# IMES DISCUSSION PAPER SERIES

**Determinants of Liquidity in the Japanese Government Bond Market:**
**An Interpretable Machine Learning Approach**

Satoko Kojima and Toshiyuki Sakiyama

**Discussion Paper No. 2026-E-3**

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

2-1-1 NIHONBASHI-HONGOKUCHO
CHUO-KU, TOKYO 103-8660
JAPAN

**IMES Discussion Paper Series 2026-E-3**
**March 2026**

# Determinants of Liquidity in the Japanese Government Bond Market: An Interpretable Machine Learning Approach

**Satoko Kojima* and Toshiyuki Sakiyama****

**Abstract**

Liquidity in government bond markets is critical for the functioning of financial markets. This paper studies the determinants of market liquidity, measured by price dispersion, by constructing various bond features using high-granularity data from the Bank of Japan Financial Network System and applying machine learning approaches. The main findings are threefold. First, the decomposition of the liquidity indicator into bond features reveals that the historical volatility of benchmark prices of Japanese government bonds has been the main driver of the liquidity indicator, while the contributions of the share of non-clearing participants' transactions and the share of the central bank's transactions and holdings have increased since around 2022. Second, some bond features affect the liquidity indicator non-linearly. For bond features such as the share of foreign financial institutions' transactions, the number of trading financial institutions, and the share of the central bank's holdings, the liquidity indicator improves as the values of these bond features increase, but deteriorates once they exceed certain thresholds. Third, bond features such as maturity, the historical volatility of benchmark prices, and the number of trading counterparties per institution affect the liquidity indicator by strongly interacting with other bond features.

**Keywords:** Market liquidity; Government bond markets; Bond features; Machine learning approach

**JEL classification:** C59, G12

*Director, Institute for Monetary and Economic Studies, Bank of Japan (Email: satoko.kojima@boj.or.jp)

**Director and Senior Economist, Institute for Monetary and Economic Studies, Bank of Japan (Email: toshiyuki.sakiyama@boj.or.jp)

The authors would like to thank participants of 2025 BOJ-IMES Finance Workshop, CFE-CMStatistics 2025, and QMF 2025, and the staff of the Bank of Japan for their useful comments. All the remaining errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan.

## 1. Introduction

The liquidity of government bond markets has attracted considerable attention amid persistently high volatility in government bond yields worldwide. In 2020, government bond yields experienced notable fluctuations globally, driven largely by the outbreak of the COVID-19 pandemic. In the fall of 2022, UK government bond yields surged sharply amid mounting concerns over fiscal conditions. Subsequently, there were several instances of large yield movements, particularly in the US and Europe, as policy interest rates were raised sharply. In contrast, Japan has experienced relatively moderate interest rate volatility, as financial conditions have been accommodative under a prolonged low interest rate environment. However, the Japanese government bond (JGB) market has yet to be tested in an environment where the policy interest rate is significantly above zero following this period of low rates.

At this juncture, understanding potential determinants of liquidity in the JGB market is of critical importance for both market participants and policymakers. However, while the number of studies of the determinants is increasing, studies that examine the determinants using transaction-level granular data are limited. To fill this gap, this paper uses granular data from the Bank of Japan Financial Network System (BOJ-NET) to construct a liquidity indicator measured by price dispersion and a wide range of bond features as potential determinants of liquidity for each individual JGB. It then employs machine learning approaches such as Random Forest and SHAP, which can capture complex relationships between a dependent variable and independent variables, to reveal which bond features affect the liquidity indicator and how.

Our main findings are threefold. First, the decomposition of the liquidity indicator into bond features reveals that the historical volatility of benchmark prices has been the main driver of the liquidity indicator, while the contributions of the share of non-clearing participants' transactions and the share of the central bank's transactions and holdings have increased since around 2022. Second, some bond features affect the liquidity indicator non-linearly. For bond features such as the share of foreign financial institutions' transactions, the number of trading financial institutions, and the share of the central bank's holdings, the liquidity indicator improves as the values of these bond features increase, but deteriorates once they exceed certain thresholds. Third, bond features such as maturity, the historical volatility of benchmark prices, and the number of trading counterparties per institution affect the liquidity indicator by strongly interacting with other bond features.

The primary advantage of utilizing the BOJ-NET data is that it encompasses granular information on the name of the financial institution, the name of the counterparty financial institution, the transaction volume, and the transaction price for each individual settlement transaction, as well as information on the holding amount of each financial institution for each individual JGB. This allows us to construct a diverse array of bond features as well as a liquidity indicator for each individual JGB.

We then construct various bond features for individual JGBs based on prior research. For example, O'Hara (1995) points out the importance of transaction volume, transaction size, dealers' inventory holdings, and the proportion of informed market participants in considering market liquidity. Gravelle (1999) analyzes the relationship between market liquidity and the transaction volume of non-residents in the Canadian government bond market. Galliani, Resti and, Petrella (2014) examine the relationships between market liquidity and factors such as newly issued bonds, outstanding amount, and remaining maturity in the European government bond markets. Fukuma et al. (2024) explore the relationships between market liquidity and the central bank purchase amounts or holdings in the JGB market.

Based on these prior studies, this study constructs the following bond features for each individual JGB: (i) maturity, (ii) age, (iii) the transaction volume to outstanding amount ratio, (iv) the historical volatility of benchmark prices, (v) the share of foreign financial institutions' transactions, (vi) the share of non-clearing participants' transactions, (vii) the share of large-lot transactions, (viii) the share of the central bank's transactions, (ix) the number of trading financial institutions, (x) the number of trading counterparties per institution, (xi) the share of financial institutions' own-account holdings, and (xii) the share of the central bank's holdings.

Regarding the liquidity indicator, there is no definitive measure of government bond market liquidity, reflecting the multifaceted nature of liquidity. In this study, following Jankowitsch, Nashikkar, and Subrahmanyam (2011), we construct a liquidity indicator for individual JGBs, called price dispersion. Although price dispersion is just one measure of liquidity, it is expected to reflect market conditions in more detail compared to other liquidity indicators, such as those derived from the best quotes, as it is calculated from actually executed transactions.

Which bond features are related to price dispersion and how they are related is unknown. In addition, the relationships could be highly non-linear. To address these issues, we employ machine learning approaches, which allow us to systematically assess potential relationships between a diverse array of bond features and price dispersion, without imposing assumptions. The approaches thus enable us to investigate potential complex relationships between bond features and the liquidity indicator.

Among various machine learning models, this study employs Random Forest as the selected machine learning model. Random Forest is widely recognized as one of the most successful and most versatile models for capturing non-linear relationships in data, as noted by Varin, Reid, and Firth (2011) and Mullainathan and Spiess (2017). Their strengths notwithstanding, machine learning models, including Random Forest, are frequently criticized for their black box nature and lack of interpretability. To address this limitation, the field of eXplainable AI (XAI) has made substantial progress in recent years, yielding the development of tools and techniques to enhance the interpretability of machine learning models. This study utilizes widely adopted visualization tools, including Partial Dependence Plot (PDP), SHapley Additive exPlanations (SHAP), and Feature Interaction (FI), to interpret the results generated by Random Forest.

PDP visualizations reveal the presence of notable non-linear relationships between certain bond features and price dispersion. We show that bond features such as the share of foreign financial institutions' transactions, the number of trading financial institutions, and the share of the central bank's holdings tend to improve price dispersion as their values rise. Beyond specific thresholds, however, these relationships reverse, leading to worsening price dispersion. In contrast, other bond features, including maturity, age, and the transaction volume to outstanding amount ratio, exhibit monotonic relationships, with price dispersion consistently worsening as the values of these bond features increase.

SHAP visualizations reveal that various bond features significantly influence price dispersion. Key features include the historical volatility of benchmark prices, maturity, and age. Further, the decomposition of price dispersion into bond features identifies the historical volatility of benchmark prices as the most influential feature. Since around 2022, however, the contributions of the share of non-clearing participants' transactions and the share of the central bank's transactions and holdings have increased significantly.

Finally, FI visualizations reveal interactions among bond features and show that maturity, the historical volatility of benchmark prices, and the number of trading counterparties per institution in particular affect price dispersion by strongly interacting with other bond features, underscoring the importance of accounting for such interactions when conducting detailed analysis. For instance, when analyzing the decomposition of price dispersion into bond features using SHAP or the relationships between bond features and price dispersion using PDP, the interactions made by maturity, the historical volatility of benchmark prices, and the number of trading counterparties per institution should be carefully considered, such as by analyzing both the SHAP decomposition and the PDP relationships by maturity.

These results indicate that various bond features are associated with the liquidity of the JGB market, highlighting that the relationships are often complex and non-monotonic. Therefore, to better understand the relationships between market liquidity and bond features, it is useful to adopt machine learning approaches capable of capturing these complex relationships among numerous variables.

**Related Literature**

This study is related to prior research in three respects. First, it is related to studies that analyze liquidity in government bond markets using high-granularity data. These include Galliani, Resti, and Petrella (2014) for the European government and corporate bond markets; Pinter (2023) for the UK government bond market; Fleming (2023) for the US government bond market; and Jankowitsch, Nashikkar, and Subrahmanyam (2011) for the US corporate bond market. These studies analyze the relationships between liquidity indicators and key bond features for individual bonds, such as maturity, outstanding amount, and transaction volume. In addition, the role of the central bank's purchases of government bonds in determining liquidity on an individual bond basis is studied by Kandre and Schleshu (2013) for the US government bond market; Schlepper et al. (2020) for the European government bond markets; and Finlay, Titkov, and Xiang (2023) for the Australian government bond market. Regarding the JGB market, studies that focus on individual bonds include Kakuma (2012) who examines the relationship between maturity and a liquidity indicator, while Uno and Tobe (2021), Han and Seneviratne (2018), and Fukuma et al. (2024) investigate relationships between the central bank's purchases of JGBs and liquidity indicators. In this study, we focus on individual bonds in the JGB market, construct a greater number of bond features than prior research, and examine their relationship with a liquidity indicator using machine learning approaches.

Second, this paper contributes to the literature on liquidity indicators in government bond markets, especially in the JGB market. BIS (1999) analyzes liquidity in government bond markets in advanced economies, using indicators such as bid-ask spreads and turnover ratios. Nishizaki, Tsuchikawa, and Yagi (2013) evaluate market liquidity from four dimensions: volume, tightness, resiliency, and depth. Building on this, Kurosaki et al. (2015) utilize high-granularity data from the JGB futures market to construct indicators for each of the four dimensions. Sakiyama and Kobayashi (2018) construct indicators pertaining to the four dimensions for the JGB cash market using high-granularity data on inter-dealer transactions. Regarding studies that use high-granularity data, Kakuma (2012) constructs an indicator called the best-worst quote spreads of dealer-to-client transactions, but the data coverage is limited to only some dealer-to-client transactions.

