio	0.036	0.01 ***	0.032	0.01 ***	0.030	0.01 ***	0.030	0.01 ***
degree of dependence on borrowings	-0.004	0.00 ***	-0.030	0.00 ***	-0.003	0.00 ***	-0.030	0.00 ***
pseudo-R2	0.35				0.35			

***, **, * and † indicate significance at the 0.1%, 1%, 5%, 10% levels

6. Interpretation of analysis results and conclusions

The graphs and estimation results in this report suggest that HR-shortage bankrupt companies have different characteristics from other companies in terms of labor cost burdens and operational efficiency. First, the labor cost burden of HR-shortage bankrupt companies gradually became heavier as they approached bankruptcy, as indicated by the change in the labor costs over net sales ratio. The change in the labor cost over net sales ratio is driven by a decrease in sales per worker and an increase in labor costs per worker. There are two possible reasons for the decrease in sales per worker. The first is the quality of human resources, such as the fact that newly hired personnel are not sufficiently productive, even though they have been hired to solve the human resource shortage. The second factor is the inability to hire new workers despite the shortage, which has led to a decline in the utilization ratio. In addition to wage increases, the increase in labor costs per worker may also be an indication of a move to secure human resources by offering high compensation. However, in light of the fact that the company ultimately went bankrupt, it is possible that even if such measures to secure human resources were taken, the aforementioned factors may have prevented the company from securing sufficient sales.

The low labor equipment ratio and high tangible fixed asset turnover of the HR-shortage bankrupt companies may have resulted from a relatively immature automation and inadequate improvement in operational efficiency. In addition, given the fact that HR-shortage bankrupt companies tend to be highly dependent on borrowings even prior to their bankruptcy, it is possible that they have not been able to make investment to improve operational efficiency due to cash flow challenges.

In interpreting this analysis, some points should be noted. The sample size of HR-shortage bankrupt companies is very limited, making it difficult to conduct a detailed analysis by industry, by the nature of the personnel shortage, or by region. In the estimation equation, it is possible that there are factors other than the variables used in this study, such as the quality of management and the number of years in business that affect the human resource shortage bankruptcies, and further refinement of the model equation remains an issue for future study.

IV. Conclusion

In this paper, an analysis to understand the financial characteristics of firms that went bankrupt due to a lack of human resources are conducted, given that the labor situation in Japan has changed significantly. By continuing the analysis using a wide range of data, including trends in corporate finances, the FSA will continue to monitor, in a forward-looking manner, the impact of changes in the macro environment on the financial system, such as whether the changing labor situation leads to an increase in credit risk for financial institutions.