BOJ Reports & Research Papers

Financial System Report Annex Series

Financial System Report – Annex

A Forecast Model for the Probability of Default Based on Granular Firm-Level Data and Its Application to Stress Testing

FINANCIAL SYSTEM AND BANK EXAMINATION DEPARTMENT
BANK OF JAPAN

MAY 2019

Financial System Research Division,
Financial System and Bank Examination Department, Bank of Japan
post.bsd1@boj.or.jp

Background

The Bank of Japan publishes the *Financial System Report* semiannually, with the objective of assessing the stability of Japan's financial system from a macroprudential perspective and facilitating communication with concerned parties on relevant tasks and challenges in order to ensure such stability. The *Report* provides a regular and comprehensive assessment of the financial system.

The *Financial System Report Annex Series* supplements the *Financial System Report* by providing more detailed analysis and additional investigations on a selected topic on an ad-hoc basis.

Abstract

This *Annex* presents a model for forecasting the probability of default (PD) based on granular firm-level data. The model is intended for practical use by financial institutions, particularly for stress testing. The model comprises the following parts: (1) a main model that explains PD in terms of financial indicators of individual firms and (2) satellite models that specify relationships between macroeconomic indicators and financial indicators of individual firms. In the main model, the key financial indicator is the interest coverage ratio (ICR, an indicator of firms' debt repayment capacity), which is an efficient variable integrating firms' financial information on their solvency. The model is simple and tractable while also having high explanatory power with regard to actual firm defaults in the past. Therefore, the model is considered highly applicable to financial institutions' practices. This *Annex* applies the model to simple stress testing exercises based on two scenarios: a deterioration in the economy and a rise in interest rates.

In recent years, Japan's financial institutions have been actively lending to domestic middle-risk firms and overseas firms. Thus, it is becoming more important to accurately assess the stress resilience of financial institutions through detailed analyses of borrowing firms' creditworthiness and loan quality. The Bank of Japan's Financial System and Bank Examination Department has been collaborating with financial institutions in efforts to improve credit risk analysis and stress testing through increased use of granular data. This *Annex* is one of the outcomes. Using the analytical results in this *Annex,* the Bank of Japan will continue such efforts through a close exchange of views and information with financial institutions.

1. Introduction

Since the Global Financial Crisis, financial authorities in major countries have increasingly utilized granular data (e.g., data on individual borrowers and on individual financial products) for stress testing. The motivation for this is that the resilience of individual financial institutions in the event of macroeconomic stress may differ depending on their asset portfolios and risk profiles even if there are no significant changes in the aggregate amount of credit. In particular, there is considerable heterogeneity in firms' credit risk reflecting differences in their financial soundness. Consequently, it is necessary to carefully examine the creditworthiness of borrowing firms and to accurately assess the quality of loans. In recent years Japan's financial institutions have been actively lending to domestic middle-risk firms and to overseas firms that have different risk profiles from domestic firms. Under these circumstances, it is becoming more important to take account of borrowing firms' heterogeneity in the assessment of credit risk.

Against this background, the Financial System and Bank Examination Department of the Bank of Japan is working in close cooperation with financial institutions to improve credit risk analysis and stress testing through increased use of granular firm-level data. This *Annex*, as part of the outcomes of such efforts, presents a forecast model for the probability of default (PD) using granular data. Specifically, this *Annex* develops a new model for forecasting borrowing firms' PD, taking advantage of the interest coverage ratio (ICR), which is an efficient indicator integrating important information on firms' debt repayment capacity. This model is then estimated using the Credit Risk Database (CRD), a large-scale database storing firm-level financial indicators.[1] This *Annex* attempts to develop a model that can well explain the actual default rates in the past while at the same time being simple and tractable, keeping in mind its application to stress testing at financial institutions. This *Annex* applies the developed model to simple macro stress testing exercises examining the potential impact of plausible macroeconomic shocks (a deterioration in the economy and a rise in interest rates) on financial institutions' capacity to absorb losses.

The remainder of this *Annex* is organized as follows. Section 2 provides an overview of the database used and the hypothetical portfolio constructed from it for the estimation. Section 3 describes in detail the structure of the newly constructed model for forecasting firm defaults, touching on its differences from the previous models reported in the literature. Section 4 demonstrates the use of this model through macro stress testing exercises. Finally, Section 5 concludes and discusses future applications of this model.

[1] The Financial System and Bank Examination Department of the Bank of Japan is grateful to the CRD Association (a general incorporated association) for providing the database.

2. Outline of the portfolio for estimation

This section provides an overview of the database used for estimation and describes the characteristics of the sample of firms' financial statements constructed from it. Information on firm defaults and financial indicators of individual firms are obtained from the database for Corporate Credit Scoring Models included in the Credit Risk Database (CRD) administered by the CRD Association. The CRD stores financial data of small and medium-sized enterprises (SMEs) that are client firms of private and government-affiliated financial institutions and of Credit Guarantee Corporations participating as CRD members.[2] The CRD integrates the credit portfolios of these participating financial institutions. The number of observations of SME financial statements stored is the largest in Japan. Specifically, the database stores firm-level financial statement data from fiscal 1998 onward. As at the end of March 2017, the database covers a total of about 2.4 million corporate enterprises and contains about 18 million financial statements. In this *Annex*, these are filtered under certain conditions and then a model for forecasting PD is estimated using the extracted financial statements of firms to be analyzed (in the following, for convenience, these extracted firms are referred to as the "portfolio for estimation").

Specifically, the portfolio for estimation comprises the financial statements (about 7.8 million in total) of SMEs (about 1.3 million in total) from fiscal 2001 to 2016. However, financial statements having missing financial indicators necessary for the analysis are excluded from the portfolio for estimation. Industries covered in the analysis are all industries except agriculture, forestry, and fisheries as well as government services. In the estimation presented below, observations are classified into nine industries: construction, manufacturing (chemical and materials), manufacturing (machinery and equipment), manufacturing (other), infrastructure, wholesale, retail, real estate, and services. This classification ensures a sufficient number of observations within each industry, while it allows for heterogeneity across industries in terms of default characteristics and financial condition.

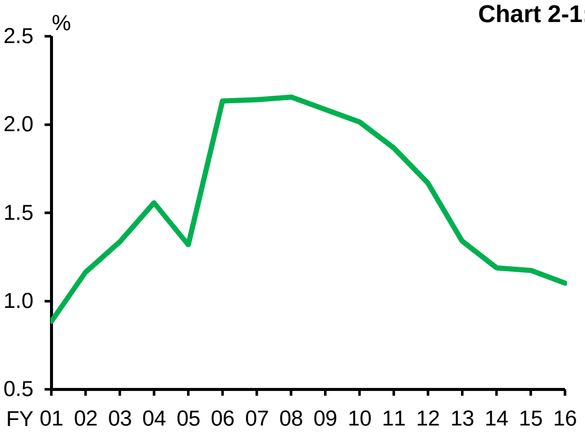
The analysis focuses on SMEs excluding "small enterprises (whose regular workforce does not exceed 5 persons)" as defined in the Small and Medium-sized Enterprise Basic Act and the Basic Act for the Activation of Small-sized Enterprises. Such small enterprises are excluded from the portfolio for estimation because their data are quite noisy. For example, it has been pointed out that for such small enterprises, it is difficult to distinguish between the enterprises' assets/earnings and the entrepreneurs' assets/earnings. In addition, developments in the financial indicators tend to be affected by the runup toward the closure of business due to such factors as the aging of the entrepreneurs and/or the lack of a successor, factors that are not necessarily related either to a deterioration in the economy or to a rise in interest rates. Exceptions are made only for the real estate industry, for which such small enterprises are included in the portfolio for estimation. There are a few reasons for doing this. First, the number of employees per firm is relatively small for the real estate industry overall. Consequently, the proportion of firms classified as small enterprises is much larger in the real estate industry than in other industries. If these firms were excluded, the number of observations for the industry would be extremely small. Second, the average amount of borrowing per firm is larger in the real estate industry than in other industries, so that even small enterprises possibly have a relatively large impact on the credit costs of financial institutions. It should be noted that the model and the analysis described below can in principle be employed for large enterprises as well, although it is employed for a large-scale database of SMEs in this *Annex*, as described above.

[2] As of April 2018, there were a total of 114 financial institutions, consisting of five major banks (including three megabanks), 66 regional banks, 40 *shinkin* banks, credit unions, etc., and three government-affiliated financial institutions, as well as Credit Guarantee Corporations, for which data on client SMEs are available.

In this *Annex*, firm defaults are defined as being *for the first time* (1) downgraded to the borrower classification "special attention" or below, (2) overdue by more than 3 months, and/or (3) subrogated by Credit Guarantee Corporations. The default rate is calculated as the ratio of the number of firms that defaulted in the fiscal year to the total number of observations in that fiscal year. In principle, a firm may have been classified as being in default, subsequently been upgraded, and then defaulted again. However, the data for this *Annex* provided by the CRD Association do not include upgrade information for firms that have defaulted. Consequently, in the portfolio for estimation, once a firm has been classified as having defaulted, it is excluded for the remainder of the observation period. Thus, the default rate defined here will understate the actual default rate faced by financial institutions if some firms have defaulted multiple times, which is likely.[3]

Chart 2-1 shows the default rate for all the observations defined in this way. The chart indicates that the default rate increased in the late 2000s and reached a peak in fiscal 2008 during the Global Financial Crisis.[4] Since then, it has declined and is now at a level similar to that during the early 2000s. Chart 2-2 presents the default rate by industry. Although the longer-term patterns across industries are not that different, there are some differences in terms of the sensitivity to the macroeconomic environment. For example, whereas default rates in manufacturing (machinery and equipment) and real estate were strongly affected by the Global Financial Crisis, default rates in manufacturing (chemical and materials), wholesale, and retail were more affected by the deterioration in the terms of trade brought about by the rise in oil prices in the mid-2000s.

Chart 2-1: Default rate

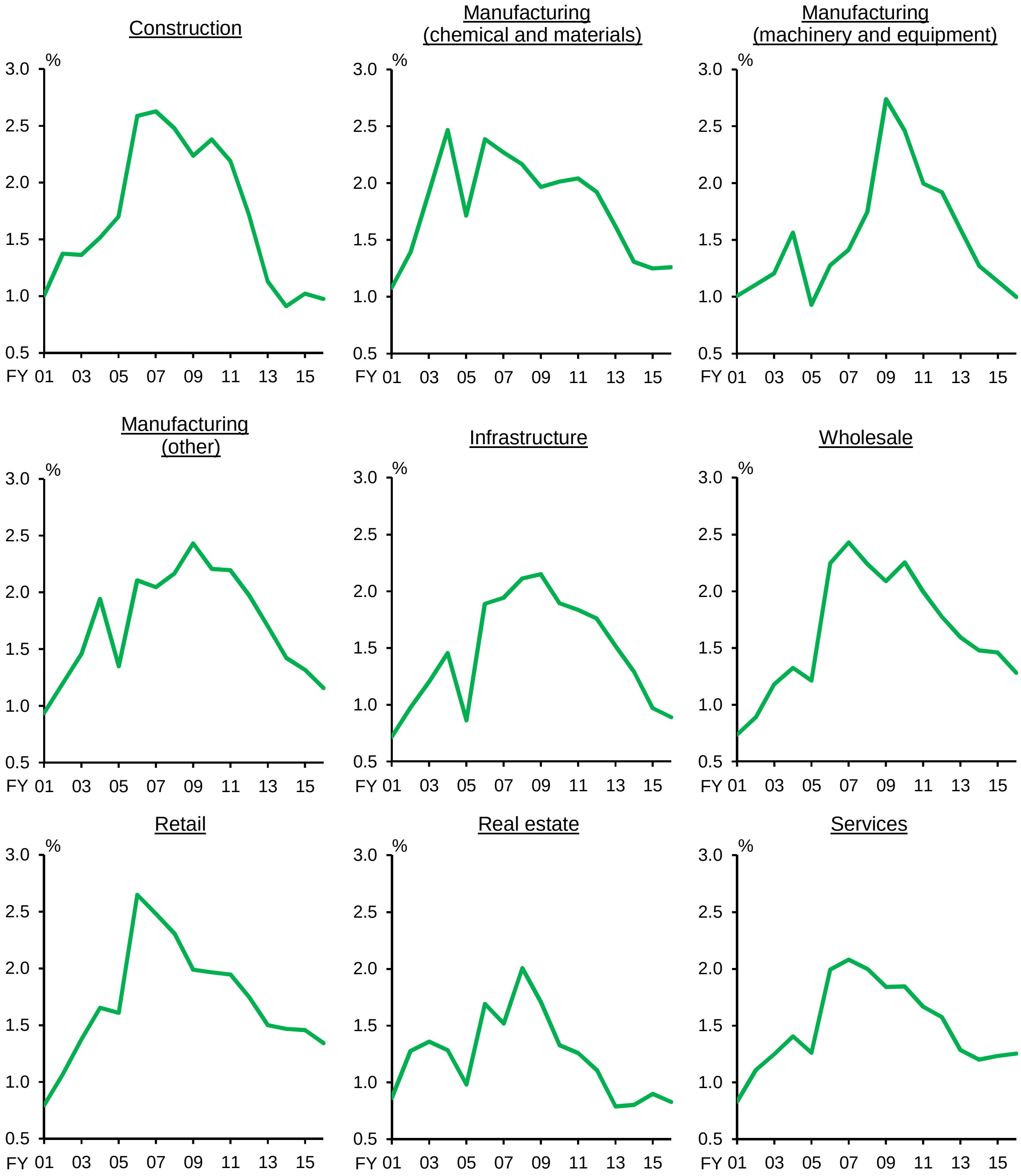


Note: Firm defaults are defined as being, for the first time, (1) downgraded to the borrower classification "special attention" or below, (2) overdue by more than 3 months, and/or (3) subrogated by Credit Guarantee Corporations. The default rate is calculated as the proportion of firms that defaulted in that fiscal year. The financial statement data of firms that have been classified as having defaulted in a particular period are excluded for the remainder of the observation period.
Source: CRD Association.