Sakiyama and Yamada (2016) utilize high-granularity data from the BOJ-NET which record almost all transactions, to construct transaction networks and develop network-based indicators such as the number of trading financial institutions and the number of links. Our focus in this study is primarily on transactions involving customers, and we construct a liquidity indicator called price dispersion using high-granularity data from the BOJ-NET.

Third, this paper is related to the literature on the application of machine learning techniques in analyzing financial markets. Technological innovation and advances in digitalization in recent years have accelerated the use of alternative data, resulting in the increased application of machine learning techniques in finance (see Warin and Stojkov [2021]; Nazareth and Ramana Reddy [2023]). Varin, Reid, and Firth (2011) and Mullainathan and Spiess (2017) propose the use of machine learning models, such as Least Absolute Shrinkage and Selection Operator, Support Vector Machines, Random Forest, Neural Networks, and Deep Learning, to account for the complex relationships that traditional linear models fail to capture amid the growing utilization of big data. However, Varin, Reid, and Firth (2011) highlight a drawback of machine learning models, noting that they are black box models, which makes interpreting their results challenging. In this context, increasing attention has been given in recent years to tools and techniques for interpreting results, collectively known as eXplainable AI (XAI). Visualization tools that have been introduced include Permutation Feature Importance (PFI), Partial Dependence Plot (PDP), Individual Conditional Explanation (ICE), SHapley Additive exPlanations (SHAP), and Feature Interaction (FI) (see Arrieta et al. [2020]; Yeo et al. [2025]; Inglis, Parnell, and Hurley [2024]). In the field of finance, Buckmann and Joseph (2023) and Bluwstein et al. (2023) utilize machine learning models for their analysis and employ XAI to interpret the obtained results. Furthermore, Luque Raya and Luque Raya (2024) analyze relationships between government bond yields and market liquidity using machine learning models. In this study, we analyze the relationships between bond features and a liquidity indicator using Random Forest as a machine learning model, and then present the interpretation of the results obtained, using visualization tools such as

PDP, FI, and SHAP.

The rest of the paper is organized as follows. Section 2 explains a wide range of bond features and a liquidity indicator in the JGB market using data from the BOJ-NET, and provides an overview of these indicators. Section 3 describes our machine learning-based analytical approach, while Section 4 applies this approach in examining the relationships between the bond features and liquidity indicator. Section 5 concludes the paper.

## 2. Data and Descriptive Analysis

This section explains the data used in this study and introduces a wide range of bond features and a liquidity indicator. Section 2.1 provides an overview of the data. Sections 2.2 and 2.3 introduce and explain the bond features and liquidity indicator, respectively. Section 2.4 provides a descriptive summary of the bond features and liquidity indicator.

### 2.1 Overview of the Data

Our data were collected through the BOJ-NET Funds Transfer System and BOJ-NET JGB services, which the BOJ provides with the aim of settling funds and JGBs between the BOJ and financial institutions efficiently and safely online. The data cover all dematerialized JGBs (book-entry JGBs) and encompass detailed information on financial institutions' JGB settlement transactions and their JGB holdings. The data contain detailed information for each individual settlement transaction, including the name of the financial institution, the name of the counterparty financial institution, the transaction volume, and the transaction price, as well as information on the amount held by each financial institution for each individual JGB. The sample period covers early October 2015 to the end of September 2024.

Approximately 32,000 government bond settlement transactions are processed via the BOJ-NET per day, representing approximately 131 trillion yen.[1] About 80% of the

[1] The BOJ publishes some settlement data, such as can be found in the following link: https://www.boj.or.jp/en/statistics/set/kess/index.htm

transactions are in the form of delivery-versus-payment (DVP) settlements.[2] In this study, we use DVP settlements for our analysis while excluding settlement transactions conducted through the Japan Securities Clearing Corporation (JSCC)—the central counterparty for JGBs—and repurchase agreement (repo) transactions.[3] Regarding settlement transactions executed through the JSCC, transaction prices are determined at the same price[4] for each individual JGB, which makes them difficult for use in analysis from the perspective of price dispersion as a liquidity indicator. In addition, outright transactions and repo transactions often exhibit price differences depending on the presence of collateral; repo transactions are thus excluded from the scope of this study.

### 2.2 Bond Features

Leveraging our machine learning approach, we consider as many bond features that have been explored in previous research as possible. As mentioned, we construct the following twelve indicators for each individual JGB as bond features: (i) maturity, (ii) age, (iii) the transaction volume to outstanding amount ratio, (iv) the historical volatility of benchmark prices, (v) the share of foreign financial institutions' transactions, (vi) the share of non-clearing participants' transactions, (vii) the share of large-lot transactions, (viii) the share of the central bank's transactions, (ix) the number of trading financial institutions, (x) the number of trading counterparties per institution, (xi) the share of financial institutions' own-account holdings, and (xii) the share of the central bank's holdings. In the following, we discuss these bond features in turn.

Regarding (i) maturity, Kakuma (2012) analyzes the JGB market and reports that bonds with shorter maturities tend to exhibit higher liquidity. Similarly, Jankowitsch, Nashikkar, and Subrahmanyam (2011) analyze the US corporate bond market and also point out that

[2] To mitigate settlement risk, a DVP mechanism has been implemented for JGB transactions, ensuring that the delivery of a JGB occurs concurrently with the corresponding transfer of funds. The settlement data for JGB transactions encompass not only outright transactions (i.e., the buying and selling of bonds) but also repo and collateralized call transactions.

[3] We detect repo transactions in the data by identifying reverse transactions involving the same JGB and the same nominal amount within 100 days of the initial transaction.

[4] This transaction price is referred to as the JGB base price and is published by the Japan Exchange Group.

bonds with shorter maturities tend to have higher liquidity.

Regarding (ii) age, it is defined as the time that has passed since issuance divided by the original maturity, so that age is normalized between 0 and 1. Galliani, Resti, and Petrella (2014) analyze the European bond markets and report that newly issued bonds tend to exhibit higher liquidity. Similarly, Uno and Tobe (2021) examine the JGB market and also find a tendency for higher liquidity in newly issued bonds. Jankowitsch, Nashikkar, and Subrahmanyam (2011) also report that US corporate bonds with a shorter time since issuance tend to show higher liquidity.

Regarding (iii) the transaction volume to outstanding amount ratio, it is calculated on a monthly basis for each individual JGB by dividing the daily transaction volume by its outstanding amount. Jankowitsch, Nashikkar, and Subrahmanyam (2011), who focus on the US corporate bond market, report that bonds with larger transaction volumes tend to exhibit higher liquidity. On the other hand, Fleming (2003) reports a negative correlation coefficient between transaction volume and liquidity indicators in the US government bond market. In this regard, the theoretical literature has highlighted mechanisms that explain both positive and negative relationships between transaction volume and market liquidity. The positive correlation arises from a mechanism in which increased trading activity by uninformed market participants reduces the risk of losses for dealers during trading, thereby supporting the dealers' market-making function and improving liquidity (see O'Hara [1995]; Kyle [1985]; Glosten and Milgrom [1985]). On the other hand, the negative correlation arises from new market information that creates variations in market participants' expectations of future prices and boosts transaction volumes due to the differing expectations among participants (see Clark [1973]; Tauchen and Pitts [1983]).

Regarding (iv) the historical volatility of benchmark prices, it is calculated on a monthly basis for each individual JGB as the standard deviation of benchmark prices, where benchmark prices refer to the JGB base prices published daily by the Japan Exchange Group. It is widely recognized that volatility is closely tied to liquidity indicators. Duffie et al. (2023) report a strong positive correlation between the principal components of various liquidity indicators and volatility in the US government bond market. Duffie et al. (2023) and O'Hara (1995) further highlight that higher volatility heightens the risks associated with inventory holding, compelling dealers to incorporate these risks into transaction prices, which results in a decline in liquidity.

Regarding (v) the share of foreign financial institutions' transactions, it represents the proportion of transactions conducted by foreign financial institutions, excluding all transactions by the central bank. Markets with participants who have diverse trading strategies and different perspectives on interest rates generally tend to exhibit higher liquidity, as transactions are less likely to be skewed in one direction. BIS (1999) highlights a mechanism in which one-sided trading raises the costs associated with market making, prompting dealers to transfer these costs to transaction prices, which ultimately leads to a deterioration in liquidity. Supporting this, Gravelle (1999) finds that non-resident participation boosts transaction volumes in the Canadian government bond market.

Regarding (vi) the share of non-clearing financial institutions' transactions, it represents the proportion of transactions conducted by non-clearing participants, excluding all transactions by the central bank. Non-clearing participants are generally considered to be less informed compared to clearing participants. Theoretically, there are mechanisms explaining both positive and negative relationships between the activity of uninformed market participants and market liquidity. On one hand, a positive relationship can be observed when increased trading activity by uninformed market participants reduces the risk of trading losses for dealers (see O'Hara [1995]; Kyle [1985]; Glosten and Milgrom [1985]). On the other hand, a negative relationship can be observed when uninformed market participants, acting as noise traders disconnected from market fundamentals, lead rational market participants to hesitate in executing arbitrage transactions, thereby giving rise to persistent price deviations (see De Long et al. [1990]).

Regarding (vii) the share of large-lot transactions, it represents the proportion of transactions with a size of 5 billion yen or more, excluding all transactions by the central bank. O'Hara (1995) highlights that an increase in large-sized transactions raises the costs of market making, prompting dealers to transfer these costs to transaction prices, thereby causing a deterioration in liquidity. In contrast, Pinter (2023), analyzing the UK government bond market at the time of the sharp rise in interest rates in the fall of 2022, reports that transaction costs notably increased for small-sized transactions and those involving smaller dealers. Kakuma (2012) finds no significant differences in liquidity based on transaction size in the JGB market.

Regarding (viii) the share of the central bank's transactions and (xii) holdings, the empirical literature provides inconclusive results on the relationship between central bank government bond purchases and market liquidity. Focusing on the JGB market, Uno and Tobe (2021) find that bonds to be purchased by the central bank generally tend to exhibit higher liquidity. Fukuma et al. (2024) identify a non-linear relationship between the central bank's holding ratio and liquidity, reporting that bonds with substantially high central bank holdings tend to exhibit lower liquidity. Focusing on the US government bond market, Kandrac and Schlusche (2013) find no evidence that central bank purchases of government bonds negatively impacted liquidity. In contrast, Schlepper et al. (2020) argue that purchases of government bonds by the European Central Bank adversely affected liquidity. Similarly, Finlay et al. (2023), focusing on the Australian government bond market under the central bank's yield target policy, report that central bank purchases of government bonds resulted in a deterioration in liquidity.

