12th biennial IFC Conference: "Statistics and beyond: new data for decision making in central banks"

22-23 August 2024

High-frequency indicators for financial stability: a methodology utilising big data in payment system data[1]

Renardi Ardiya Bimantoro, Charvin Lim and Mohammad Khoyrul Hidayat,
Bank Indonesia

[1] This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

High-Frequency Indicators for Financial Stability A Methodology Utilizing Big Data in Payment System Data

Renardi Ardiya Bimantoro[1], Charvin Lim[2], Mohammad Khoyrul Hidayat[3]

Abstract

The 2008 GFC and subsequent challenges, notably the COVID-19 pandemic and rapid digitalization, highlight the importance to monitor financial stability through granular and high frequency data. This paper proposes a systematic approach to develop novel statistics using big data analytics on high-frequency payment systems data. The methodology categorizes payment activities across government, financial institutions, non-financial corporations, and households, with potential for more detailed classifications as a new data source. The paper also explores the interrelationships between payment system data and indicators of intermediation and financial system soundness. The proposed innovation is expected to serve as a new tool to strengthen policymaking for central banks.

Keywords: financial system stability, new data source, payment system, big data, text mining, machine learning.

JEL classification: G21, C55

[1] Data Innovation & Digitalization Department – Bank Indonesia; e-mail: renardi_ardiya@bi.go.id

[2] Macroprudential Policy Department – Bank Indonesia; e-mail: charvin_lim@bi.go.id

[3] Statistics Department – Bank Indonesia; e-mail: moh_khoyrul@bi.go.id

Contents

1. Background 3
2. Literature Review 4
   2.1 Role of Central Banks in Financial Stability 4
   2.2 Application of AI and Big Data in Financial Stability Analysis 5
   2.3 Big Data Analytics Application for Economic Activiy Analysis 6
3. Data Source & Methodology 7
   3.1 Data Source 7
   3.2 Methodology 8
   3.3 Model Evaluation 11
4. Result and Analysis 11
5. Conclusion and Future Work 15
   5.1 Conclusion 15
   5.2 Future Work 16
References 17

1. Background

The 2008 Global Financial Crisis (GFC) exposed two critical lessons to central banks. Firstly, it underscored the need for expanding central bank mandates beyond maintaining price stability to include safeguarding financial system stability. Secondly, it emphasized the importance of understanding the interplay between macroeconomic conditions and financial system stability, known as macrofinancial linkages (Warjiyo and Juhro, 2019). In the aftermath of the GFC, central banks and regulatory bodies intensified their focus on developing robust frameworks to detect and mitigate financial risks. However, the recent COVID-19 pandemic has introduced new complexities, further emphasizing the critical role of continuous monitoring in a rapidly evolving financial landscape.

The COVID-19 pandemic has not only been a public health crisis but also a severe economic shock, affecting financial markets and institutions globally. The pandemic-induced economic disruptions highlighted the fragility of financial systems and the need for real-time data to manage and respond to emerging risks (Borio, 2020). This period has demonstrated that traditional data sources, often lagged and aggregated, are insufficient for capturing the swift and dynamic nature of financial instability. Instead, there is a growing recognition of the value of granular and high-frequency data in providing timely insights into economic activity and financial health (Carstens, 2020).

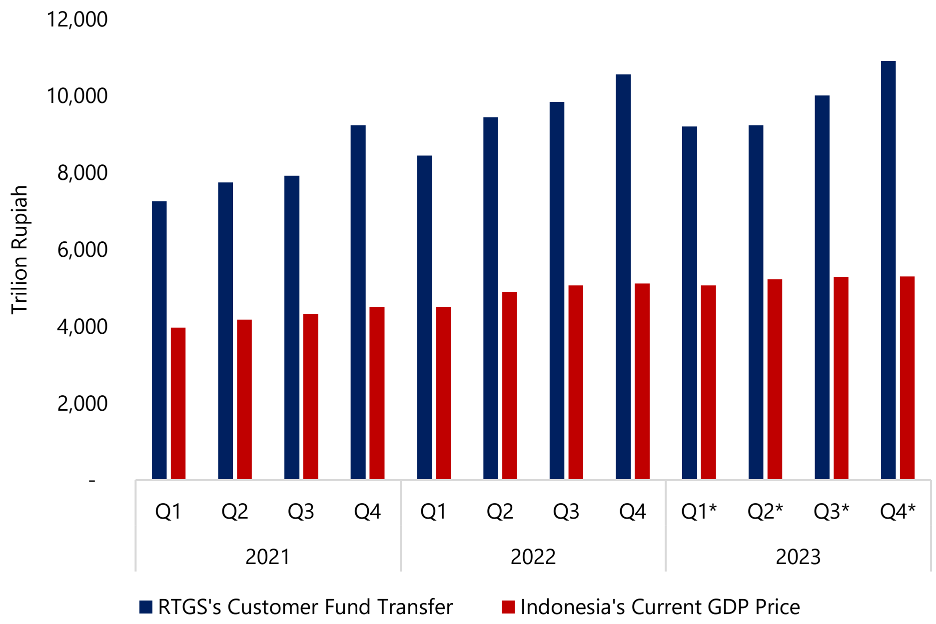
Parallel to these developments, the rapid digitalization of financial services has transformed the financial landscape. The proliferation of digital transactions, fintech innovations, and the use of big data analytics have introduced both opportunities and challenges for financial stability. Digitalization has enabled more efficient and inclusive financial services, but it has also created new channels for risk propagation and increased the complexity of monitoring financial systems (Philippon, 2016). In this context, leveraging advanced methodologies such as artificial intelligence and machine learning to analyze high-frequency transactional data can significantly enhance the ability to monitor and maintain financial stability.