[3] However, the model described below can deal with the possibility that the same firm defaults several times. Therefore, this caveat about the portfolio for estimation in this *Annex* does not mean that it prevents financial institutions from applying the model to their own portfolio and information on actual defaults of borrowing firms.

[4] It should be noted that the default rate from that period onward may be non-negligibly affected by various policy measures, especially the implementation of the SME Finance Facilitation Act. These policy measures recommend that loans, even ones that had been restructured, not be classified as "special attention" if certain conditions are met. For details, see the April 2012 issue of the *Financial System Report.* To the extent that the default rate is affected, the macro stress testing presented below may understate the increase in PD (unless similar policy measures are implemented in the event of the future stress event).

Chart 2-2: Default rate by industry



Note: For the calculation method of the default rate, see the note on Chart 2-1. Latest data as at fiscal 2016.
Source: CRD Association.

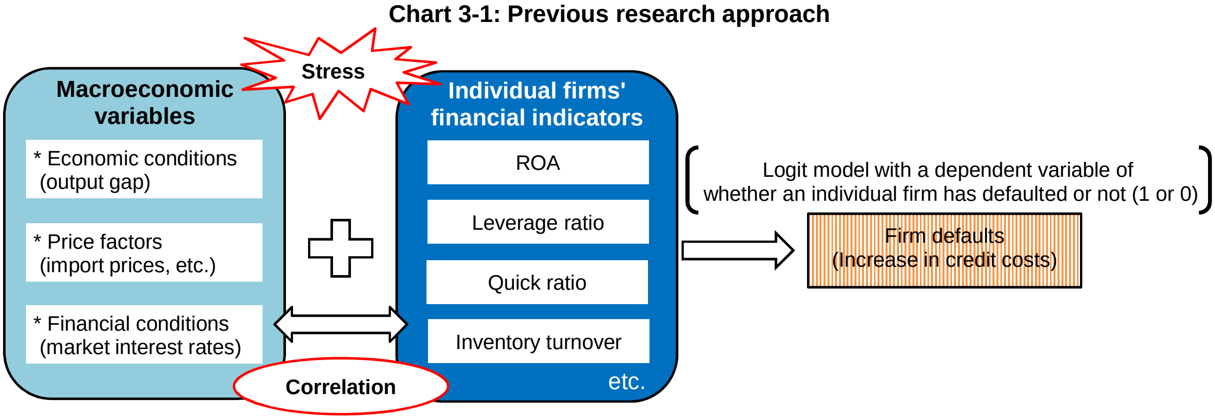
3. Outline of the forecast model for PD

This section provides an overview of the model constructed in this *Annex* using macroeconomic variables and firm-level financial indicators to forecast PD.

A. Overview of the forecast model for PD

Empirical studies examining firm defaults typically employ specifications that include both macroeconomic variables and financial indicators of individual firms as explanatory variables. The reason is that the macroeconomic environment, such as economic conditions and market interest rates, will clearly have a substantial impact on PD. On the other hand, firm defaults are themselves microeconomic (firm-level) phenomena. Therefore, taking account of firm-level financial indicators representing firms' solvency and liquidity -- such as their return on assets (ROA), ICR, leverage ratio, and liquidity ratio -- as well is likely to increase the predictive power.[5] Against this background, conventional empirical analyses of firms' PD typically employ so-called logit models in which the dependent variable is a discrete variable (taking a value of 1 or 0) indicating whether an individual firm has defaulted or not and the explanatory variables consist of (various combinations of) both macroeconomic variables and firm-level financial indicators (Chart 3-1).

Chart 3-1: Previous research approach

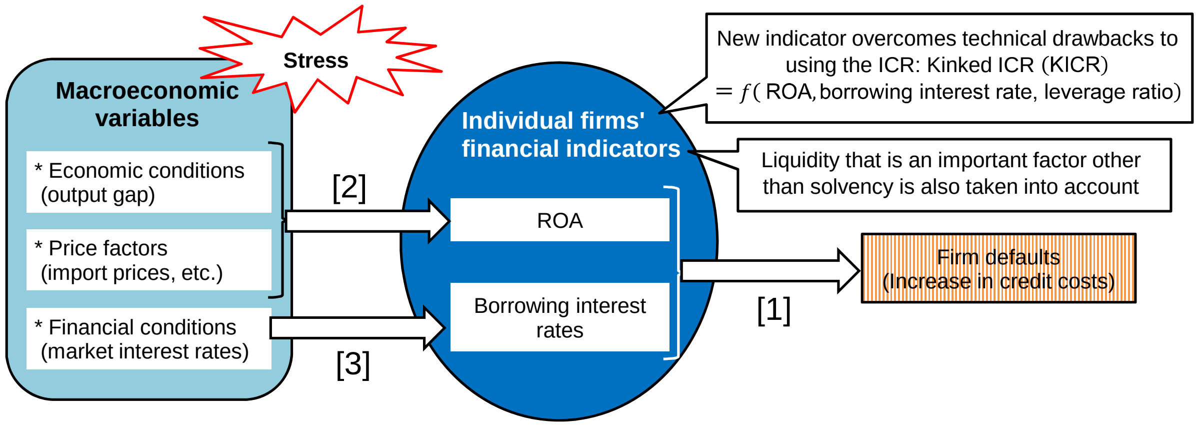


However, there are a few issues associated with logit models including both macroeconomic variables and financial indicators of individual firms as explanatory variables. First, long-term time-series data are often not available for financial indicators of individual firms. This is especially the case for financial information on SMEs, which account for the largest part of the loan portfolios of financial institutions. Second, firms' business performance and activities are affected by the macroeconomic environment -- for example, a deterioration in the output gap will push down firm profitability -- so that there will be considerable correlation between macroeconomic variables and financial indicators of individual firms. This issue, which is referred to as "multicollinearity," makes it difficult to obtain statistically reliable estimates and hence to quantify the impact of each variable. Third, this model cannot by itself explicitly specify causal relationships between macroeconomic variables and financial indicators of individual firms, thus making the model difficult to use for stress testing. With regard to this third issue, the typical procedure for stress testing in terms of credit risk is to assume a particular change to the macroeconomic environment, and then examine how the change impacts the financial indicators of individual firms. The goal is to measure how much the change affects financial institutions' credit costs through firm defaults. However, since a conventional model treats macroeconomic variables and financial indicators as separate sets of

[5] On this point, see, for example, T. Jacobson, J. Lindé, and K. Roszbach, "Firm Default and Aggregate Fluctuations," *Journal of the European Economic Association*, Vol. 11(4), August 2013.

explanatory variables that are independent of each other, simulations conducted with the conventional model fail to take account of transmissions from the former to the latter.

Chart 3-2: Overview of the model



	Model	Characteristics of the model
[1]	PD and firm-level financial indicators	Constructs the KICR (Kinked ICR). Differences in sensitivity due to firm characteristics (leverage ratio) are taken into account.
[2]	Firms' ROA and macroeconomic variables	Differences in sensitivity to macroeconomic variables due to firm characteristics (average profitability) are taken into account.
[3]	Borrowing interest rates and market interest rates	Uses the long-term time series data including the periods of interest rate fluctuations by industry (obtained from the *Financial Statements Statistics of Corporations by Industry*).

Taking all of the above issues into consideration, this *Annex* constructs the following model consisting of three estimation equations in two stages (an overview of the model is provided in Chart 3-2) with a view to enhancing the applicability to risk management practices at individual financial institutions including stress testing. First, for the main model that predicts PD of borrowing firms, a specification is adopted so that it does not include macroeconomic variables but only includes financial indicators of individual firms (most importantly the ICR) as explanatory variables (Model [1]). The reason for using such a parsimonious specification to forecast PD is that it avoids multicollinearity with macroeconomic variables even when long-term time-series data are not fully available, thus making it possible to obtain sufficiently reliable estimates from cross-sectional information on firms' ICR. The ICR is not only a simple and intuitive financial variable that integrates basic information on the debt repayment capacity of firms but also, by definition, closely related to firm default. In fact, there appear to be technical drawbacks to using the ICR in its original form as an explanatory variable in the PD function. As a result, this *Annex* uses a new variable called the Kinked ICR (KICR), which avoids these drawbacks (details are described below).

The remaining two estimation equations necessary are satellite models facilitating the application to stress testing. One of these represents the link between macroeconomic variables and firm profitability (ROA, Model [2]). Changes in economic conditions (the output gap) and in price factors (such as the terms of trade or import prices) affect the profitability of individual firms. Such changes in firm profitability lead to changes in PD through changes in the numerator of the ICR (sources for debt repayment, such as operating profits), which is an explanatory variable in the main model. The other represents the link between market interest rates and borrowing interest rates (Model [3]). Changes in market interest rates are passed through to firms' borrowing interest rates, although there are some differences in the lag structure depending on the type of loan. These changes in borrowing interest rates also affect PD through changes in the denominator of the ICR (interest payments), which is an explanatory variable of the main model. As illustrated, these two satellite models explicitly specify the causal relationships between macroeconomic variables and firm profitability and between market interest rates and borrowing interest rates. This makes it possible to conduct stress testing assuming the occurrence of a macroeconomic shock such as a deterioration in the economy or a rise in interest rates. The following three subsections provide a detailed description of each of the three models.

B. Link between PD and firm-level financial indicators

This subsection describes Model [1], the main model, which represents the link between PD and firm-level financial indicators. Generally speaking, the factors that determine whether a firm defaults can be broadly divided into the firm's (1) solvency (debt repayment capacity) and (2) liquidity (short-term funding capacity). In the model here, firm solvency and firm liquidity are represented by the ICR and the liquid asset ratio (= current assets / current liabilities) respectively, which are used as explanatory variables.[6]

The ICR measures the extent to which current profits cover current interest payments. Concretely, it is defined as the sum of operating profits and interest and dividends received divided by interest payments:[7,8]

$$\begin{aligned} \text{ICR} &= \frac{\text{Operating profits} + \text{Interest and dividends received}}{\text{Interest payments}} \\ &= \frac{(\text{Operating profits} + \text{Interest and dividends received})/\text{Total assets}}{(\text{Borrowing interest rate} \times \text{Amount borrowed})/\text{Total assets}} \\ &= \frac{\text{Operating ROA}}{\text{Borrowing interest rate} \times \text{Leverage ratio}} \end{aligned}$$

As shown by the last line of the above equation, the ICR is an indicator that efficiently integrates three components determining a firm's debt repayment capacity: its profits, its borrowing interest rate, and its leverage. As such, it is a particularly useful variable for estimating the PD function when it is not possible to use a large number of explanatory variables due to a limited sample size and the possibility of multicollinearity.