Regarding (ix) the number of trading financial institutions and (x) the number of trading counterparties per institution, the former refers to the number of financial institutions that engage in transactions in a given month, while the latter represents the average number of counterparties with whom each financial institution engages in transactions in a given month. Theoretically, there are several mechanisms that explain the positive relationship between the number of financial institutions participating in transactions and market liquidity. Duffie, Gârleanu, and Pedersen (2005) and Vayanos and Wang (2007) study a mechanism in which an increase in the number of participating financial institutions reduces search costs incurred by dealers when finding appropriate counterparties or the best transaction prices, thereby improving liquidity. Coen and Coen (2022) highlight that lower search costs help alleviate dealers' concerns about inventory holding, enhancing their market-making function and further boosting liquidity. On the other hand, when the number of uninformed market participants increases, a different mechanism may arise in which rational market participants refrain from engaging in arbitrage transactions, allowing price deviations to persist.

Regarding (xi) the share of financial institutions' own-account holdings, it is calculated as the ratio of the institutions' holdings for their own accounts to the total outstanding amount for each individual JGB, excluding the central bank's holdings. Focusing on the US government bond market, Fleming (2003) and Fleming, Nguyen, and Rosenberg (2024) report that dealer inventory holdings support their market-making function, thus improving liquidity. Conversely, Duffie et al. (2023) find that excessive dealer inventory holdings can negatively affect market liquidity, as dealers pass on the risks associated with inventory holding to transaction prices.

### 2.3 Price Dispersion as a Liquidity Indicator

Various liquidity indicators have been proposed both in the literature and in practice.[5] In this study, following Jankowitsch, Nashikkar, and Subrahmanyam (2011), we adopt price dispersion as a liquidity indicator for the JGB market. The liquidity indicator is constructed for each individual JGB by calculating the daily differences of transaction prices from the benchmark price published by the Japan Exchange Group. Our liquidity indicator is then calculated as the one-month standard deviation of these differences in transaction prices.

A government bond with low price dispersion is considered to have high liquidity, in the sense that it can be traded at the intended price without price fluctuations, even in the course of substantial transactions. One advantage of price dispersion as a liquidity indicator is that it is constructed from actually executed trades, so that it is considered to

[5] The BOJ publishes some indicators relevant to the JGB market, such as can be found in the following link: https://www.boj.or.jp/en/paym/bond/index.htm#p02

reflect market conditions more accurately than bid-ask spreads derived from the best quotes, which are often used as a liquidity indicator in the literature. [6] These considerations informed the use of price dispersion as the primary indicator of market liquidity in this study.[7]

## 2.4 Descriptive Statistics

Before studying the relationships between the liquidity indicator and the bond features introduced above, their basic statistics are provided. Table 1 shows the mean, standard deviation, and percentiles for all monthly indicators from October 2015 to September 2024. Here, several observations are in order.

Regarding the liquidity indicator, its mean is 34.2 JPY cents, with the 10th percentile at 1.0 JPY cents and the 90th percentile at 81.3 JPY cents.[8]

Regarding bond features, (i) maturity and (ii) age are in the ranges of 1.1–25.7 years with a median of 8.6 years, and 0.10–0.85 with a median of 0.46, respectively. The median maturity is relatively long and the median age is around half of the maturity.

Regarding (iii) the transaction volume to outstanding amount ratio, it has a mean value of 0.30%, with the 10th percentile at 0.02% and the 90th percentile at 0.61%. On average, the daily transaction volume across all JGBs accounts for approximately 0.3% of the outstanding amount. Considering that there are around 250 trading days in a year, this suggests that the annual transaction volume is approximately 75% of the outstanding amount. Regarding (iv) the historical volatility of benchmark prices, the values range from 1.0 JPY cent at the 10th percentile to 78.2 JPY cents at the 90th percentile, with an average of 31.4 JPY cents. This feature clearly shows large variations relative to its mean.

---

[6] For example, in the context of the JGB market, Uno and Tobe (2021) and Fukuma et al. (2024) analyze bid-ask spreads among dealers.

[7] Price dispersion has been used as an indicator for market liquidity in various studies, including Friewald, Jankowitsch, and Subrahmanyam (2012) and Schestag, Schuster, and Uhrig-Homburg (2016).

[8] Jankowitsch, Nashikkar, and Subrahmanyam (2011) measure price dispersion in the US corporate bond market, and point out that the average from October 2004 to October 2006 is approximately 50 bps.

Regarding (v) the share of foreign financial institutions' transactions, it has a mean value of 29.8%, with a range between 1.8% at the 10th percentile and 68.4% at the 90th percentile. On average, across all JGBs, this suggests that foreign financial institutions participating as buyers or sellers account for approximately 30% of transaction volumes, excluding transactions conducted by the central bank. Additionally, for certain JGBs, the share of transactions involving foreign institutions can reach up to 70%. Regarding (vi) the share of non-clearing participants' transactions, it has a mean value of 13.4%, with a range between 0.1% at the 10th percentile and 38.9% at the 90th percentile. On average, across all JGBs, non-clearing participants involved as buyers or sellers account for approximately 15% of transaction volumes, excluding transactions conducted by the central bank. Furthermore, transactions for certain JGBs involve no non-clearing participants, as only clearing participants are engaged in the transactions.

Regarding (vii) the share of large-lot transactions, it shows a median value of 0%, with a range from 0% at the 10th percentile to 28.3% at the 90th percentile. This suggests that more than half of JGBs have no transactions exceeding 5 billion yen per trade when excluding transactions conducted by the central bank. Regarding (viii) the share of the central bank's transactions, it has a median value of 0%, ranging from 0% at the 10th percentile to 33.4% at the 90th percentile. This suggests that the central bank conducts no transactions for more than half of JGBs on a monthly basis. For certain JGBs, however, the central bank accounts for approximately a third of the total monthly transaction volume.

Continuing with bond features, (ix) the number of trading financial institutions and (x) the number of trading counterparties per institution range from 11 to 28, with a median of 20, and from 1.6 to 4.2, with a median of 2.6, respectively. This suggests that, on average, across all JGBs, approximately 20 financial institutions participate in transactions at least once per month, with each financial institution trading with 2 to 3 counterparties.

Regarding (xi) the share of financial institutions' own-account holdings and (xii) the share of the central bank's holdings, the ranges are 14.5–45.2%, with a mean of 27.6%, and 12.2–76.6%, with a mean of 41.0%, respectively. This indicates that, on average, across all JGBs, the central bank holds approximately 40% of the outstanding amount, while financial institutions, excluding the central bank's holdings, hold approximately 30% of the outstanding amount for their own accounts.

## 3. Empirical Approach

With the bond features and liquidity indicator laid out, in this section, we present our empirical approach to uncovering potential relationships between them. Specifically, we use a machine learning model called Random Forest and then apply model explainability techniques to the results, such as PDP, FI, and SHAP, to visualize which the bond features affect the liquidity indicator and how.

### 3.1 Machine Learning Model

The primary benefit of employing machine learning models is their ability to systematically assess relationships between a dependent variable and independent variables, without ruling out the possibility that these relationships are non-linear or the possibility that independent variables are interrelated. Another benefit is their capacity to allow for the concurrent use of numerous independent variables.

We employ Random Forest, a machine learning algorithm composed of multiple decision trees.[9] In our case, each decision tree generates a predicted value of the liquidity indicator for a given data point of the bond features. The average of predicted values across all decision trees is our prediction of the liquidity indicator for a given data point of the bond features. In constructing each decision tree, data are randomly sampled from our dataset, and the sampled data are split into two groups iteratively to make the predicted value of the liquidity indicator. Random Forest mitigates the overfitting problem, a common issue with decision trees, by constructing multiple decision trees randomly.

---

[9] See, for example, Breiman (2001) for the details of Random Forest.

Random Forest is among the most frequently used machine learning models. For instance, in a large-scale empirical study comparing 179 algorithms, Fernández-Delgado et al. (2014) identify Random Forest as the best family of classifiers. One reason for its popularity is that, unlike many other machine learning models, Random Forest achieves high accuracy without requiring the tuning of numerous complex hyperparameters (see Hastie, Tibshirani, and Friedman [2009]). Random Forest also offers a computational advantage: The training process can be parallelized since each decision tree is constructed independently, allowing for efficient learning even with large-scale datasets (see Liaw and Wiener [2002]). Furthermore, a distinct feature of Random Forest is its ability to evaluate model accuracy by utilizing data not selected during sampling, without preparing a separate validation dataset (see Breiman [2001]).

### 3.2 Model Explainability Techniques

Machine learning models often exhibit complexity and produce outputs that are difficult to interpret. To address this issue, we employ model explainability techniques such as PDP, FI, and SHAP, which are explained below.

#### 3.2.1 PDP

PDP is a method, introduced by Friedman (2001), that visualizes the relationship between the $j$-th feature $X_j$ and the predicted value of $Y$. In our case, for example, PDP generates a liquidity indicator as a function of a bond feature, such as maturity. Let $Y = ML(X)$ denote a machine learning model that predicts $Y$ by using an input $X = \{X_j\}_{j=1}^{J}$. In our case, we have $J = 12$ bond features. For notational simplicity, let $X_{-j}$ define a set of input variables except for $X_j$ such that $X = \{X_j, X_{-j}\}$. The $PD$ function, $PD(X_j)$, can be mathematically expressed as follows:

$$PD\left(X_j\right) = \frac{1}{M}\sum_{m=1}^{M} ML(X_j, X_{-j}^{m})$$

where $m = 1, \ldots, M$ denotes the $m$-th data point of $X = \{X_j\}_{j=1}^{J}$. The $PD\left(X_j\right)$ function is given by the average of a variable predicted from the machine learning model

$ML$ over all other features except for $X_j$. Using this $PD$ function, we can obtain a simplified representation of how a specific bond feature influences the liquidity indicator. However, it should be noted that $PD(X_j)$ captures the average effect of $X_j$ on the predicted variable. If the feature $X_j$ is independent of the other features, $PD(X_j)$ captures the exact effect of $X_j$, but otherwise, the effect of $X_j$ depends on the values of other features.

### 3.2.2 FI

FI in machine learning models refers to the effect of the combination of two or more features on the predicted variable. In regression analysis, introducing the product of two explanatory variables is a commonly used method of capturing interactive effects. Following the methodology of Friedman and Popescu (2008), we measure and visualize the strength of feature interactions among bond features.

The strength of feature interactions of $X_j$, denoted by $H_j^2$, can be calculated as follows. First, given a single data point of $X = \{X_j\}_{j=1}^{J}$, the feature interaction function $FI_j(X)$ for $X_j$ is calculated as follows:

$$FI_j(X) = ML(X) - \frac{1}{M}\sum_{m=1}^{M} ML\left(X_j, X_{-j}^{m}\right) - \frac{1}{M}\sum_{m=1}^{M} ML\left(X_j^{m}, X_{-j}\right).$$

If $X_j$ is independent of the other features, the right-hand side of the above equation is zero so that $FI_j(X) = 0$. If $X_j$ shows some degree of interaction with the other features, $FI_j(X)$ becomes non-zero. Based on this property, $H_j^2$ is derived as follows:

$$H_j^2 = \sum_{m=1}^{M}\left[FI_j(X^m)\right]^2 / \sum_{m=1}^{M}[ML(X^m)]^2 .$$

Thus, $H_j^2$ captures how large the deviation is on average from the independence assumption for $X_j$ in predicting a dependent variable. It is worth noting that $H_j^2$ pertains to the size of interactions, so it does not provide information about the direction of the effect of $X_j$ on the predicted variable.[10]

### 3.2.3 SHAP

Following Buckmann and Joseph (2023), who propose utilizing SHAP as an XAI technique to address the issue of interpretability, we use SHAP to quantify and visualize the contribution of each bond feature to the liquidity indicator. In SHAP, the predicted value $\hat{Y}^m$ for a data point $(Y^m, X^m)$ is decomposed into the average of all predicted values, $\bar{Y} = (1/M)\sum_{m=1}^{M} ML(X^m)$, and the contributions of each feature, $\varphi_j^m$, as follows:

$$\hat{Y}^m = ML(X^m) = \bar{Y} + \sum_{j=1}^{J} \varphi_j^m .$$

The contribution $\varphi_j^m$ is referred to as the SHAP value of $X_j$ for the data point $(Y^m, X^m)$. For instance, if the SHAP value for maturity of the 10-year JGB #360 in March 2024 is 10 JPY cents, and the SHAP value for the share of the central bank's transactions is 5 JPY cents, this indicates that maturity plays a more significant role in explaining the predicted liquidity indicator. In our analysis, we conduct similar calculations for all JGBs and across all time points.

The SHAP value $\varphi_j^m$ is calculated as the change in the predicted value resulting from the inclusion of $X_j$. Let $S$ denote a subset of feature indexes. In our case, $S \subseteq \{1,2,\dots,J\}$. Define $v_m(S) = ML(\{X_j^m\})$ with $j \in S$ as the predicted value based on the features in *S*. Features not included in *S* are averaged over all data points. To determine $\varphi_j^m$, a change in the predicted value when $X_j$ is added as follows:

$$v_m(S \cup j) - v_m(S).$$

For example, if $J = 4, j = 3, S = \{1,2\}$, the value above can be calculated as follows:

[10] To analyze feature interactions in greater detail, other methods are often used, such as 2D PDP and SHAP interaction values. See, for example, Molnar (2025) for details.

$$v_m(\{1,2,3\}) - v_m(\{1,2\}) = \frac{1}{M}\sum_{i=1}^{M} ML\left(X_1^m, X_2^m, X_3^m, X_4^i\right) - \frac{1}{M}\sum_{i=1}^{M} ML\left(X_1^m, X_2^m, X_3^i, X_4^i\right).$$

By performing this calculation for all possible combinations of $S$, $\varphi_j^m$ can ultimately be expressed as follows:

$$\varphi_j^m = \frac{1}{J!}\sum_{S\subseteq\{1,\ldots,J\}\backslash\{j\}} [|S|!\,(J-|S|-1)!]\,[v_m(S\cup j) - v_m(S)]$$

where $|S|$ represents the number of elements in $S$.

## 4. Empirical Results

This section presents the main results of our empirical analysis. As explained in the previous section, we employed Random Forest to investigate potential relationships between various bond features and the liquidity indicator. The empirical results are presented by using PDP, FI, and SHAP to elucidate the intricate patterns of these relationships and provide interpretable insights. SHAP-based time-series decomposition of the liquidity indicator is also presented.

Before presenting the results, the fitting accuracy of Random Forest is evaluated. The Root Mean Squared Error (RMSE)[11] is 35 JPY cents, the Mean Absolute Error (MAE)[12] is 13 JPY cents, and the R-squared ($R^2$)[13] is about 0.6. For reference, the panel analysis that includes entity fixed effects and time fixed effects yields the RMSE of 48 JPY cents, the MAE of 19 JPY cents, and the $R^2$ of about 0.3. These values indicate that our machine learning model shows a relatively high fitting accuracy.[14]

---

[11] The RMSE is defined as: $\sum_{m=1}^{M}(Y^m - \hat{Y}^m)^2/M$.

[12] The MAE is defined as: $\sum_{m=1}^{M}\left|Y^m - \hat{Y}^m\right|/M$.

[13] The $R^2$ is defined as: $1 - \sum_{m=1}^{M}(Y^m - \hat{Y}^m)^2/\sum_{m=1}^{M}(Y^m - \bar{Y})^2$.

[14] In Random Forest, data not selected during the sampling process using the bootstrap method can be used to evaluate the fitting accuracy of unseen data, without the need for separate test data. This unselected data is known as Out-of-Bag (OOB) data, and the $R^2$ calculated using such OOB data is approximately 0.4.

## 4.1 PDP

We use PDP to quantify how changes in a single bond feature affect the liquidity indicator predicted from our machine learning model. Figure 1 presents the results of our PDP analysis for all bond features. The figure indicates several notable patterns between bond features and the predicted liquidity indicator.

Regarding (i) maturity, the liquidity indicator, measured by price dispersion, is increasing in maturity, as shown in Figure 1(i), implying that bonds with longer maturities have lower liquidity. This observation aligns with the findings of Kakuma (2012) and Jankowitsch, Nashikkar, and Subrahmanyam (2011). In addition, two break points are observed, where the liquidity indicator significantly increases when maturity exceeds 5 years and 15 years.

Regarding (ii) age, the liquidity indicator remains flat when age is below around 0.5, implying that liquidity is unchanged for bonds whose time since issuance is less than half of their original maturities (Figure 1(ii)). However, the liquidity indicator increases as age approaches 1, indicating that liquidity declines as bonds approach redemption. Despite differences in terms of non-linearity, this finding is broadly consistent with the results of Galliani, Resti, and Petrella (2014), Uno and Tobe (2021), and Jankowitsch, Nashikkar, and Subrahmanyam (2011), who all report higher liquidity for newly issued bonds.

Regarding (iii) the transaction volume to outstanding amount ratio, the liquidity indicator tends to rise for bonds with larger transaction volumes (Figure 1(iii))—that is, liquidity tends to deteriorate for such bonds. This result is consistent with Fleming (2003), who finds a negative relationship between transaction volume and liquidity in the US government bond market. In contrast, analyzing the US corporate bond market, Jankowitsch, Nashikkar, and Subrahmanyam (2011) report that bonds with higher transaction volumes tend to exhibit better liquidity.

Regarding (iv) the historical volatility of benchmark prices, the liquidity indicator tends to rise for bonds with larger fluctuations in benchmark prices (Figure 1(iv)). This result is consistent with Duffie et al. (2023), who analyze the US government bond market and report a positive relationship between the principal components of multiple liquidity indicators and volatility.

Regarding (v) the share of foreign financial institutions' transactions, the liquidity indicator decreases as the share increases up to approximately 30%. Beyond that point, however, the liquidity indicator begins to rise, suggesting a non-linear U-shape relationship (Figure 1(v)). This pattern appears to align with previous studies that suggest that markets with diverse participants adopting various trading strategies tend to exhibit higher liquidity. For instance, Gravelle (1999) reports that the participation of non-residents contributes to enhancing liquidity in the Canadian government bond market.

Regarding (vi) the share of non-clearing participants' transactions, a positive relationship with the liquidity indicator is observed (Figure 1(vi)). This may be attributed to the fact that such participants are generally considered to possess relatively less information, and an increase in transactions by noise traders may lead other market participants to hesitate to engage in arbitrage transactions, resulting in a persistent decline in liquidity (see De Long et al. [1990]).

Regarding (vii) the share of large-lot transactions, no significant differences in the liquidity indicator are observed (Figure 1(vii)). This finding aligns with Kakuma (2012) with regard to the JGB market. In contrast, analyzing the UK government bond market during the sharp interest rate hikes in the fall of 2022, Pinter (2023) finds that transaction costs rose significantly for small-sized transactions and those involving smaller dealers.

Regarding (viii) the share of the central bank's transactions, no significant differences in the liquidity indicator are observed (Figure 1(viii)). In addition, regarding (xii) the share of the central bank's holdings, a non-linear U-shape relationship is observed. Specifically, an increase in the share up to 25% tends to lower the liquidity indicator. When the share exceeds 50%, however, further increases tend to raise the liquidity indicator (Figure 1(xii)). The empirical literature finds mixed results on the relationship between central bank purchases of government bonds and market liquidity. Kandrac and Schlusche (2013) report no evidence that central bank bond purchases reduced liquidity in the US government bond market. Fukuma et al. (2024), in line with our study, reveal a non-linear relationship between the central bank's holding ratios and market liquidity, reporting that bonds with substantially high central bank holding ratios tend to exhibit lower liquidity.

Regarding (ix) the number of trading financial institutions and (x) the number of trading counterparties per institution, both exhibit a non-linear U-shaped relationship with the liquidity indicator (Figures 1(ix) and (x)). The liquidity indicator initially decreases as these numbers increase but begins to rise once they exceed a certain threshold. These results are consistent with Coen and Coen (2022) who highlight that an increase in the number of financial institutions participating in transactions reduces search costs for dealers, thereby facilitating market making and improving liquidity; however, when the number of uninformed market participants rises significantly, a deterioration in liquidity may persist, potentially offsetting liquidity improvements.

Regarding (xi) the share of financial institutions' own-account holdings, there is a tendency for the liquidity indicator to decrease as the share increases up to 20% (Figure 1(xi)). This result suggests that financial institutions holding an appropriate amount of inventory help to enhance the market-making function, as noted by Fleming (2003) and Fleming, Nguyen, and Rosenberg (2024). Conversely, Duffie et al. (2023) identify a mechanism whereby excessive increases in financial institutions' holdings could reduce their willingness to engage in market-making activities, due to the heightened inventory costs and regulatory burdens, thereby disrupting smooth transactions. It is worth noting that our dataset does not appear to include any instances of excessive increases in financial institutions' holdings of JGBs, so the mechanism proposed by Duffie et al. (2023) would not apply to our results.

### 4.2 FI

Figure 2 shows the strength of feature interactions, measured by $H_j^2$, for each bond feature. Bond features exhibiting relatively strong feature interactions are maturity, the historical volatility of benchmark prices, the number of trading counterparties per institution, and age. This result suggests that the relationships between bond features and the predicted liquidity indicator could vary significantly depending on the values of these bond features with relatively strong interactions.

In addition, to analyze the feature interactions in greater detail, we applied two-dimensional PDP to our machine learning model. Since the share of the central bank's holdings shows a large variation in our dataset, we examine the relationship between the share of such holdings and the predicted liquidity indicator (Figure 3). The analysis focuses on bond features with relatively strong interactions, such as maturity and the number of trading counterparties per institution, as well as bond features with weaker interactions, such as the share of non-clearing participants' transactions.

Regarding the interaction of non-clearing participants' transactions, the PDP for the share of the central bank's holdings appears primarily to undergo a parallel shift in response to variations in the level of the share of such transactions, with its overall shape remaining largely unaffected (Figure 3(iii)). This result implies no significant degree of interactions between the share of non-clearing participants' transactions and the share of the central bank's holdings.