However, using the ICR as an explanatory variable in the PD function has the following technical drawbacks (Chart 3-3). When operating profits in the numerator are positive, the ICR does indeed represent the extent to which current profits cover current interest payments. The ICR deteriorates (i.e., takes a smaller positive value) due to an increase in interest payments in the denominator through an increase in the firm's borrowing interest rate and/or leverage ratio, as well as due to a

---

[6] Current assets here consist of the sum of firms' cash and deposits, notes receivable-trade and accounts receivable, securities, inventories, and other current assets (such as advance payments and accounts receivable-other). Current liabilities consist of the sum of notes payable-trade and accounts payable, short-term loans payable (such as current portion of loans payable -- including corporate bonds, commercial paper, etc.-- overdrafts, and borrowings on deeds and bills), and other current liabilities (such as accounts payable-other, advances received, and deposits received from employees). These are common definitions in line with corporate accounting principles in Japan.

[7] In the remainder of this *Annex*, operating profits will include interest and dividends received.

[8] The leverage ratio here is defined as the amount borrowed divided by total assets.

decline in operating profits in the numerator. On the other hand, when operating profits in the numerator are negative, a further decrease in operating profits still results in a deterioration in the ICR (i.e., a larger negative value), but an increase in interest payments in the denominator actually results in an improvement in the ICR (i.e., a smaller negative value). Moreover, as a somewhat extreme example, if a firm's interest payments are very small (i.e., the denominator is very small) and if the firm moves from making profits to making even the tiniest of losses, its ICR would turn substantially negative.[9] Thus, using the ICR is problematic when there are changes in the sign on operating profits; that is, when operating profits turn negative, the ICR loses its intrinsic economic meaning. This is well illustrated in Chart 3-4, which shows a simple graph of a relationship between the ICR and the default rate. When operating profits are positive, the default rate rises as the ICR deteriorates; however, when operating profits are negative, the link between the two becomes unclear.[10]

Chart 3-3: Technical drawbacks to using the ICR

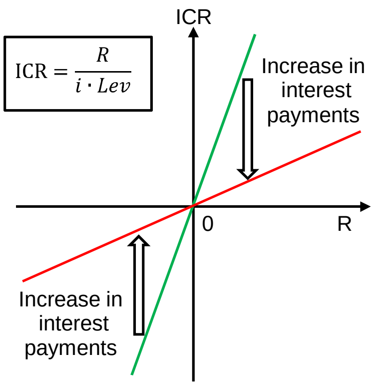


Chart 3-4: Relationship between the ICR and the default rate

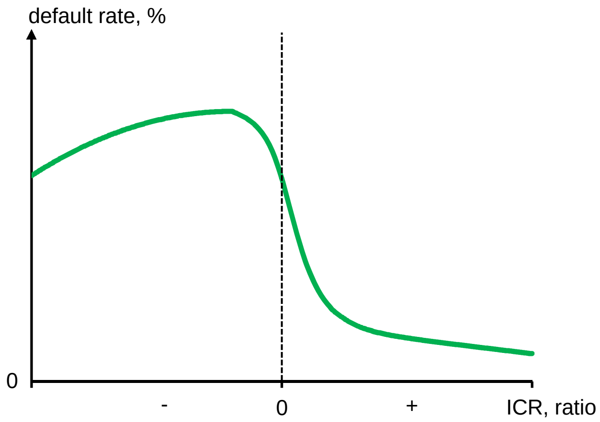
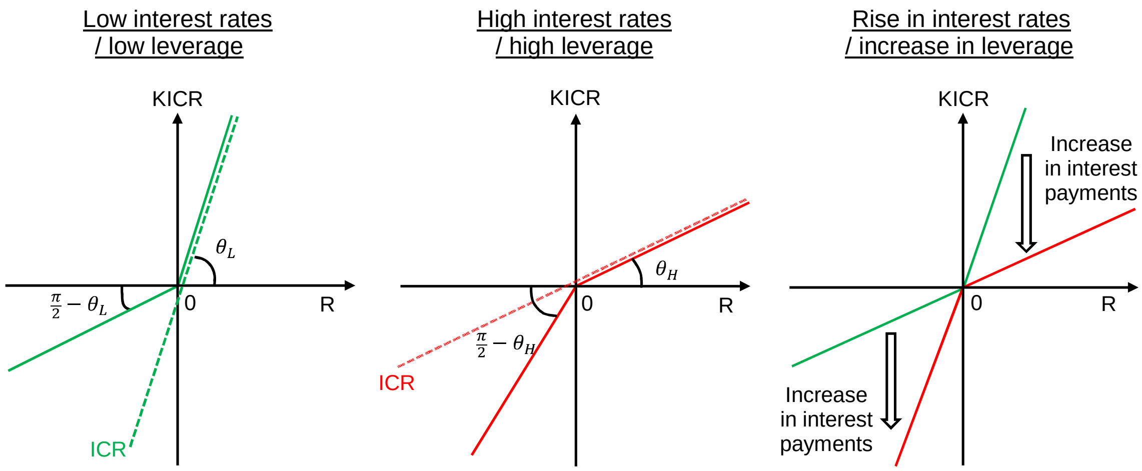


Chart 3-5: Concept of the KICR



In order to overcome this problem while still making use of the advantages of the ICR, this *Annex* develops a transformed version of the ICR for its analysis. Specifically, when a firm's operating profits are positive, the ICR is used as it is; on the other hand, when operating profits are negative,

[9] However, it is difficult to suppose that, for Japan's corporate sector as a whole, the aggregate of operating profits could be negative, so that, at an aggregate national level, this issue would not arise.

[10] The graph is generated by dividing the ICR into intervals and calculating the default rate among sample firms falling into each interval.

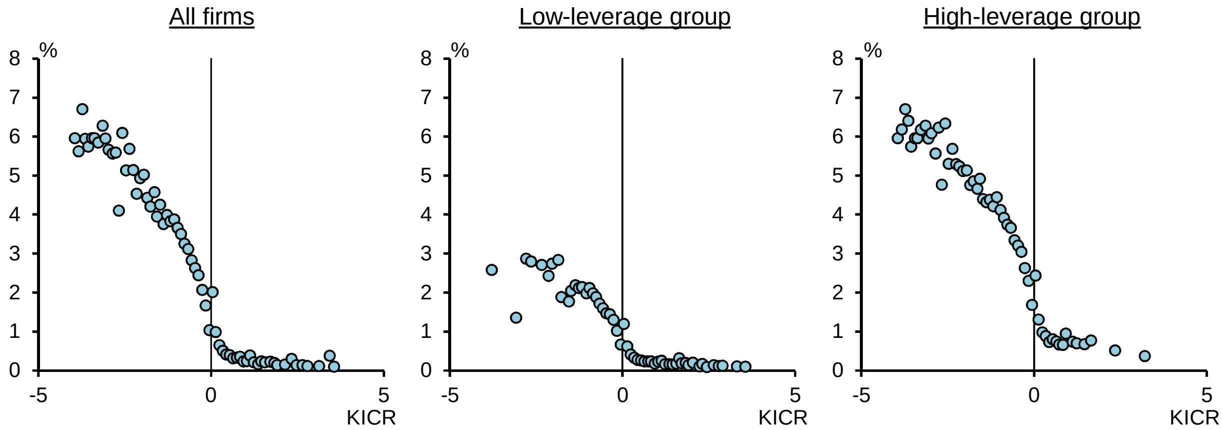
the following approach is used so that a rise in interest rates and/or leverage signals a decline in the firm's debt repayment capacity: the ICR is replaced by the product of the three components making up the ICR: operating ROA × borrowing interest rate × leverage ratio. As shown in Chart 3-5, where this new variable and the operating ROA are plotted on the vertical axis and horizontal axis, respectively, there is a kink when the operating ROA falls to zero. In this *Annex*, this variable will be referred to as the Kinked ICR (KICR) due to its shape (see Box 1 for details of the construction of the KICR). The KICR is defined as follows:

$$\text{KICR} \equiv \begin{cases} \dfrac{\text{Operating ROA}}{\text{Borrowing interest rate} \times \text{Leverage ratio}} & \text{for Operating ROA} \geq 0 \\ \text{Operating ROA} \times \text{Borrowing interest rate} \times \text{Leverage ratio} & \text{for Operating ROA} < 0 \end{cases}$$

The value of this newly constructed KICR becomes smaller (i.e., signals a decline in a firm's debt repayment capacity) when (1) operating profits deteriorate (i.e., profits become smaller or losses become larger) and/or (2) the borrowing interest rate and/or leverage ratio increase, regardless of the sign on operating profits. Thus, the KICR bypasses the technical issues that arise when firms make losses while preserving the desirable properties of the ICR.

In fact, as shown in Chart 3-6, unlike in the case of the ICR, there is a clear downward sloping relationship between the KICR and the default rate, implying that a decline in the KICR is associated with an increase in the default rate even when the KICR is negative. In other words, this suggests that firms with a negative KICR reflecting operating losses are more likely to default when these losses increase as a result of a deterioration in the economy or when interest payments increase due to a rise in interest rates. Furthermore, between the two groups of firms, those with high leverage and those with low leverage, there are clear differences in the link between the KICR and the default rate. The difference is not only in the level of the default rate but also in its sensitivity to the KICR.[11] In particular, the sensitivity of the default rate tends to be greater in the group with high leverage than in the group with low leverage when the KICR is negative. This indicates that the higher the debt level and leverage ratio of a firm are, the more likely it is to default when its KICR deteriorates due to a decline in operating profits or an increase in borrowing interest rates.

Chart 3-6: Relationship between the KICR and the default rate

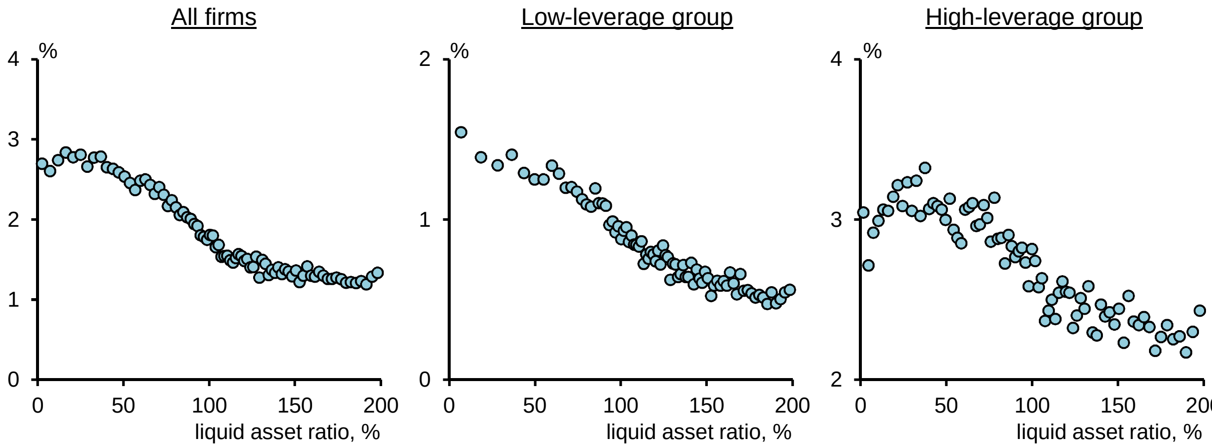


Note: The vertical axis indicates the default rate.
Source: CRD Association.

[11] Firms are divided into the high-leverage and low-leverage groups by pooling all observations (for all industries and all periods) and then splitting them at the median of the leverage ratio distribution. For this reason, on an industry-level basis, the numbers of observations in the two groups are not the same.

Chart 3-7 shows the relationship between the default rate and the liquid asset ratio, which is another important determinant of firm defaults. As in the case of the ICR (and KICR), a downward sloping relationship is observed, implying that the lower firms' liquid asset ratios are, the higher the default rate is. Again dividing firms into two groups -- those with high leverage and those with low leverage -- shows that there are clear differences in the degree of correlation between the liquid asset ratio and the default rate. In particular, a stronger negative correlation is found for the low-leverage group. This suggests that for firms with a low leverage ratio, a deterioration in their liquidity may be more likely to cause their default than a deterioration in their solvency. On the other hand, for the high-leverage group, the correlation between the liquid asset ratio and the default rate is not that strong (although there is some correlation). This is likely because, for the high-leverage group, firm solvency (represented here by the KICR) is relatively strongly related to default.