Next, an examination of maturity reveals distinct patterns (Figure 3(i)). For bonds with shorter maturities, the liquidity indicator remains stable even as the share of the central bank's holdings increases significantly. In contrast, for bonds with longer maturities, a reduction in the share of the central bank's holdings results in a pronounced increase in the liquidity indicator, suggesting a deterioration in liquidity. Furthermore, an analysis of the number of trading counterparties per institution demonstrates that bonds with fewer counterparties exhibit a notable rise in the liquidity indicator as the share of the central bank's holdings increases (Figure 3(ii)). These results suggest that maturity and the number of trading counterparties per institution have strong interactions with the share of the central bank's holdings.

### 4.3 SHAP-Based Feature Importance

Figure 4 displays the average absolute SHAP values for all bond features. Among these, the historical volatility of benchmark prices, maturity, and age are identified as the three most important bond features in predicting the liquidity indicator. The SHAP values are approximately 15 JPY cents for the historical volatility of benchmark prices, 12–13 JPY cents for maturity, and 5 JPY cents for age. Interestingly, these three bond features also align with the features exhibiting a high degree of feature interactions. Moreover, while maturity demonstrates the strongest interaction with other bond features, the historical volatility of benchmark prices holds the highest SHAP value. This finding suggests that, although maturity acts as the most significant facilitator of the effects of other bond features on the liquidity indicator, the historical volatility of benchmark prices has the largest direct and indirect impact on the indicator.

The remaining bond features have SHAP values of less than 5 JPY cents. Among these, bond features with relatively high SHAP values, ranging from 2 to 4 JPY cents, include the number of trading counterparties per institution, the transaction volume to outstanding amount ratio, the share of financial institutions' own-account holdings, the share of the central bank's holdings, and the share of non-clearing participants' transactions. Although the individual SHAP values of these five bond features are relatively small, their combined SHAP values are comparable to that of the historical volatility of benchmark prices, the bond feature with the highest SHAP value. Therefore, these five bond features also play an important role in explaining the liquidity indicator.

### 4.4 Time-Series Decomposition Using SHAP

The time-series developments in the predicted liquidity indicator can be decomposed into the contributions from various bond features using SHAP values. Given the substantial number of bond features included in our analysis, they are grouped into five categories for the simplicity of visualization: (i) maturity and age, (ii) the historical volatility of benchmark prices, (iii) the share of non-clearing participants' transactions and related features (the share of non-clearing participants' transactions, the share of foreign financial institutions' transactions, the share of large-lot transactions, and the transaction volume to outstanding amount ratio), (iv) the shares of central bank's transactions and holdings, and (v) other features (the number of trading financial institutions, the number of trading counterparties per institution, and the share of financial institutions' own-account holdings).

Figure 5 shows the time-series decomposition of the predicted liquidity indicator into the five categories of bond features from January 2016 to September 2024, based on all JGBs.[15] Until around the end of 2021, the historical volatility of benchmark prices accounts for the bulk of variation in the liquidity indicator, during which overall fluctuations in the liquidity indicator remain small. From 2022, however, the liquidity indicator increases, and its movement is influenced not only by the historical volatility of benchmark prices but also by various bond features such as maturity and age, the share of non-clearing participants' transactions and related features, the share of the central bank's transactions and holdings, and other features.[16]

Figure 6 presents the same time-series decomposition by maturity range: medium-term (1–5 years), long-term (5–10 years), and ultra-long-term (10–40 years). As might be expected from the previous PDP analysis (Figure 1(i)), there is a relatively large range for the liquidity indicator in the ultra-long-term range. From 2022, the liquidity indicator increases across all maturity ranges, especially for long-term and ultra-long-term JGBs, with the historical volatility of benchmark prices, maturity and age, and the share of non-clearing participants' transactions and related features contributing to the upward movements. In addition, for ultra-long-term JGBs, while the historical volatility of benchmark prices contributes to a decrease in the liquidity indicator from 2017 to 2021, it contributes to an increase in the liquidity indicator from 2022.

## 5. Conclusion

[15] The contribution decomposition, on an all JGBs basis, is constructed by aggregating the decomposition result for individual JGBs, weighted by their respective transaction volumes.

[16] This study assumes that the effects of macroeconomic conditions are captured by the historical volatility of benchmark prices. When additional macro-related features are included, such as JGB futures prices or the call rate, the influence of the historical volatility of benchmark prices declines. However, the main finding of this study—that price dispersion has a complex relationship to a wide range of bond features—remains unchanged.

This study utilized high-granularity data from the BOJ-NET for JGBs since 2016 to construct a diverse array of bond features and a market liquidity indicator, referred to as price dispersion, for each individual JGB. We then investigated potential relationships between these bond features and the liquidity indicator by employing machine learning approaches, including Random Forest and SHAP.

The analysis reveals three key points. First, the decomposition of the liquidity indicator into bond features reveals that the historical volatility of benchmark prices has been the main driver of the liquidity indicator, while the contributions of the share of non-clearing participants' transactions and the share of the central bank's transactions and holdings have increased since around 2022. Second, some bond features affect the liquidity indicator non-linearly. For bond features such as the share of foreign financial institutions' transactions, the number of trading financial institutions, and the share of the central bank's holdings, the liquidity indicator improves as the values of these bond features increase, but deteriorates once they exceed certain thresholds. Third, bond features such as maturity, the historical volatility of benchmark prices, and the number of trading counterparties per institution affect the liquidity indicator by strongly interacting with other bond features.

Finally, we would like to mention a few caveats about our analysis. First, as noted in the main text, we focus on data on outright transactions of JGBs, excluding data on repo transactions. Data on transactions cleared through the JSCC are also excluded. Analyzing these data would be a next step toward a comprehensive understanding of liquidity in the JGB market. Second, our analysis is based on a dataset from October 2015 to September 2024, a period when interest rates in Japan were low. Although most of our results are consistent with previous studies, some of our results might change in a significantly positive interest rate environment. Third, the machine learning approaches used in this study capture correlational relationships but not necessarily causal relationships between the liquidity indicator and bond features. With these caveats in mind, our hope is that our analysis will be useful for understanding and monitoring liquidity in government bond markets.

## References


Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI." *Information Fusion* 58 (2020): 82–115.

Bank for International Settlements (BIS). "Market Liquidity: Research Findings and Selected Policy Implications." *CGFS Papers* No. 11 (1999): https://www.bis.org/publ/cgfs11.htm.

Bluwstein, Kristina, Marcus Buckmann, Andreas Joseph, Sujit Kapadia, and Özgür Şimşek. "Credit Growth, the Yield Curve and Financial Crisis Prediction: Evidence from a Machine Learning Approach." *Journal of International Economics* 145, (2023): 103773.

Breiman, Leo. "Random Forests." *Machine Learning* 45 (2001): 5–32. https://link.springer.com/content/pdf/10.1023/A%3A1010933404324.pdf.

Buckmann, Marcus and Andreas Joseph. "An Interpretable Machine Learning Workflow with an Application to Economic Forecasting." *International Journal of Central Banking* 19, no. 4 (2023): 449–552. https://www.ijcb.org/sites/default/files/journal/v19n4/ijcb-v19n4-interpretable-machine-learning-workflow-application-economic-forecasting.pdf.

Clark, Peter K. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica* 41, no.1 (1973): 135–155.

Coen, Jamie and Patrick Coen. "A Structural Model of Liquidity in Over-the-Counter Markets." Bank of England Working Paper No. 979 (2022).

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98, no. 4 (1990): 703–738.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen. "Over-the-Counter Markets." *Econometrica* 73, no. 6 (2005): 1815–1847.

Duffie, Darrell, Michael J. Fleming, Frank M. Keane, Claire Nelson, Or Shachar, and Peter Van Tassel. "Dealer Capacity and U.S. Treasury Market Functionality." Federal Reserve Bank of New York Staff Reports No. 1070 (2023).

Finlay, Richard, Dmitry Titkov, and Michelle Xiang. "The Yield and Market Function Effects of the Reserve Bank of Australia's Bond Purchases." *Economic Record* 99, no. 326 (2023): 359–384.

Fernández-Delgado, Manuel, Eva Cernadas, Senén Barro, and Dinani Amorim. "Do We Need Hundreds of Classifiers to Solve Real World Classification Problems?" *Journal of Machine Learning Research* 15 (2014): 3133–3181.

Fleming, Michael J. "Measuring Treasury Market Liquidity." *Economic Policy Review* 9, no.3 (2003): 83–108.

Fleming, Michael (Federal Reserve Bank of New York). 2023. "How Has Treasury Market Liquidity Evolved in 2023?" Liberty Street Economics. https://libertystreeteconomics.newyorkfed.org/2023/10/how-has-treasury-market-liquidity-evolved-in-2023/.

Fleming, Michael, Giang Nguyen, and Joshua Rosenberg. "How Do Treasury Dealers Manage Their Positions?" *Journal of Financial Economics* 158 (2024): 103885.

Friedman, Jerome H. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29, no. 5 (2001): 1189–1232.

Friedman, Jerome H. and Bogdan E. Popescu. "Predictive Learning via Rule Ensembles." *The Annals of Applied Statistics* 2, no. 3 (2008): 916–954.

Friewald, Nils, Rainer Jankowitsch, and Marti G. Subrahmanyam. "Illiquidity or Credit Deterioration: A Study of Liquidity in the US Corporate Bond Market during Financial Crises." *Journal of Financial Economics* 105, no. 1 (2012): 18–36.

Fukuma, Noritaka, Tomiyuki Kitamura, Kohei Maehashi, Naoki Matsuda, Keita Takemura, and Kota Watanabe. "The Impact of Quantitative and Qualitative Easing and Yield Curve Control on the Functioning of the Japanese Government Bond Market." Bank of Japan Working Paper Series No.24-E-9 (2024).

Galliani, Clara, Andrea Resti, and Giovanni Petrella. "The Liquidity of Corporate and Government Bonds: Drivers and Sensitivity to Different Market Conditions." JRC Technical Reports (2014).

Glosten, Lawrence R. and Paul R. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics* 14, no. 1 (1985): 71–100.

Gravelle, Toni. "Liquidity of the Government of Canada Securities Market: Stylized Facts and Some Market Microstructure Comparisons to the United States Treasury Market." Bank of Canada Staff Working Papers 99-11 (1999).

Han, Fei and Dulani Seneviratne. "Scarcity Effects of Quantitative Easing on Market Liquidity: Evidence from the Japanese Government Bond Market." IMF Working Paper No. 18/96 (2018).

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edition, New York: Springer, 2009.

Inglis, Alan, Andrew Parnell, and Catherine Hurley. "Visualizations for Bayesian Additive Regression Trees." *Journal of Data Science, Statistics, and Visualisation*, 4, no. 1 (2024).

Jankowitsch, Rainer, Amrut Nashikkar, and Marti G. Subrahmanyam. "Price Dispersion in OTC Markets: A New Measure of Liquidity." *Journal of Banking & Finance* 35, no. 2 (2011): 343–357.

Kakuma, Kazuo. "Liquidity Structure and Execution Cost in the Japanese Government Bond Market —Analysis and Consideration of Trading Inquiry Data on an

Electronic Trading Platform—." *Securities Analysts Journal* (2012). https://www.saa.or.jp/english/professional/pdf/aj1209_kakuma.pdf.

Kandrac, John and Bernd Schlusche. "Flow Effects of Large-Scale Asset Purchases." *Economics Letters* 121, no. 2 (2013): 330–335.

Kurosaki, Tetsuo, Yusuke Kumano, Kota Okabe, and Teppei Nagano. "Liquidity in JGB Markets: An Evaluation from Transaction Data." Bank of Japan Working Paper Series, No.15-E-2 (2015).

Kyle, Albert S. "Continuous Auctions and Insider Trading." *Econometrica* 53, no. 6 (1985): 1315–1335.

Liaw, Andy and Matthew Wiener. "Classification and Regression by randomForest." *R News* 2/3 (2002): 18–22. https://journal.r-project.org/articles/RN-2002-022/

Luque Raya, Ignacio Manuel, and Pablo Luque Raya. "Machine Learning Algorithms Applied to the Estimation of Liquidity: The 10-year United States Treasury Bond." *European Journal of Management and Business Economics* 33, no. 3 (2024): 341–365.

Molnar, Christoph. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 3rd edition (2025). https://christophm.github.io/interpretable-ml-book/

Mullainathan, Sendhil and Jann Spiess. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31, no. 2 (2017): 87–106.

Nazareth, Noella and Yeruva Venkata Ramana Reddy. "Financial Applications of Machine Learning: A Literature Review." *Expert Systems with Applications* 219, (2023): 119640.

Nishizaki, Kenji, Akira Tsuchikawa, and Tomoyuki Yagi. "Indicators Related to Liquidity in JGB Markets." Bank of Japan Review 2013-E-3 (2013).

O'Hara, Maureen. *Market Microstructure Theory*, Cambridge MA: Blackwell (1995).

Pinter, Gabor. "An Anatomy of the 2022 Gilt Market Crisis." Bank of England Working Paper No. 1019 (2023).

Sakiyama, Toshiyuki and Shun Kobayashi. "Liquidity in the JGB Cash Market: An Evaluation from Detailed Transaction Data." BOJ Reports and Research Papers, March 2018.

Sakiyama, Toshiyuki and Tetsuya Yamada. "Market Liquidity and Systemic Risk in Government Bond Markets: A Network Analysis and Agent-Based Model Approach." IMES Discussion Paper Series 2016-E-13 (2016).

Schestag, Raphael, Philipp Schuster, and Marliese Uhrig-Homburg. "Measuring Liquidity in Bond Markets." *The Review of Financial Studies* 29, no. 5 (2016): 1170–1219.

Schlepper, Kathi, Heiko Hofer, Ryan Riordan, and Andreas Schrimpf. "The Market Microstructure of Central Bank Bond Purchases." *Journal of Financial and Quantitative Analysis* 55, no. 1 (2020): 193–221.

Tauchen, George E. and Mark Pitts. "The Price Variability-Volume Relationship on Speculative Markets." *Econometrica* 51, no. 2 (1983): 485–505.

Uno, Jun and Reiko Tobe. "Liquidity of Government Bond Market and Monetary Policy: Recent Research Trends." Financial Review 146 (2021): 74-94 (in Japanese).

Varin, Cristiano, Nancy Reid, and David Firth. "An Overview of Composite Likelihood Methods." *Statistica Sinica* 21, no. 1 (2011): 5–42.

Vayanos, Dimitri and Tan Wang. "Search and Endogenous Concentration of Liquidity in Asset Markets." *Journal of Economic Theory* 136, no. 1 (2007): 66–104.

Warin, Thierry and Aleksandar Stojkov. "Machine Learning in Finance: A Metadata-Based Systematic Review of the Literature." *Journal of Risk and Financial Management* 14, no. 7 (2021): 302.

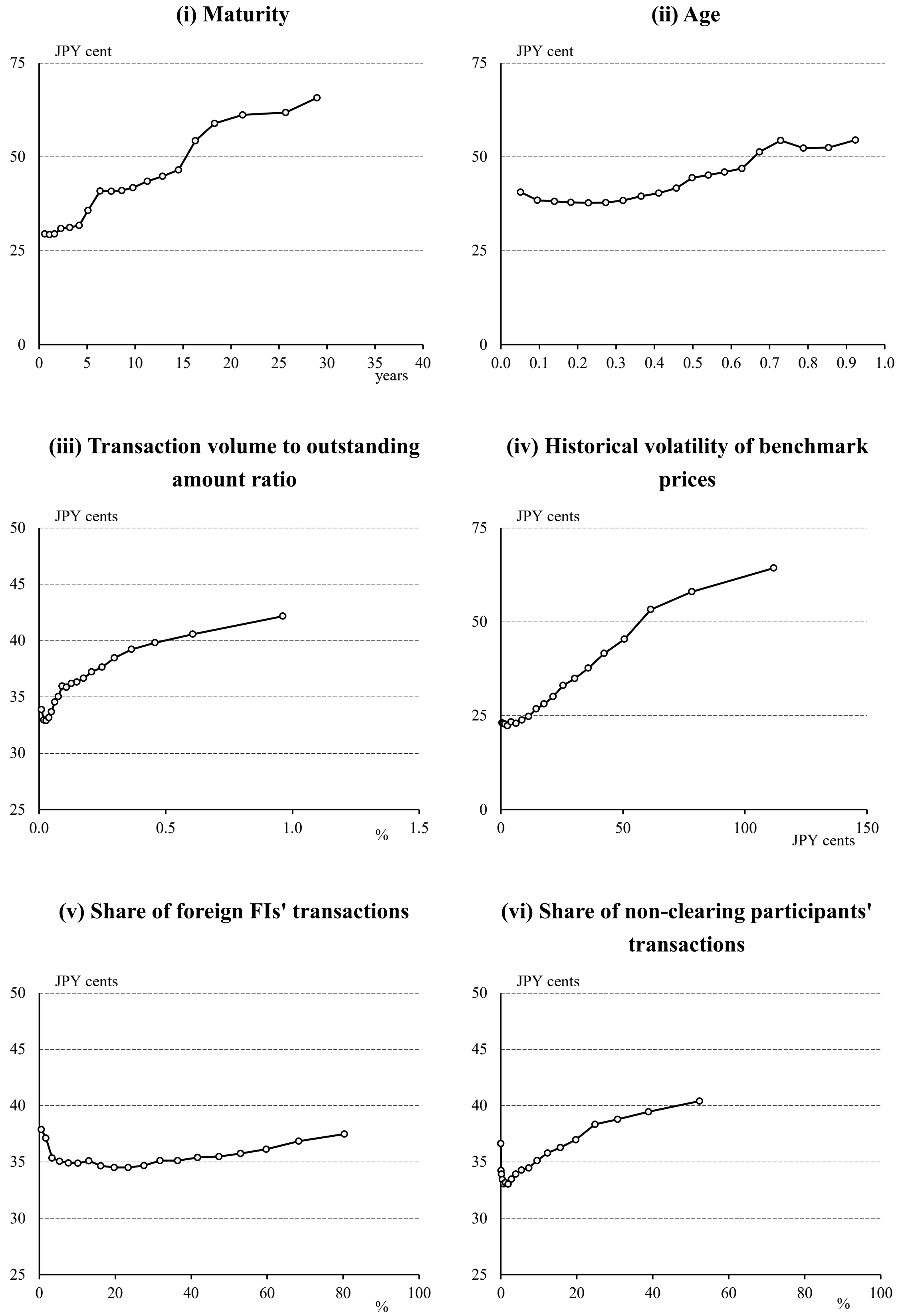
Yeo, Wei Jie, Wihan Van Der Heever, Rui Mao, Erik Cambria, Ranjan Satapathy, and

Gianmarco Mengaldo. "A Comprehensive Review on Financial Explainable AI." *Artificial Intelligence Review* 58 (2025): 189.

## Table 1 Descriptive statistics

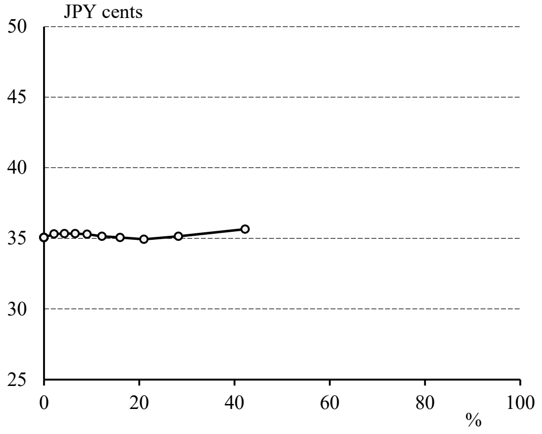
| | | | $Q_{0.10}$ | $Q_{0.25}$ | $Q_{0.50}$ | $Q_{0.75}$ | $Q_{0.90}$ | Mean | Std. dev. |
|---|---|---|---|---|---|---|---|---|---|
| | | Price dispersion (JPY cents) | 1.0 | 4.1 | 17.7 | 44.2 | 81.3 | 34.2 | 58.9 |
| Bond characteristic variables | (i) | Maturity (years) | 1.1 | 3.2 | 8.6 | 16.3 | 25.7 | 10.9 | 9.2 |
| | (ii) | Age | 0.10 | 0.23 | 0.46 | 0.67 | 0.85 | 0.46 | 0.27 |
| Trading activity variables | (iii) | Transaction volume to outstanding amount ratio (%) | 0.02 | 0.05 | 0.13 | 0.30 | 0.61 | 0.30 | 0.70 |
| | (iv) | Historical volatility of benchmark prices (JPY cents) | 1.0 | 4.2 | 17.8 | 42.4 | 78.2 | 31.4 | 41.1 |
| | (v) | Share of foreign FIs' transactions (%) | 1.8 | 7.7 | 23.5 | 47.3 | 68.4 | 29.8 | 25.5 |
| | (vi) | Share of non-clearing participants' transactions (%) | 0.1 | 0.7 | 5.4 | 19.7 | 38.9 | 13.4 | 18.0 |
| | (vii) | Share of large-lot transactions (%) | 0.0 | 0.0 | 0.0 | 12.2 | 28.3 | 9.0 | 15.6 |
| | (viii) | Share of CB's transactions (%) | 0.0 | 0.0 | 0.0 | 10.6 | 33.4 | 9.5 | 17.9 |
| | (ix) | Number of trading FIs | 11.0 | 15.0 | 20.0 | 24.0 | 28.0 | 19.9 | 7.6 |
| | (x) | Number of trading counterparties per institution | 1.6 | 2.0 | 2.6 | 3.4 | 4.2 | 2.8 | 1.1 |
| Position holding variables | (xi) | Share of FIs' own-account holdings (%) | 14.5 | 18.7 | 24.8 | 34.0 | 45.2 | 27.6 | 12.4 |
| | (xii) | Share of CB's holdings (%) | 12.2 | 23.0 | 39.0 | 55.2 | 76.6 | 41.0 | 22.7 |