Chart 3-7: Relationship between the liquid asset ratio and the default rate



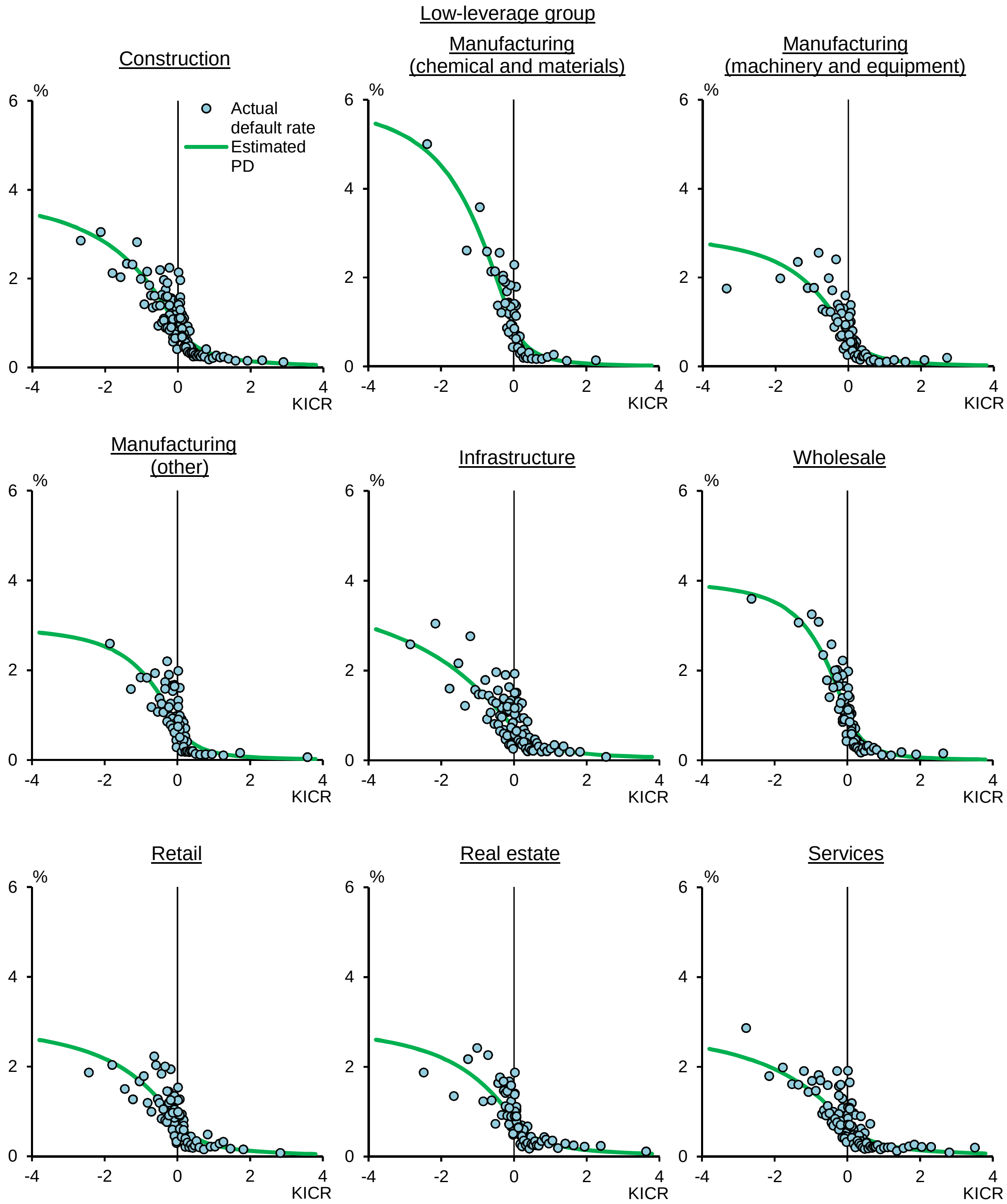
Note: The vertical axis indicates the default rate. Liquid asset ratio = current assets / current liabilities * 100.
Source: CRD Association.

Based on the above observations, the approach taken in this *Annex* is to regress PD on the KICR and the liquid asset ratio for each industry by dividing observations into the high-leverage and low-leverage groups. The PD functions are estimated with the method of least squares (for details on the estimation method, see Box 2). The estimation results indicate that the model generally shows a good empirical fit in that the estimates capture well the actual relationships between the KICR and firm defaults since the early 2000s (Chart 3-8), although there are some differences depending on the industry and the leverage group.[12] Moreover, the estimation results indicate that in all industries, PD of the high-leverage group is more sensitive to the KICR than that of the low-leverage group. This is consistent with the observations mentioned above.[13]

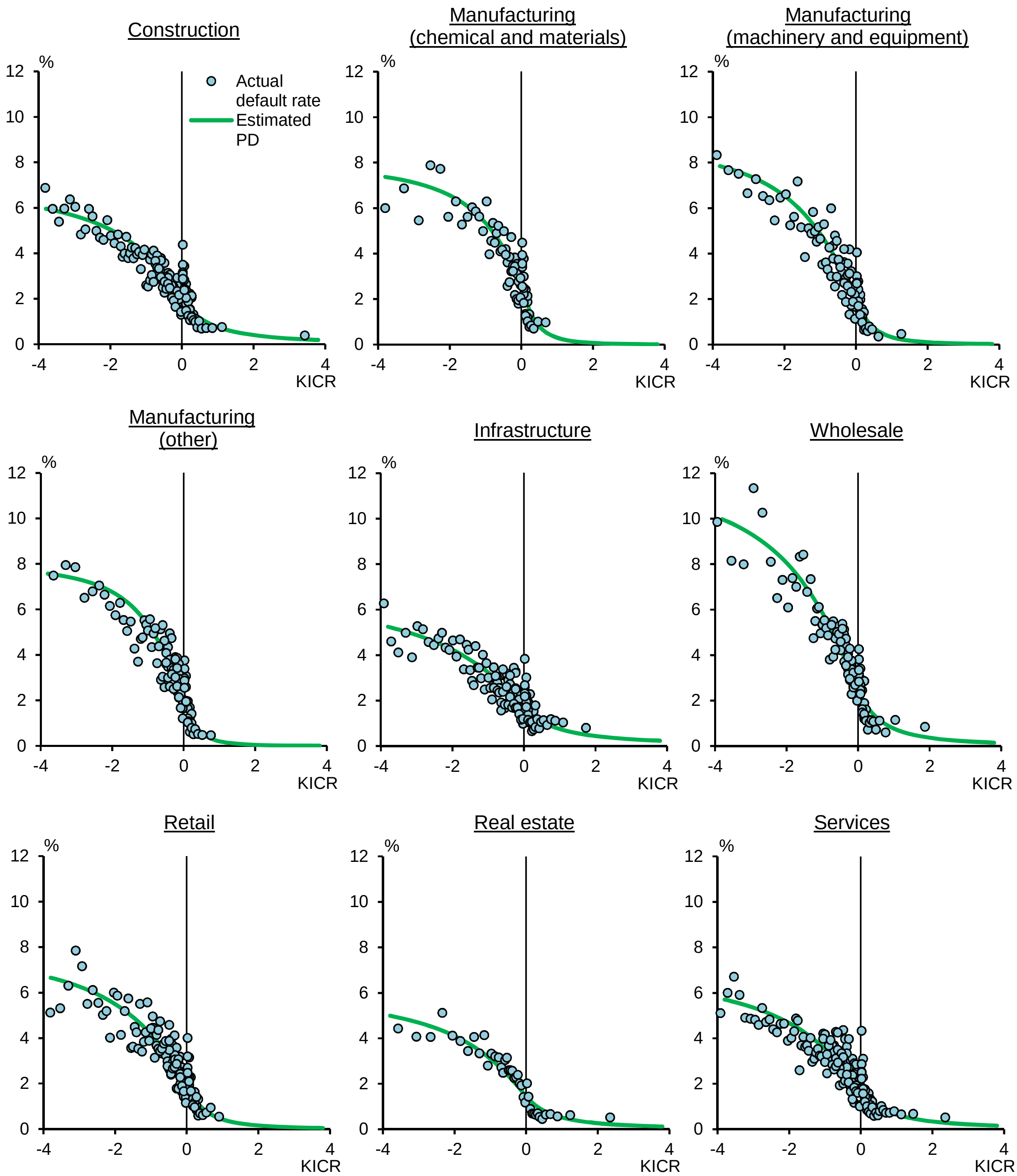
[12] Changing the definition of default -- for example, from a downgrade to "special attention" or below to a downgrade to "in danger of bankruptcy" or below -- leaves the estimation results regarding the sensitivity of PD to the explanatory variables and the fit of the estimation equation largely unchanged, especially in the case of the high-leverage group. In this sense, the forecast model for PD estimated in this *Annex* is reasonably robust to changes in the definition of defaults.

[13] The estimation results presented here are based on firm-level financial data from fiscal 2001 to 2016. The robustness of the model performance to changes in the length of the observation period is also checked by using data only for the period from fiscal 2007 to 2016. According to this robustness check, the model performs well overall regardless of the period, although shortening the observation period reduces the statistical significance of some of the estimates.

Chart 3-8: Actual default rate and estimated PD



High-leverage group

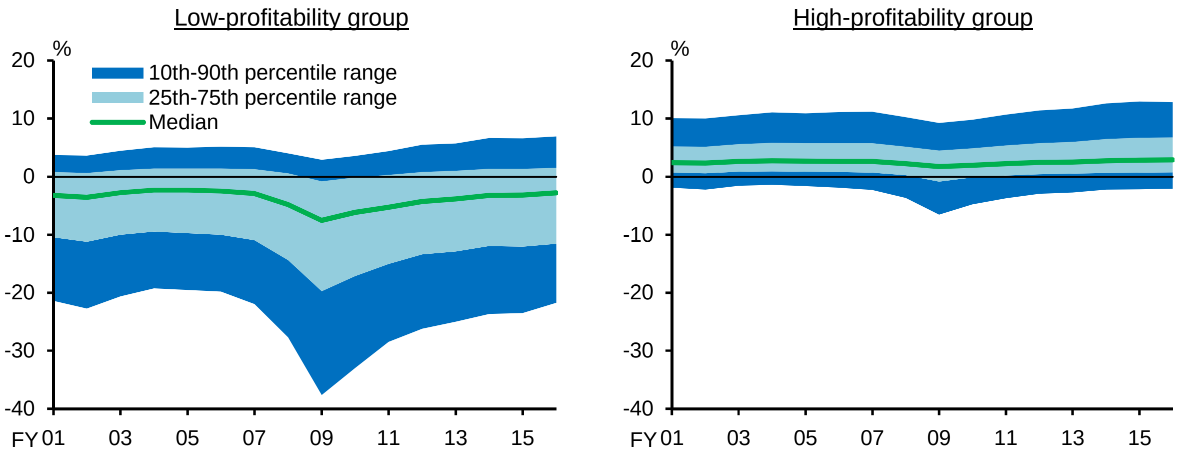


Note: The vertical axis indicates the default rate. The solid line shows the estimated PD calculated assuming the median liquid asset ratio of firms by group and industry where the liquid asset ratio is statistically significant.
Source: CRD Association.

C. Link between firms' ROA and macroeconomic variables

Next, this subsection describes the estimation approach and estimation results of the first satellite model, Model [2], representing the relationship between firms' ROA and macroeconomic variables. In Model [2], operating ROA of individual firms is the dependent variable, while the output gap (to represent developments in economic conditions) and the terms of trade or import prices (to represent price factors) are used as explanatory variables. The model is then estimated for each industry by using a panel regression technique.[14] There is not only substantial heterogeneity across firms in the dependent variable, the ROA level, but there are also considerable differences in the sensitivity to fluctuations in economic conditions (Chart 3-9). Specifically, firms with a low ROA experienced a much larger decline in their ROA during the Global Financial Crisis than other firms. Given that the sensitivity to economic conditions may differ between more profitable firms and less profitable firms, it seems desirable to take account of heterogeneity in the coefficient on the output gap when estimating the ROA function.[15]

Chart 3-9: Distribution of operating ROA



Note: Operating ROA = (operating profits + interest and dividends received) / total assets * 100. Latest data as at fiscal 2016.
Source: CRD Association.

Therefore, the following two-stage estimation is conducted for the ROA function:

(1) The average operating ROA of an individual firm is estimated as a fixed effect in a fixed effects model.[16] For each industry, firms are then divided into two groups based on the distribution of their estimated fixed effects, with those falling into the bottom quartile categorized as firms with low profitability and the rest categorized as firms with high profitability.