**Figure 1 Partial dependence plot**

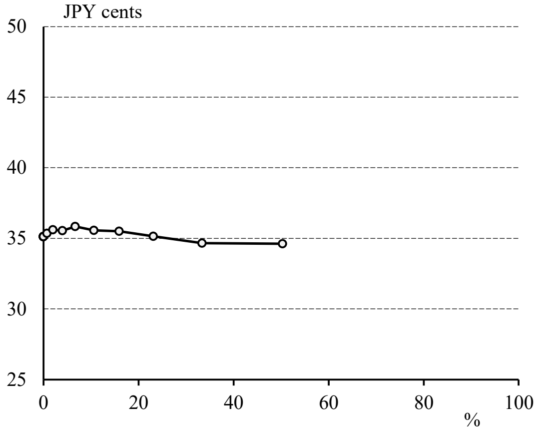


## Figure 1 (cont'd)  Partial dependence plot

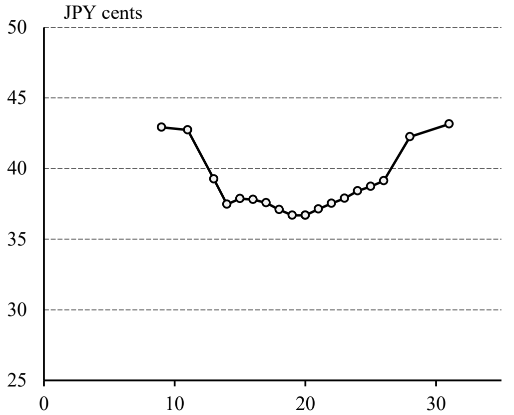
### (vii) Share of large-lot transactions