(2) Then, the sensitivity to macroeconomic variables is estimated for each of the two groups in each industry, using as the dependent variable the difference between firms' operating ROA and the estimated fixed effect.

Charts 3-10 and 3-11 present the estimation results for the high- and low-profitability groups by industry based on this procedure. The estimates all have the right sign and are statistically significant. Moreover, the estimates for the sensitivity to economic conditions (the coefficient on the output gap) are all larger for the low-profitability group than for the high-profitability group. This result is consistent with the observations made earlier. By industry, the coefficient on the output

[14] For manufacturing industries, the terms of trade -- defined as the ratio of export prices to import prices -- were used, while for non-manufacturing industries other than the real estate industry, import prices were used. For the real estate industry, commercial property prices (year-on-year percentage change) were used.

[15] If the estimation is conducted assuming that the sensitivity of firms' ROA to economic conditions is the same for all firms, the estimates for highly profitable firms would be biased upward, while those for less profitable firms would be biased downward. Using such biased estimates for stress testing would distort the increase in PD of individual firms.

[16] In this panel estimation, the fixed effects model has been adopted after conducting F-tests to confirm the presence of individual effects.

gap is larger in industries with a high sensitivity to economic conditions such as manufacturing (machinery and equipment) and construction than in other industries.

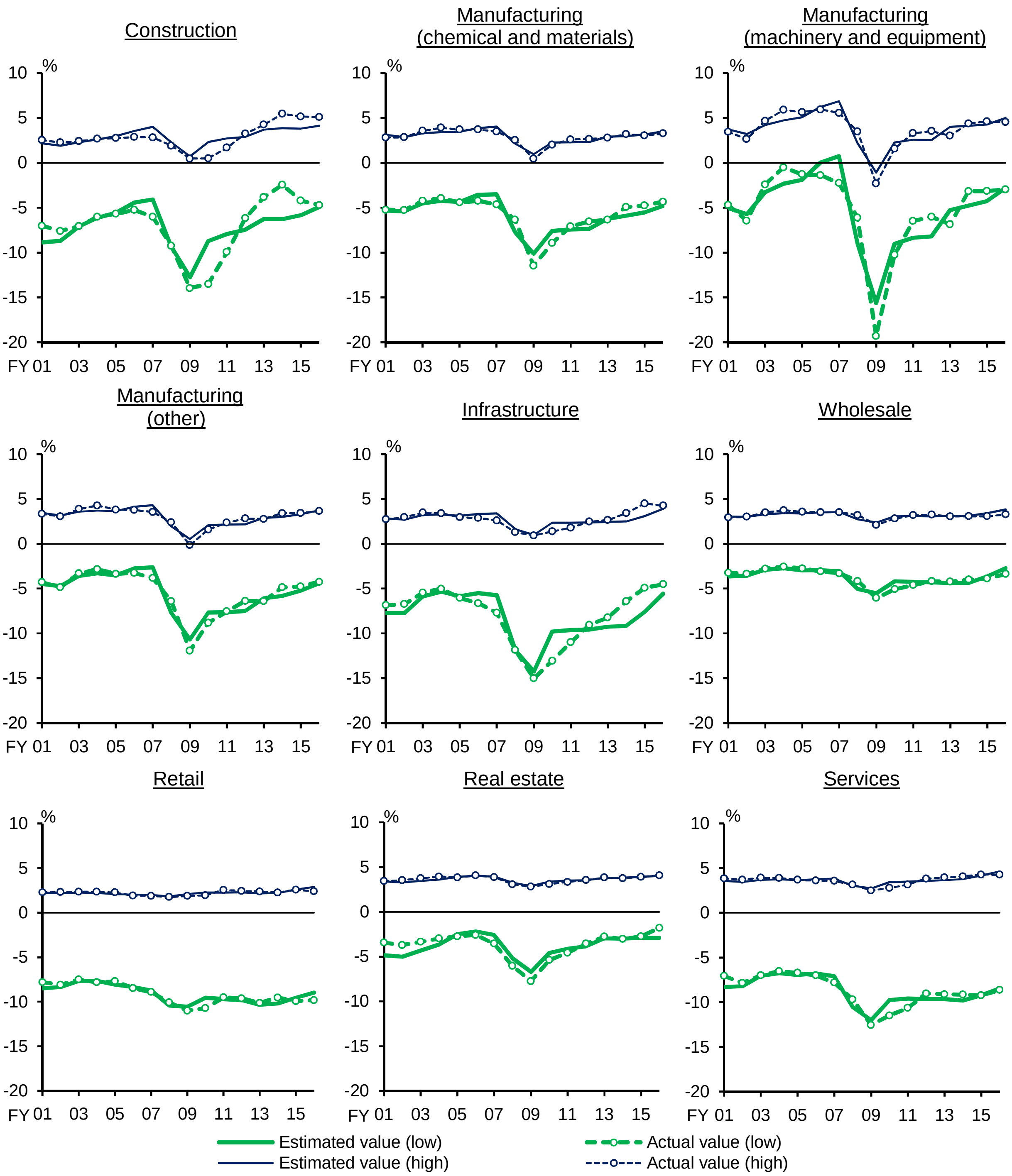
Chart 3-10: Estimates: operating ROA (1)

		Dependent variables: operating ROA								
		Construc-tion	Manufacturing			Infra-structure	Wholesale	Retail	Real estate	Services
			Chemical and materials	Machinery and equipment	Other					
		Low-profitability group								
Explanatory variables	Output gap	1.49 ***	1.05 ***	2.56 ***	1.27 ***	1.70 ***	0.56 ***	0.44 ***	0.45 ***	0.91 ***
	Price factor	-0.07 ***	0.07 ***	0.11 ***	0.08 ***	-0.15 ***	-0.06 ***	-0.07 ***	0.06 ***	-0.08 ***
Sample size (thou.)		355	80	117	220	165	220	162	219	402
		High-profitability group								
Explanatory variables	Output gap	0.57 ***	0.53 ***	1.32 ***	0.65 ***	0.57 ***	0.28 ***	0.09 ***	0.12 ***	0.28 ***
	Price factor	-0.00 ***	0.03 ***	0.05 ***	0.04 ***	-0.05 ***	-0.03 ***	-0.02 ***	0.02 ***	-0.03 ***
Sample size (thou.)		1,066	240	351	661	495	659	485	656	1,205

Note: 1. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
2. As "Price factor," terms of trade (export prices / import prices) is used for the manufacturing industry, the year-on-year percentage change in commercial property prices (nationwide) is used for the real estate industry, and import prices are used for other industries. The commercial property prices until fiscal 2007 are estimated using the indices of Tokyo, Aichi, and Osaka.

Source: CRD Association; Ministry of Land, Infrastructure, Transport and Tourism, "Japan property price index"; BOJ, "Corporate Goods Price Index."

Chart 3-11: Estimates: operating ROA (2)

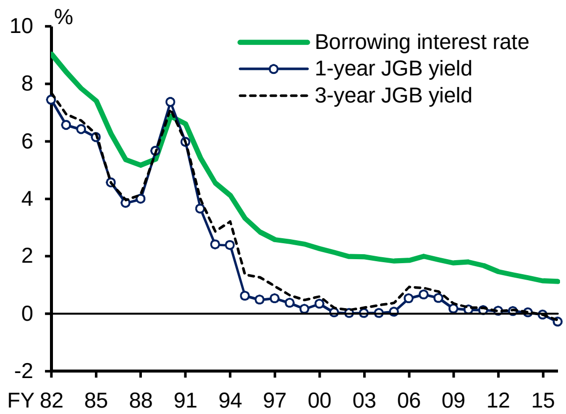


Note: "Low" and "high" indicate the low- and high-profitability groups, respectively. Latest data as at fiscal 2016.
Source: CRD Association.

D. Link between borrowing interest rates and market interest rates

Finally, this subsection describes the estimation approach and estimation results of the second satellite model, Model [3], representing the relationship between borrowing interest rates and market interest rates. In Japan, market interest rates have been following a secular downward trend -- albeit with some fluctuations -- since the late 1990s, so that lending interest rates have also been declining (Chart 3-12). Therefore, if borrowing interest rates were calculated from the CRD data, which are available only from fiscal 1998 onward, the observation period would not include enough phases of rising interest rates. That would make it difficult to obtain sufficiently reliable estimates of the sensitivity of borrowing interest rates to changes in market interest rates (i.e., the pass-through rate) that could be used for forecasting the response in a phase of rising interest rates.[17] Given this circumstance, the model is not estimated using a panel regression technique with firm-level borrowing interest rates. Instead, the approach taken in this *Annex* is calculating industry-level borrowing interest rates from the *Financial Statements Statistics of Corporations by Industry* -- which provides long-term time-series data (from fiscal 1981 to 2016) that include phases of rising interest rates -- and then using the calculated data as the dependent variable to estimate the *industry-level* sensitivity of borrowing interest rates to changes in market interest rates.[18]

Chart 3-12: Interest rates



Note: Latest data as at fiscal 2016.
Source: Ministry of Finance, "Financial statements statistics of corporations by industry," "Interest rate."

Chart 3-13: Estimates: borrowing interest rates (1)

		Dependent variables: borrowing interest rates								
		Construc-tion	Manufacturing: Chemical and materials	Manufacturing: Machinery and equipment	Manufacturing: Other	Infra-structure	Wholesale	Retail	Real estate	Services
Explanatory variables	1-year JGB yield	0.51 ***	0.64 ***	0.64 ***	0.61 ***	0.36 **	0.75 ***	0.47 ***	0.40 ***	0.39 ***
Explanatory variables	1-year JGB yield lagged by 1 year	0.33 ***	0.44 ***	0.40 ***	0.43 ***	0.42 **	0.39 ***	0.36 ***	0.32 ***	0.14 *
Explanatory variables	Term spread (between 1-year and 3-year JGBs)	1.08 ***	1.43 ***	1.67 ***	1.36 ***	2.66 ***	1.18 ***	0.79 ***	--	0.88 **
Explanatory variables	Constant	1.73 ***	1.23 ***	1.25 ***	1.35 ***	1.93 ***	1.27 ***	1.41 ***	1.69 ***	1.08 ***
Sample size		35	35	35	35	35	35	35	35	35
Adj.R^2		0.98	0.98	0.98	0.99	0.91	0.99	0.99	0.98	0.93

Note: 1. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The "--" in the chart indicates that the explanatory variable is excluded from the estimation because the estimate is not statistically significant at the 10 percent level.
2. Term spread is the difference between 1-year and 3-year JGB yields.
Source: Ministry of Finance, "Financial statements statistics of corporations by industry," "Interest rate."

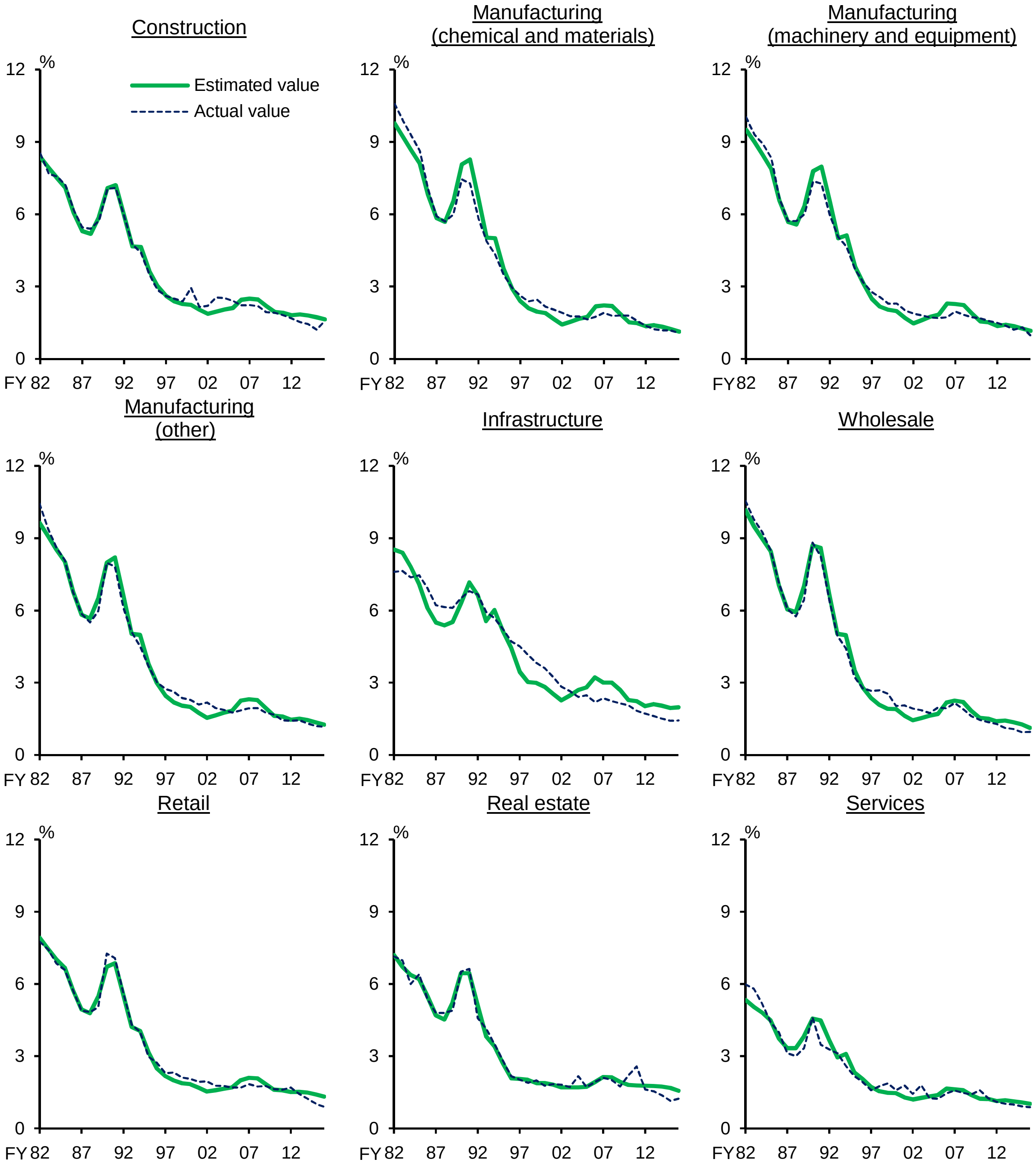
Specifically, the sensitivity of borrowing interest rates to changes in market rates is estimated for each industry by using a simple specification in which the borrowing interest rates (calculated as

[17] For example, due to the zero lower bound on interest rates, the sensitivity of borrowing interest rates to changes in market interest rates in recent years may have been different from that in periods with higher interest rates.

[18] In the stress scenario assuming a rise in interest rates presented below, these estimates for the industry to which a firm belongs are used to calculate the change in the borrowing interest rate of the firm.

described above) and market interest rates are the dependent variable and explanatory variables, respectively. As market interest rates, the term spread between 3-year and 1-year JGB yields is also used in addition to 1-year JGB yields and its lagged values, taking the average remaining maturity of fixed rate loans into account. The estimation results presented in Chart 3-13 indicate that almost all of the estimates have the right sign and are statistically significant, although the sensitivity to market interest rates differs somewhat across industries, reflecting differences in the share of fixed rate loans and in the length of loan periods. Moreover, for all industries, the model shows a good empirical fit, including during phases of rising interest rates (Chart 3-14).

Chart 3-14: Estimates: borrowing interest rates (2)



Note: Latest data as at fiscal 2016.
Source: Ministry of Finance, "Financial statements statistics of corporations by industry."

4. Stress testing

This section presents macro stress tests employing the model for forecasting PD estimated in the preceding section. In this stress testing, the increase in PD (the deviation from the baseline scenario) is measured in response to the following two stress scenarios: (1) a deterioration in the economy and (2) a rise in interest rates.[19]

A. Assumptions underlying the baseline and stress scenarios

The baseline scenario is used as a benchmark for assessing the simulation results of the two stress scenarios. For simplicity, the macroeconomic variables are assumed to remain unchanged over the simulation period (for the 3 years from the start of the simulation).

Next, for the stress scenario of a deterioration in the economy, the output gap is assumed to deteriorate to the same extent as during the Global Financial Crisis. Specifically, the output gap deteriorates by 7 percentage points from the value at the start of the simulation to the next fiscal year and then returns to the original level linearly over the following 2 years. The stress scenario of a rise in interest rates assumes a parallel upward shift of JGB yields by 100 basis points (bps). Specifically, both 1-year and 3-year JGB yields are assumed to increase by 100 bps from the values at the start of the simulation to the next fiscal year and then to remain at that level over the following 2 years.

B. Stress scenario: a deterioration in the economy

Chart 4-1 shows the simulation results -- i.e., increases in PD -- for the scenario assuming an economic deterioration. Based on the portfolio for estimation at the end of the observation period (called the FY2016 portfolio hereafter), the average PD for all industries would increase by 0.32 percentage point. This increase is somewhat smaller than the actual increase in the default rate in the wake of the Global Financial Crisis, which was 0.39 percentage point.[20] On the other hand, if the FY2007 portfolio is used for the simulation, the results indicate that the average PD for all industries would increase by 0.40 percentage point. This increase is larger than the increase based on the FY2016 portfolio and close to the actual increase following the Global Financial Crisis.

Chart 4-1: Simulation results for the stress scenario of a deterioration in the economy

	Simulation results		Actual default rate
	FY2007 portfolio	FY2016 portfolio	
PD (%, average over all industries)	Baseline scenario: 1.46	Baseline scenario: 1.21	Average from FY2003 to FY2007: 1.70
	Stress scenario: 1.86	Stress scenario: 1.53	Average from FY2008 to FY2010: 2.08
Deviation from baseline, % pts	+0.40	+0.32	+0.39

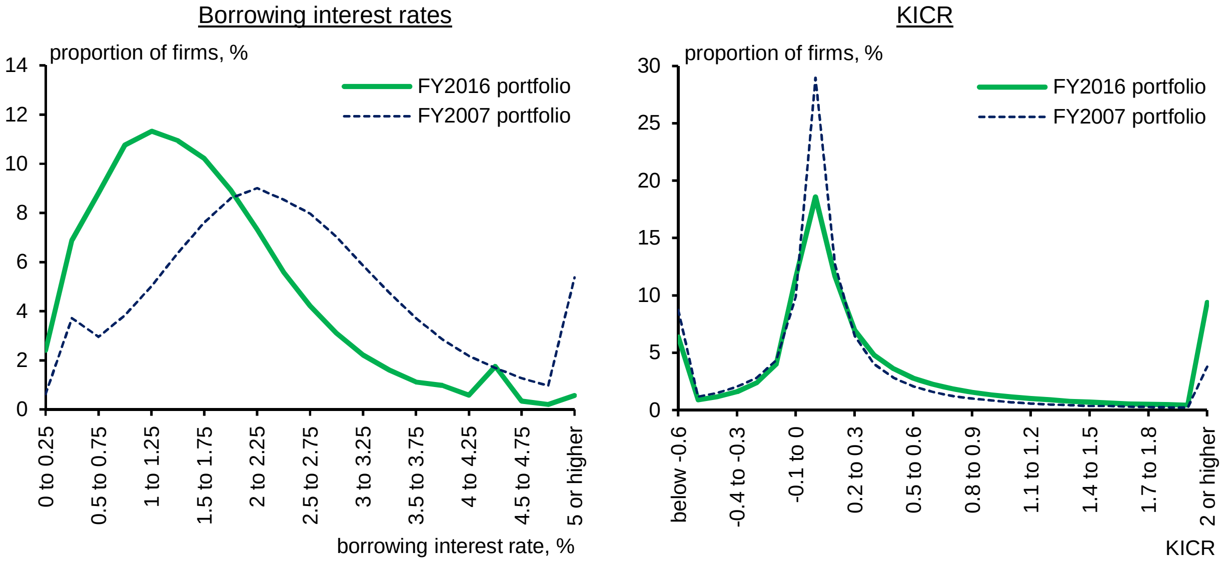
A major reason why the increase in PD based on the FY2016 portfolio is smaller than that based on the FY2007 portfolio is that firms' KICR levels (initial values in the simulations) improved from fiscal 2007 to 2016. This improvement results from firms' borrowing interest rates having declined

[19] While increases in PD of individual firms are calculated in this stress testing, this section presents only the average increase for all industries or within a particular industry.

[20] Specifically, the actual default rate increased from an average of 1.70 percent for the period from fiscal 2003 to 2007 to an average of 2.08 percent for the period from fiscal 2008 to 2010.

across the board, reflecting a decline in market interest rates. This is shown in Chart 4-2 as a substantial shift of the distribution of firms' borrowing interest rates to the left from fiscal 2007 to 2016. Along with the decline in borrowing interest rates, the right tail of the distribution of the KICR has become thicker (i.e., the proportion of firms with a significantly positive KICR has increased). As discussed earlier, the sensitivity of PD to the KICR is non-linear and falls substantially once the KICR becomes positive (Chart 3-6). Therefore, the improvement in firms' KICR levels from fiscal 2007 to 2016 reduces the increase in PD in the event of stress.

Chart 4-2: Distributions of borrowing interest rates and KICR



Source: CRD Association.

Chart 4-3: Simulation results for the stress scenario of a deterioration in the economy (by industry)

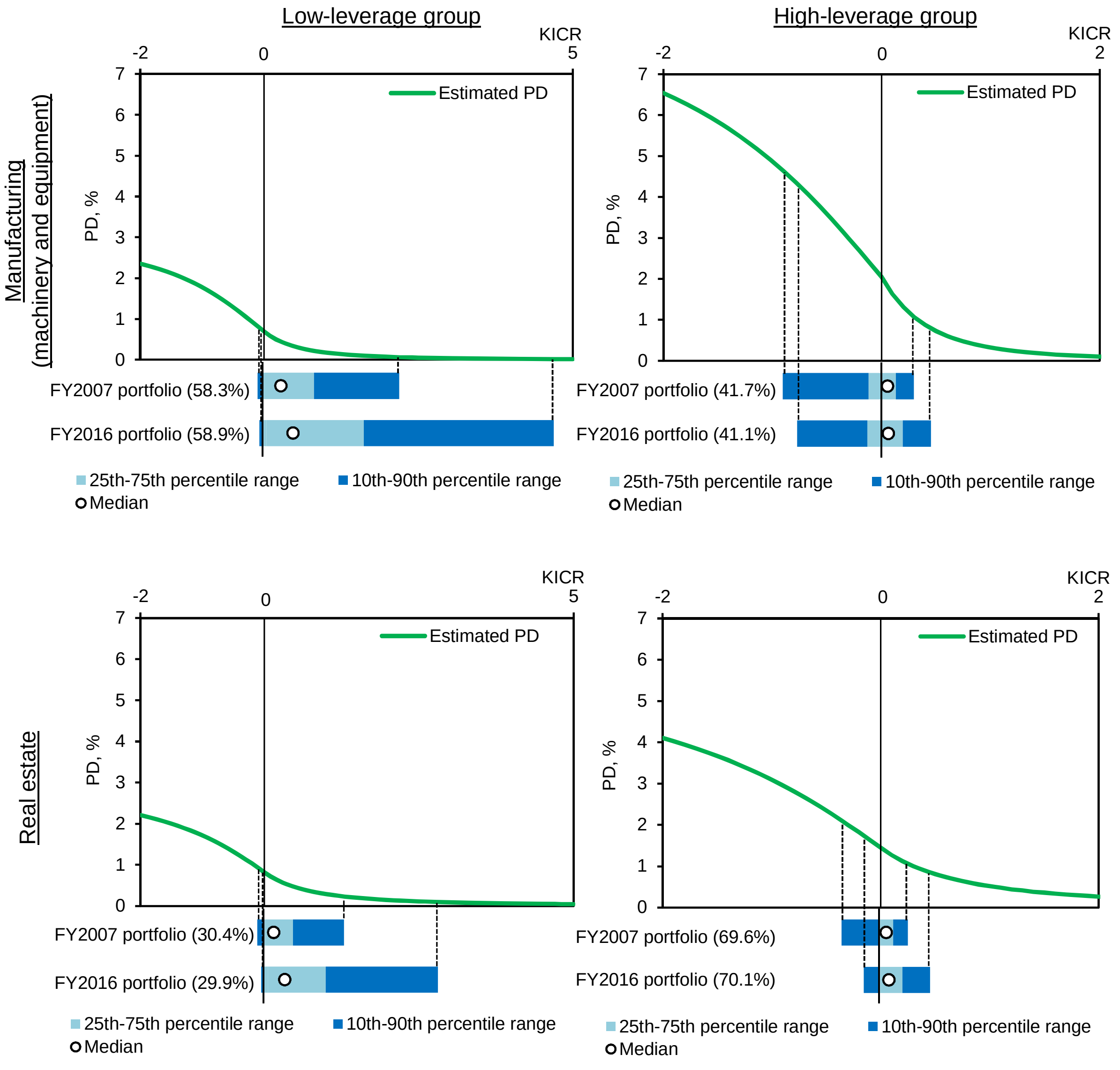
		Deviations of PD from baseline (% pts)		
		FY2007 portfolio	FY2016 portfolio	(Difference from FY2007 portfolio)
Construction		+0.44	+0.35	-0.09
Manufacturing	Chemical and materials	+0.60	+0.53	-0.07
	Machinery and equipment	+1.21	+0.96	-0.25
	Other	+0.75	+0.62	-0.13
Infrastructure		+0.40	+0.32	-0.08
Wholesale		+0.30	+0.29	-0.01
Retail		+0.11	+0.10	-0.01
Real estate		+0.10	+0.11	+0.02
Services		+0.21	+0.19	-0.02

The results by industry presented in Chart 4-3 indicate that the increase in PD is relatively large in industries with a high sensitivity to economic conditions such as manufacturing (machinery and equipment), regardless of the portfolio reference year.[21] There are a few reasons of this result. One is that the sensitivity of the firms' ROA to economic conditions is high in these industries.