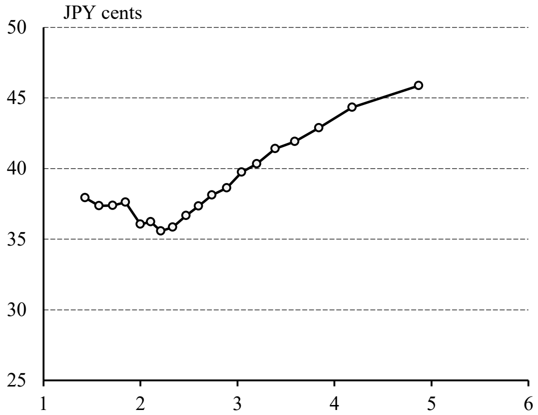

### (viii) Share of CB's transactions



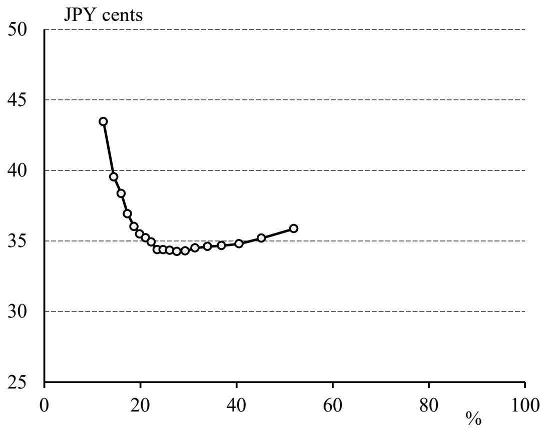

### (ix) Number of trading FIs



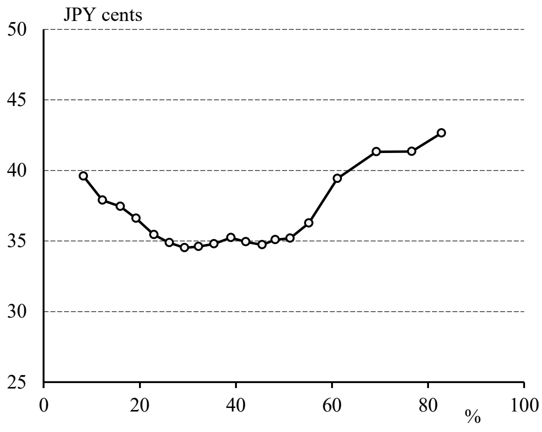

### (x) Number of trading counterparties per institution



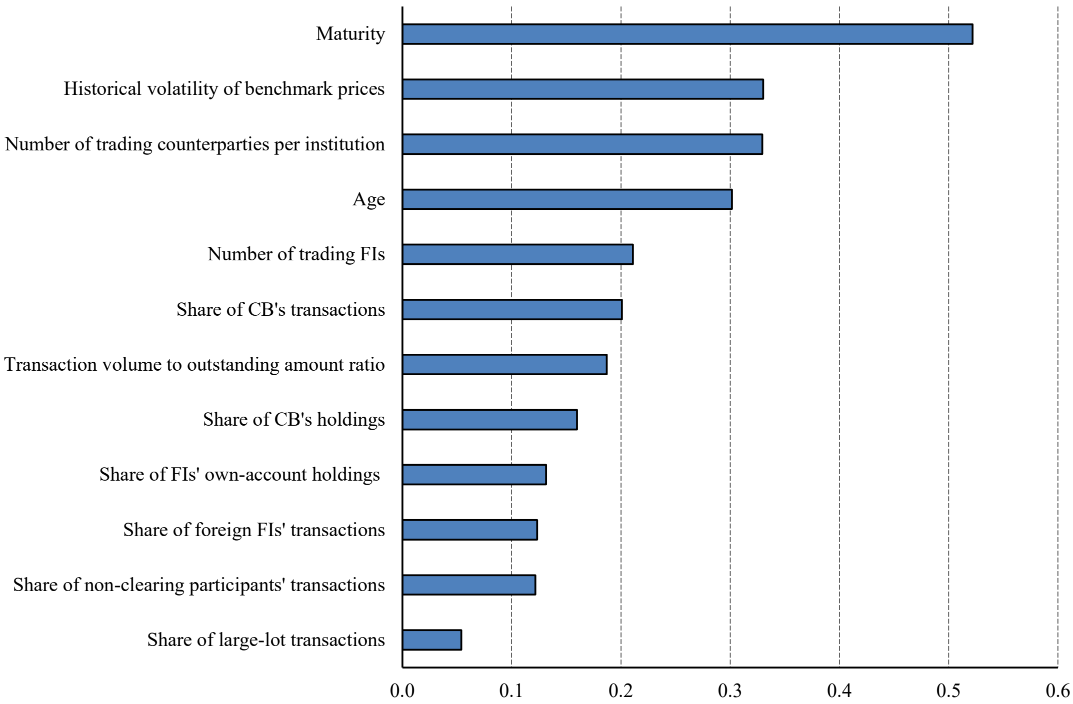

### (xi) Share of FIs' own-account holdings



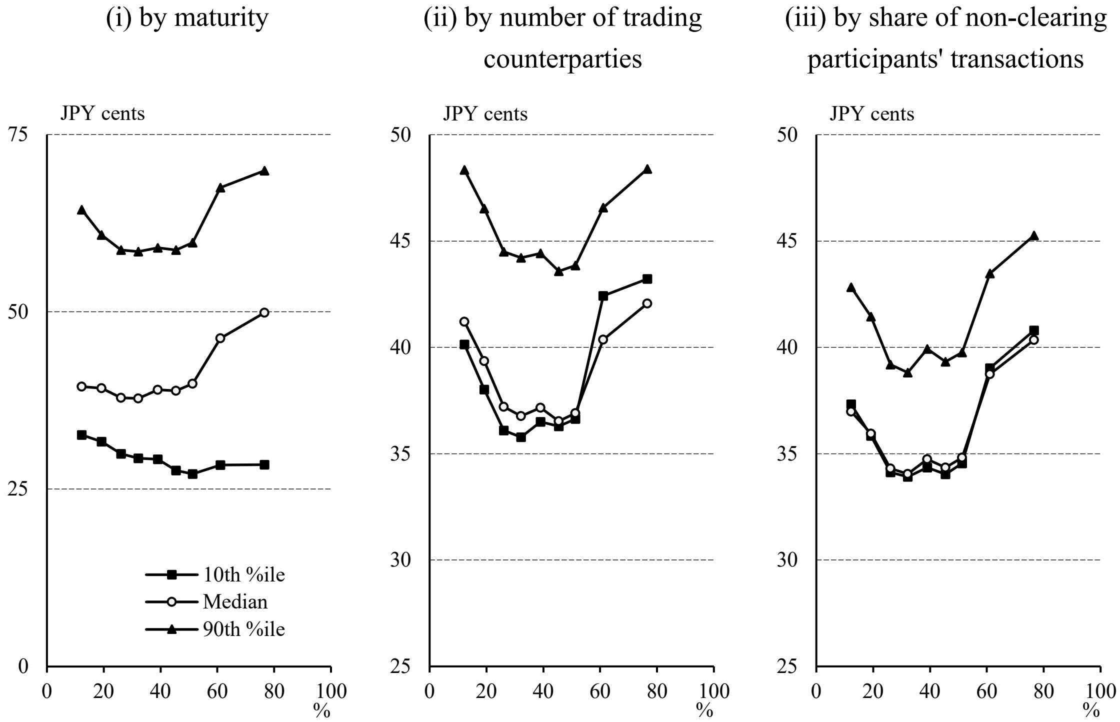

### (xii) Share of CB's holdings



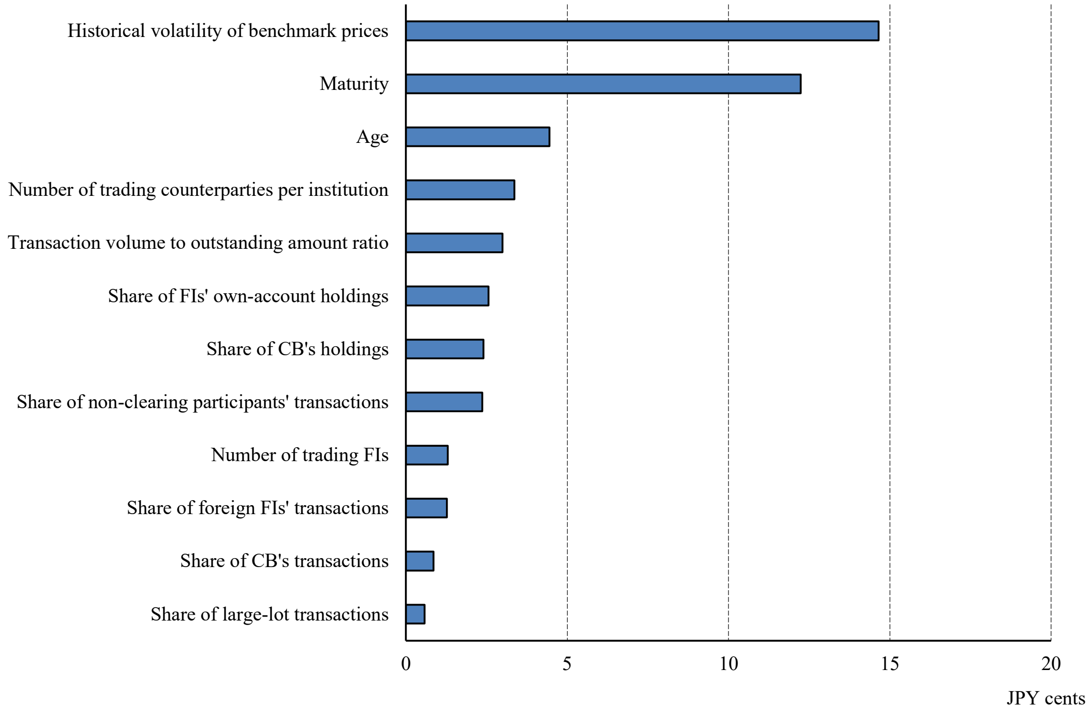
## Figure 2　Feature interaction



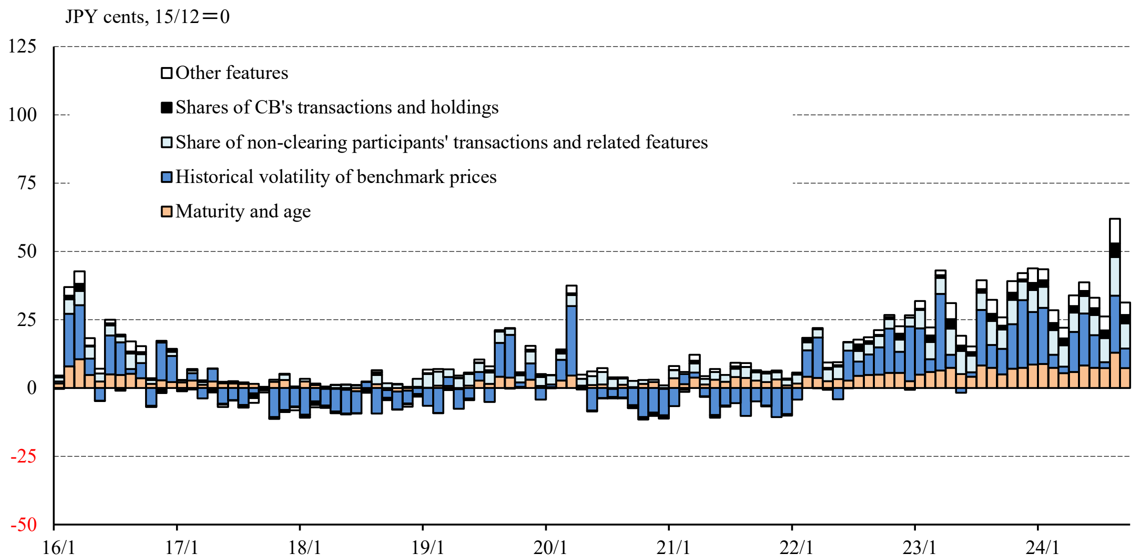

## Figure 3　PDP regarding share of CB's holdings



## Figure 4　Importance of bond features in relation to price dispersion




## Figure 5　Time-series decomposition of price dispersion (all JGBs basis)



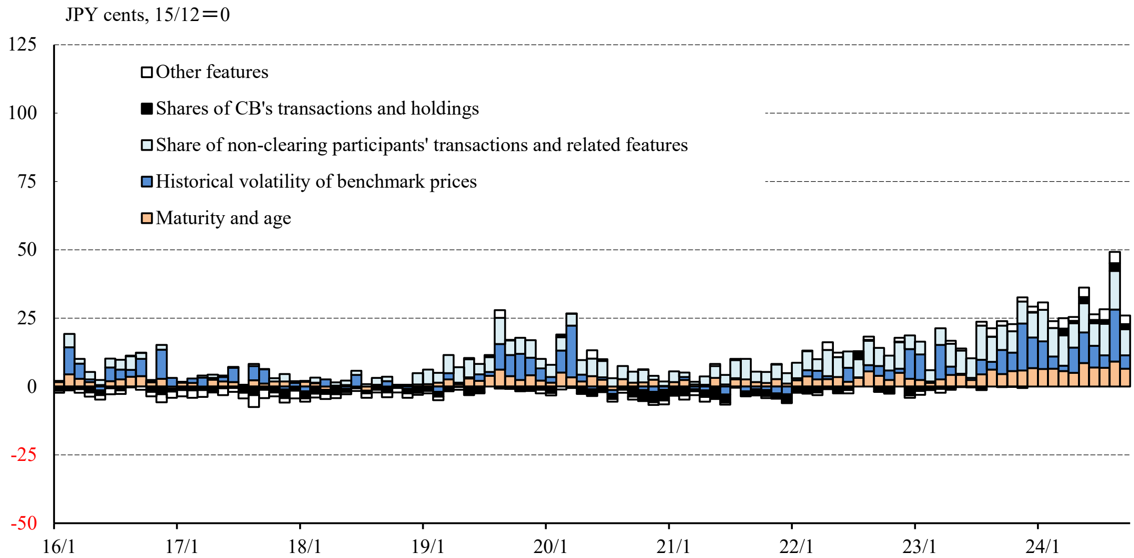

Note: 1. This figure is calculated by aggregating the decomposition for individual JGBs, weighted by their respective transaction volumes.

2. Here and in the following figures, "Share of non-clearing participants' transactions and related features" includes the share of non-clearing participants' transactions, the share of foreign financial institutions' transactions, the share of large-lot transactions, and the transaction volume to outstanding amount ratio.

3. Here and in the following figures, "Other features" includes the number of trading financial institutions, the number of trading counterparties per institution, and the share of financial institutions' own-account holdings.

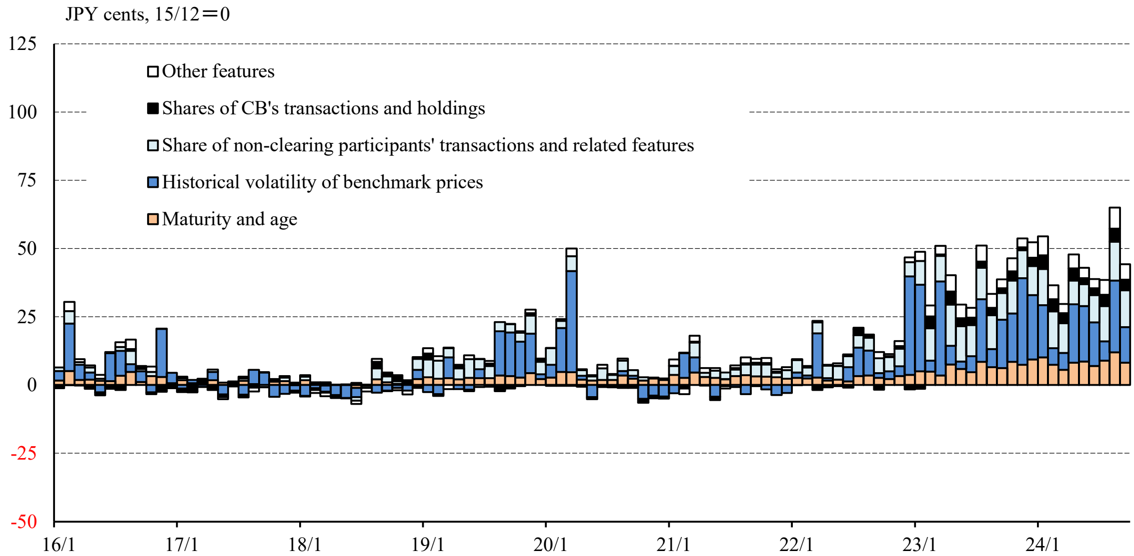
## Figure 6　Time-series decomposition of price dispersion by maturity range

### Medium-term: 1–5 years




Note: 1. This figure is calculated by aggregating the decomposition for individual JGBs in the medium-term range, weighted by their respective transaction volumes.

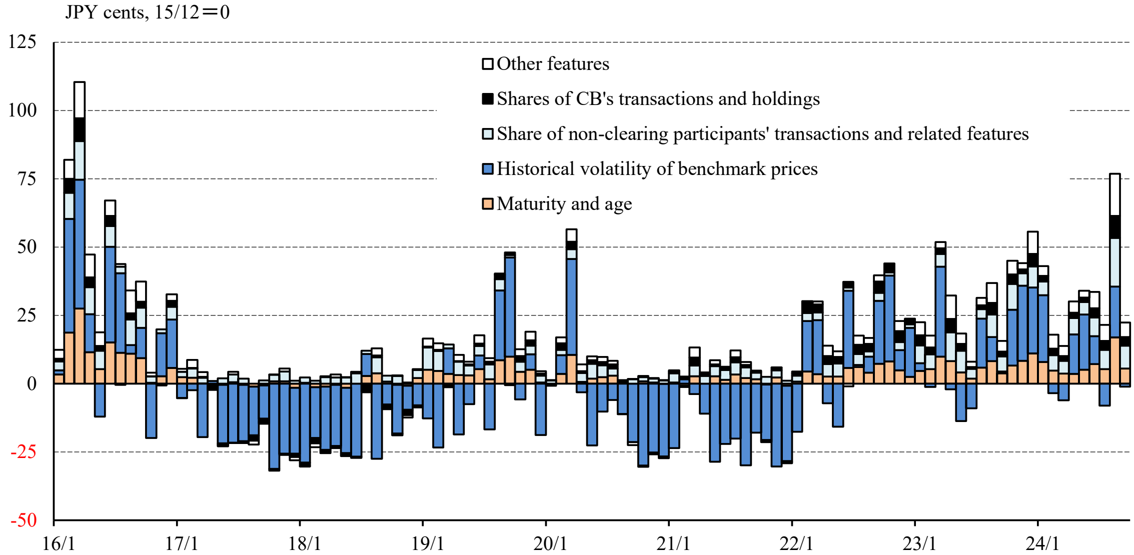
### Long-term: 5–10 years




Note: 1. This figure is calculated by aggregating the decomposition for individual JGBs in the long-term range, weighted by their respective transaction volumes.

## Figure 6 (cont'd) Time-series decomposition of price dispersion by maturity range

### Ultra-long-term: 10–40 years




Note: 1. This figure is calculated by aggregating the decomposition for individual JGBs in the ultra-long-term range, weighted by their respective transaction volumes.