[21] For simplicity, the scenario of a deterioration in the economy assumes that only the output gap deteriorates. If commercial property prices are also assumed to fall substantially, the increase in PD of the real estate industry would be larger than otherwise.

Another is that the share of firms with a negative KICR is relatively large in the high-leverage groups of these industries, the groups for which the sensitivity of PD to the KICR is high. This implies that a deterioration in the KICR is prone to giving rise to a non-linear increase in PD (Chart 4-4). On the other hand, in the real estate industry, the increase in PD is small because (1) the sensitivity of firms' ROA to economic conditions is small,[22] and (2) the share of firms with a negative KICR is relatively small. Meanwhile, in almost all industries, the increase in PD on the basis of the FY2016 portfolio is smaller than that on the basis of the FY2007 portfolio.

Chart 4-4: PD and changes in distribution of KICR



Note: 1. The horizontal bar charts show distributions of KICR by portfolio.
2. The proportions of each leverage group by industry are given in parentheses.
Source: CRD Association.

In the model presented in this *Annex*, it is possible to examine reasons for the differences in the increase in PD based on the FY2007 and FY2016 portfolios, focusing on changes in the quality of these portfolios. To this end, it is useful to measure the increase in PD based on the following four portfolios:

[22] The sensitivity of the ROA to economic conditions (the output gap) is estimated to be relatively small for the real estate industry. As pointed out in the previous footnote, it should be noted that the model and the scenario do not incorporate the possibility that commercial property prices fall as a result of a deterioration in the economy.

(a) the FY2007 portfolio (comprising about 0.5 million firms);

(b) a portfolio consisting of firms present in both the FY2007 and FY2016 portfolios (about 0.2 million firms) with their financial indicators fixed at the fiscal 2007 values;

(c) a portfolio consisting of firms present in both the FY2007 and FY2016 portfolios (about 0.2 million firms) with their financial indicators fixed at the fiscal 2016 values;

(d) the FY2016 portfolio (comprising about 0.4 million firms).

By definition, the difference in PD between portfolios (a) and (b) corresponds to the contribution of firms dropping out of the portfolio during the period between fiscal 2007 and 2016 (called the portfolio exit effect here). Next, the difference between portfolios (b) and (c) corresponds to the contribution of changes in firms' financial soundness between fiscal 2007 and 2016 (called the portfolio internal effect). Furthermore, the difference between portfolios (c) and (d) corresponds to the contribution of firms that were not included in the FY2007 portfolio and that were newly added to the portfolio in the intervening period up to fiscal 2016 (called the portfolio entry effect).

Chart 4-5 presents the relative contribution of each of the above effects. According to this chart, the three effects all make similar contributions in terms of direction and impact -- namely, they all slightly reduce the increase in PD. That is, an improvement in the stress resilience of the portfolio is attributable to each of the three corresponding factors: (1) the exit of firms with relatively low creditworthiness, (2) improvements in the creditworthiness of firms that remained in the portfolio, and (3) the entry of firms with relatively high creditworthiness.

Chart 4-5: Decomposition of simulation results for the stress scenario of a deterioration in the economy

Portfolio group	(a)	(b)	(c)	(d)
Sample	All firms as at fiscal 2007 (about 0.5 mil. firms)	Firms remained for both FY2007 and FY2016 portfolios (about 0.2 mil. firms)	Firms remained for both FY2007 and FY2016 portfolios (about 0.2 mil. firms)	All firms as at fiscal 2016 (about 0.4 mil. firms)
Financial condition	FY2007	FY2007	FY2016	FY2016
Deviation of PD from baseline, % pts	+0.40	+0.37	+0.35	+0.32
		Portfolio exit effect: (b)-(a)	Portfolio internal effect: (c)-(b)	Portfolio entry effect: (d)-(c)
		-0.04	-0.02	-0.03

It is worthwhile noting that this kind of decomposition is not possible without a forecast model of PD based on granular data on individual borrowers. For individual financial institutions, applying this approach to their own loan portfolios would be expected to have a wide range of practical applications, such as allowing them to analyze changes in the quality of their loan portfolio and the associated changes in their own stress resilience. As an example, some financial institutions have actively taken on more credit risk by expanding their business with middle-risk firms. For such financial institutions, the default risk in the event of an economic deterioration may have increased in recent years through the portfolio entry effect contributing positively to the increase in PD. Conversely, other financial institutions may have reduced their business with less creditworthy firms during the past economic downturn. For such financial institutions, an increase in PD could be small even in the event of a future deterioration in the economy because of the negative contribution of the portfolio exit effect.

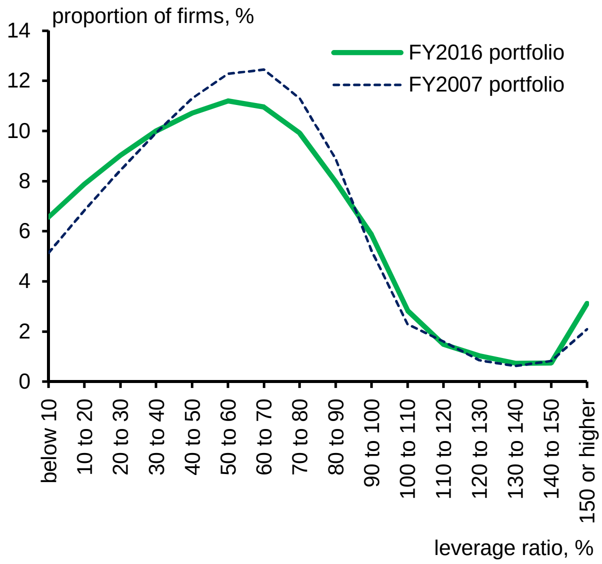
C. Stress scenario: a rise in interest rates

Chart 4-6 shows the simulation results -- i.e., increases in PD -- for the scenario assuming a rise in interest rates. Based on the FY2016 portfolio, PD would increase by 0.18 percentage point, while based on the FY2007 portfolio, it would increase by 0.13 percentage point; that is, the increase is slightly larger when based on the more recent portfolio. This is largely due to the fact that the share of highly leveraged firms increased from fiscal 2007 to 2016 (Chart 4-7). The simulation results suggest that the more recent FY 2016 portfolio is somewhat more vulnerable to an interest rate shock across all industries. This finding is particularly pronounced in some industries (manufacturing, wholesale, retail, and real estate) in which the share of highly leveraged firms became relatively large in the low interest rate environment after the Global Financial Crisis (Chart 4-8).

Chart 4-6: Simulation results for the stress scenario of a rise in interest rates

	Simulation results	
	FY2007 portfolio	FY2016 portfolio
PD (%, average over all industries)	Baseline scenario: 1.46	Baseline scenario: 1.21
	Stress scenario: 1.59	Stress scenario: 1.39
Deviation from baseline, % pts	+0.13	+0.18

Chart 4-7: Distribution of leverage ratios



Source: CRD Association.

Chart 4-8: Simulation results for the stress scenario of a rise in interest rates (by industry)

		Deviations of PD from baseline (% pts)		
		FY2007 portfolio	FY2016 portfolio	(Difference from FY2007 portfolio)
Construction		+0.11	+0.14	+0.03
Manufacturing	Chemical and materials	+0.19	+0.27	+0.09
	Machinery and equipment	+0.15	+0.22	+0.07
	Other	+0.18	+0.26	+0.08
Infrastructure		+0.10	+0.13	+0.03
Wholesale		+0.17	+0.26	+0.09
Retail		+0.16	+0.23	+0.07
Real estate		+0.09	+0.15	+0.06
Services		+0.09	+0.13	+0.04

Because the model presented in this *Annex* produces a non-linear relationship between the KICR and PD, as mentioned earlier, the impact of a rise in interest rates on PD depends on the level of borrowing interest rates at the start of the simulation. As the current level of interest rates is extremely low, the increase in PD in the event of a rise in interest rates would also be small. However, it should be noted that once interest rates have risen to a certain extent, PD could increase in a non-linear fashion in the event of a further rise in interest rates.

5. Conclusion

This *Annex* has presented a forecast model for PD using granular firm-level financial data in order to quantitatively assess credit risk, taking changes in the quality of loans into account. The model has been estimated with a large-scale database (obtained from the CRD) containing financial data of client firms of various financial institutions. The estimation demonstrated that the model performs well empirically and can be used for stress testing. While the stress testing in this *Annex* has assumed two types of macroeconomic shock -- an economic deterioration and an interest rate hike -- cost shocks such as changes in commodity prices or in exchange rates could also be assumed. Moreover, because the model is estimated at the industry level, it could also be used relatively easily for stress testing assuming industry-specific events and scenarios; for instance, the model could be used to examine the impact of a large fall in real estate prices on PD in the real estate industry. Furthermore, it could also be used to assess the relative risk profile and vulnerability of individual financial institutions by comparing their loan portfolios with the portfolio for estimation analyzed in this *Annex* as a benchmark.

Although the analysis presented in this *Annex* has focused on domestic SMEs for which granular data were available, the approach employed in the model could be used for a wide range of databases. For example, it could in principle be useful for assessing the credit risk of overseas firms to which Japan's financial institutions, in particular major banks, have increased their lending in recent years. Many Japanese financial institutions have started to compile databases on financial information of individual overseas firms, but only just recently. It is thus important for these institutions to examine the efficacy of the model using granular data on overseas firms and conduct stress testing with such data in order to strengthen the ability to address a wide range of complex risks related to overseas lending. Needless to say, when applying the model presented in this *Annex* to the credit portfolio of an individual financial institution, it is necessary to appropriately adjust the specification of the model, including the explanatory variables, depending on the purpose of the analysis and the risk profile of its loans.[23]

While this *Annex* has demonstrated how the model for forecasting PD can be used in practice by implementing stress testing exercises, the model is available for other purposes. Apart from stress testing, the model would be particularly useful also for a forward-looking calculation of loan-loss provisions for individual borrowing firms. Internationally, forward-looking loan-loss provisions based on expected credit losses (ECL) are being introduced among major countries; for example, IFRS 9, as an International Financial Reporting Standard, came into force in January 2018, and in the United States, the Current Expected Credit Losses (CECL) accounting standard is scheduled to be applied in 2020. Meanwhile, in Japan as well, with the recall of the financial inspection manual by Japan's Financial Services Agency being scheduled, it is an important task to establish a framework that makes it possible to make loan-loss provisions based on a reasonable economic outlook. Under these circumstances, it is becoming ever more useful to have a model that makes it possible to predict loan losses on the basis of individual firms and/or firms with particular characteristics, using forecasts of future financial and economic conditions.

In this context, the model presented in this *Annex* can be used for calculating PD of individual firms if the future expected path for macroeconomic variables and firms' key financial indicators

[23] For example, it is possible that the specification of the model that would empirically perform well for overseas firms differs from that for domestic firms. In the model here, the output gap and import prices are adopted as important macroeconomic variables that have a major impact on domestic SMEs' debt repayment capacity; however, it is possible that, for example, corporate bond spreads and exchange rates (and/or capital flows) might be more important variables for U.S. firms and firms in emerging markets, respectively.

(the leverage ratio and the liquid asset ratio) are fed into the model. Therefore, setting a variety of likely scenarios instead of tail event scenarios in stress testing and further assuming the path of the firms' leverage ratio based on a financial institution's lending policy -- for example, assuming that loans to a specific industry will rise by x percent annually -- it is also possible to calculate the needed loan-loss provisions for individual firms (or firms with specific characteristics) taking the loss given default (LGD) as given.[24] The information obtained from such exercises would also be useful for deciding the financial institution's lending strategy, in terms of which industries or firms to increase lending to or in terms of how large of risk limits to set for each industry and each firm.

Taking account of the possibilities of further development of models based on granular data, the Bank of Japan's Financial System and Bank Examination Department will continue to improve stress testing and to accumulate knowledge on ECL-based loan-loss provisions, through a close exchange of views and information with financial institutions. Moreover, the Bank will further deepen its dialogue with those concerned on how to efficiently gather and effectively use granular data.

[24] The expected loss (EL) that loan-loss provisions should cover is calculated as the product of PD, LGD, and the exposure at default (EAD). Therefore, in principle, it would be desirable to also construct a forecast model for LGD taking into account the extent to which the value of collateral varies in response to changes in economic conditions and asset prices. However, since data on debt recovery from borrowers are difficult to obtain, empirical research on LGD is much scarcer than that on PD.

Box 1: Constructing the kinked interest coverage ratio (KICR)

The KICR devised in this *Annex* is a transformed version of the regular ICR such that even when the operating ROA in the numerator of the ICR takes a negative value, an increase in the borrowing interest rate or in the leverage ratio in the denominator of the ICR induces a deterioration in the ICR, in other words, an increase in the negative value of the ICR. Specifically, as shown in Chart 3-5,

$$\text{KICR} \approx \begin{cases} \dfrac{R}{i \cdot Lev} \equiv R \cdot \tan\theta & \text{for } \; R \geq 0 \\ R \cdot \tan\left(\dfrac{\pi}{2} - \theta\right) = \dfrac{R}{\tan\theta} = R \cdot i \cdot Lev & \text{for } \; R < 0 \end{cases}$$

where R, i, and Lev represent the operating ROA, the borrowing interest rate, and the leverage ratio, respectively. In addition, θ is the smallest positive real number satisfying $\tan\theta = 1/(i \cdot Lev) \Leftrightarrow \theta = \arctan(1/(i \cdot Lev))$. Plotting the KICR on the vertical axis and R on the horizontal axis, the approximation implies that when $R \geq 0$, the slope is $1/(i \cdot Lev)$, as in the case of the regular ICR. On the other hand, when $R < 0$, the slope is transformed to $i \cdot Lev$, implying that in the region of $R < 0$, the slope is kinked such that the sum of the angles of incidence and refraction is equal to 90 degrees.[25]

In practice, to make it more tractable for estimation, the KICR is defined as a variable that satisfies the following hyperbolic relationship with respect to R for both $R \geq 0$ and $R < 0$ in order to smooth out the kink at $R = 0$. An overline in the following equation indicates that the variable is normalized by its standard deviation:

$$\left(\text{KICR} - \overline{\frac{R}{i \cdot Lev}}\right)\left(\text{KICR} - \overline{R \cdot i \cdot Lev}\right) = k$$

$$\Rightarrow \; \text{KICR} = \frac{1}{2}\left[\left(\overline{R \cdot i \cdot Lev} + \overline{\frac{R}{i \cdot Lev}}\right) + \sqrt{\left(\overline{R \cdot i \cdot Lev} - \overline{\frac{R}{i \cdot Lev}}\right)^2 + 4k}\,\right]$$

where k is a positive constant close to zero.[26]

[25] However, $i \cdot Lev$ often takes a small value of less than 1, so that the slopes shown in Chart 3-5 when $R \geq 0$ ($1/(i \cdot Lev)$) and $R < 0$ ($i \cdot Lev$) are close to respectively vertical (infinity) and horizontal (zero), thus crossing almost at a right angle. In other words, a change in the KICR in response to a change in R differs extremely between $R \geq 0$ and $R < 0$. As a result, it becomes difficult to estimate the relationship between the KICR and PD. In order to address this issue, the KICR is adjusted in advance through normalization (indexation) using its standard deviation so that the change in the KICR with respect to a change in R becomes less dramatic.

[26] As long as $(0 <)\, i \cdot Lev < 1$ holds, taking the limit $k \to 0$ yields the approximate expression of the KICR.

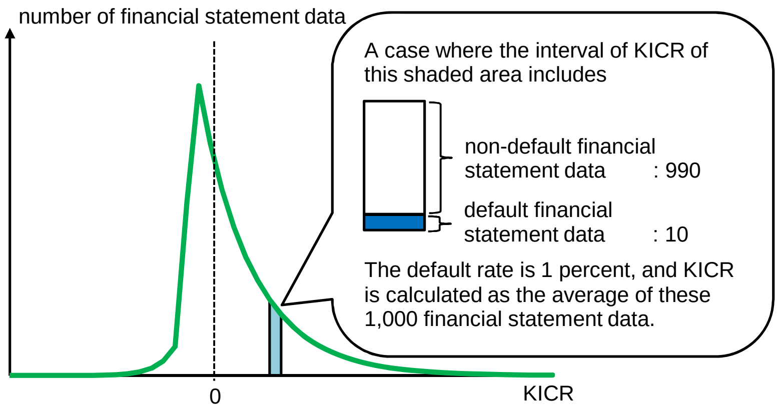
Box 2: Estimating the forecast model for the probability of default (PD)

In practice, when considering the relationship between the KICR and the default rate, data are divided into KICR intervals of a certain width and these variables are calculated for each interval as follows (Chart B2-1):[27]

KICR = Average value of the KICR of all observations within the interval

Default rate = Number of defaults within the interval
/ Number of all observations within that interval

Chart B2-1: Calculation method of default rates



As mentioned in Section 3, while there is a clear negative relationship between the KICR thus constructed and the default rate -- i.e., a decline in the KICR is associated with an increase in the default rate -- this relationship is often non-linear (Chart 3-6). Therefore, when estimating the PD function using the KICR as an explanatory variable, the estimation equation is made as linear as possible by transforming the two variables, in order to make estimation easier. Specifically, referring to previous research, the following quasi-logit transformation $\mathrm{L(PD)}$ is applied to PD and the negative log-transformation $f(\mathrm{KICR})$ is applied to the KICR:[28]

$$\mathrm{L(PD)} = \log\left(\frac{\mathrm{PD}}{\bar{p} - \mathrm{PD}}\right), \qquad f(\mathrm{KICR}) = \begin{cases} \log(1 + \mathrm{KICR}) & \text{for } \mathrm{KICR} \geq 0 \\ -\log(1 - \mathrm{KICR}) & \text{for } \mathrm{KICR} < 0 \end{cases}$$

where $\bar{p}$ represents the maximum actual default rate.

Then, assuming that the relationship between $\mathrm{L(PD)}$ and $f(\mathrm{KICR})$ takes the following functional form, this relationship is estimated using the method of non-linear least squares:[29]

$$L(\mathrm{PD}) = \beta + \frac{1}{2}\left[(\gamma + \delta) \cdot f(\mathrm{KICR}) - \sqrt{(\gamma - \delta)^2 \cdot \left(f(\mathrm{KICR})\right)^2 + 4h}\,\right] + \rho \cdot \log(\mathrm{LIQ})$$

[27] The width of the intervals is adjusted to ensure that there are a sufficient number of observations in each interval.

[28] One such prior example of using this log transformation to maintain a monotonically increasing relation is found in Satoshi Yamashita and Kakeru Miura, "Shinyou risuku moderu no yosoku seido: AR-chi to hyouka shihyou," *Finance Library, vol. 11*, Asakura Publishing Co., Ltd., 2011 (available in Japanese only).

[29] The method of non-linear least squares makes it possible to simultaneously estimate parameters by minimizing the sum of squared residuals between estimated and actual values even when the relationship between the parameters to be estimated is non-linear.

where LIQ is the liquid asset ratio, which is added to the explanatory variables in order to consider firms' liquidity as well as their solvency; h is a positive constant close to zero; and β, γ, δ, and ρ are the parameters to be estimated. This functional form is a hyperbola whose asymptotes are two straight lines that intersect at $f(\mathrm{KICR}) = 0$ when the final term representing firms' liquidity, $\rho \cdot \log(\mathrm{LIQ})$, is omitted. Furthermore, γ and δ correspond to the slopes of the asymptotes when the KICR is in positive and in negative territory, respectively. However, when there is no significant difference between the slopes of the two asymptotes -- that is, when $\gamma = \delta$ cannot be rejected in the non-linear least squares estimation -- the estimation is conducted assuming that the relationship between $\mathrm{L(PD)}$ and $f(\mathrm{KICR})$ is linear, as in the following specification:

$$L(\mathrm{PD}) = \beta + \alpha \cdot f(\mathrm{KICR}) + \rho \cdot \log(\mathrm{LIQ})$$

where α, β, and ρ are the parameters to be estimated.

The estimation results are presented in Chart B2-2. These results indicate that, with the exception of the liquid asset ratio for the high-leverage group in the real estate industry, all explanatory variables are statistically significant and the parameter estimates have the expected sign (Chart B2-2).[30] In other words, the results suggest that PD increases as firms' debt repayment capacity deteriorates or as their liquidity relative to short-term debt becomes poorer due to a decline in the KICR or a decrease in the liquid asset ratio. While the coefficients of determination differ across industries and between the high-leverage and low-leverage groups, they all take values between approximately 0.7 and 0.9, indicating that the fit of the estimated values to the actual values is generally reasonable.

[30] Because of the non-linearity of the estimation equation, the marginal sensitivity (first-order derivative) of PD to the explanatory variables depends non-linearly not only on the values of the parameters to be estimated and the value of $\bar{p}$ but also on the values of the explanatory variables themselves. Consequently, it is not possible to judge the sensitivity of PD only from the magnitude of the absolute values of the parameter estimates. Note also, when both $f(\mathrm{KICR})$ and $\log(\mathrm{LIQ})$ are zero, PD is approximately $\bar{p}/(1 + e^{-\beta})$.

Chart B2-2: Estimates: PD

		Dependent variables: quasi-logit transformation of PD								
		Construc-tion	Manufacturing			Infra-structure	Wholesale	Retail	Real estate	Services
			Chemical and materials	Machinery and equipment	Other					
		Low-leverage group								
Explanatory variables	KICR α	-1.90 ***	-2.56 ***	-2.27 ***	-2.50 ***	-1.60 ***	-2.80 ***		-1.82 ***	-1.65 ***
	Constant β	-0.92 ***	-1.42 ***	-0.86 ***	-0.58 ***	-1.17 ***	-0.90 ***	-0.89 ***	-0.94 ***	-0.87 ***
	KICR (+) γ							-1.90 ***		
	KICR (-) δ							-1.87 ***		
	Liquid asset ratio ρ	-0.66 ***	-0.51 ***	-0.55 ***	-0.85 ***	-0.70 ***	-0.50 ***	-0.50 ***	-0.11 **	-0.54 ***
Sample size		111	59	71	66	101	76	86	83	110
Adj.R^2		0.85	0.83	0.72	0.80	0.81	0.84	0.73	0.68	0.81
		High-leverage group								
Explanatory variables	KICR α	-1.70 ***				-1.42 ***			-1.75 ***	-1.67 ***
	Constant β	-0.83 ***	-0.61 ***	-1.08 ***	-0.92 ***	-1.03 ***	-1.27 ***	-1.06 ***	-1.15 ***	-1.06 ***
	KICR (+) γ		-3.69 ***	-2.98 ***	-3.79 ***		-1.93 ***	-2.57 ***		
	KICR (-) δ		-2.02 ***	-2.00 ***	-2.50 ***		-1.87 ***	-1.73 ***		
	Liquid asset ratio ρ	-0.28 ***	-0.26 **	-0.33 ***	-0.35 ***	-0.25 ***	-0.20 ***	-0.16 ***	--	-0.27 ***
Sample size		130	73	95	99	129	109	130	42	157
Adj.R^2		0.84	0.77	0.89	0.87	0.80	0.86	0.80	0.88	0.87

Note: 1. ***, **, and * indicate statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively. The "--" in the chart indicates that the explanatory variable is excluded from the estimation because the estimate is not statistically significant at the 10 percent level.
2. "KICR" is the negative log-transformed value and "Liquid asset ratio" is the logarithm value.