Bank Indonesia, in collaboration with the government, consistently endeavours to develop suitable policies to achieve its three primary objectives: monetary stability, financial system stability, and payment system stability. Bank Indonesia has implemented and operated the BI-RTGS payment system since November 17, 2000. Due to its coverage of about 90% of total payment transactions in Indonesia, BI-RTGS has been designated as a systemically important payment system.

The significant role of the large value payment system (BI-RTGS) in maintaining financial system stability and mitigating systemic risk, makes it crucial for the central bank to thoroughly analyse the structure of the interconnections among players in the payment system. Furthermore, the global financial crisis served as a reminder of the significance of comprehending the interdependence of financial institutions. Hence, it is imperative to ascertain the interconnection of players among banks in the BI-RTGS payment system through network analysis.

Processed RTGS's Customer Fund Transfer Transaction Compared to Indonesia's Current GDP Price in 2021-2023

Figure 1



Building on this need for thorough analysis, this paper explores the importance of using granular and high-frequency payment system data to provide insight on financial stability. It examines the potential of artificial intelligence and machine learning methodologies to provide deeper insights into financial system dynamics, thereby enabling more proactive and informed regulatory responses. This study centers on leveraging a more comprehensive real-time gross settlement (RTGS) dataset to map transactions between users and develop indicators to provide insights on financial system stability.

2. Literature Review

2.1 Role of Central Banks in Financial Stability

Central banks are uniquely positioned to contribute to maintaining financial stability, due to their role as monetary authority and financial sector regulator. Freixas, Laeven, and Peydró (2015) discuss the critical role central banks play in providing liquidity during times of stress, ensuring the stability of financial institutions. Schinasi (2019) underscores the inherent link between monetary stability and broader financial stability, advocating central banks to play pivotal role with their tools and authority to influence monetary policy, manage interest rates, and act as lenders of last resort during financial crises. Juhro and Warjiyo (2020) further emphasize that central banks' mandate extends beyond traditional monetary policy to include oversight of the financial system, aiming to prevent and mitigate systemic risks.

Systemic risk in the financial system manifests in two principal forms: time series risk (procyclicality) and cross section risk (interconnection and common exposure). Time series risk refers to the amplification of financial instability over time, often driven by procyclical behavior where financial institutions' actions during economic booms and busts exacerbate the fluctuations. Cross section risk, on the other hand, stems from the interconnectedness and common exposures among financial institutions, which can lead to the rapid spread of financial distress across the sector (Schinasi, 2019). Acharya, Pedersen, Philippon, and Richardson (2017) provide empirical evidence on how systemic risk can spread through interconnected financial networks, highlighting the importance of monitoring both types of risks.

Financial stability is linked to microfinancial linkages, which highlight the interconnectedness of the banking sector with various segments of the economy, including the government, non-financial corporates, households, and the external sector. These linkages mean that shocks in one sector can quickly propagate through the financial system, affecting overall stability. The banking sector's connections with these diverse economic agents underline the importance of monitoring and managing risks across the entire financial landscape (Juhro & Warjiyo, 2020). Moreover, Brunnermeier, Eisenbach, and Sannikov (2013) emphasize the significance of understanding the interdependencies between financial institutions and other economic agents to maintain stability.

Addressing systemic risks requires a comprehensive policy mix that combines monetary policy, macroprudential regulation, and payment system policy. Claessens, Ghosh, and Mihet (2013) purport the need for integrated policy approaches to address financial stability comprehensively. Warjiyo and Juhro (2020) argue that a well-coordinated policy mix is essential to mitigate both time series and cross section risks effectively. This involves not only traditional monetary measures but also targeted macroprudential policies designed to enhance the resilience of financial institutions and markets against systemic shocks.

Strong analytical frameworks are crucial to support the formulation and implementation of effective policies. This includes detailed analysis of balance sheet interlinkages among financial institutions, which provides insights into the potential pathways of risk transmission within the financial system. Understanding these interlinkages helps in identifying vulnerabilities and crafting appropriate regulatory responses (Schinasi, 2019). Adrian and Liang (2018) underscore the importance of rigorous analytical tools and data to support macroprudential oversight and policymaking.

While balance sheet analysis is fundamental, it can be complemented with flows analysis derived from transaction data reflected in payment systems. Payment system data offers a real-time perspective on the financial flows between various economic agents, providing a dynamic view of the financial interconnections. This flow-based analysis can reveal emerging risks and trends that static balance sheet data might not capture, thus enhancing the robustness of financial stability assessments (Juhro & Warjiyo, 2020). BIS (2018) discusses the value of payment system data in providing timely insights into financial stability. However, there are still limited instances of the utilization of payment system data, primarily due to the big, unstructured, and complex nature of payment system data sets.

2.2 Application of AI and Big Data in Financial Stability Analysis

The integration of AI and big data methodologies has significantly enhanced the analytical capabilities of central banks, particularly in the analysis of payment system data. AI algorithms can process vast amounts of transaction data efficiently, uncovering patterns and anomalies that may indicate financial stress or systemic risk. Big data analytics enables a more granular and timely understanding of financial system dynamics, supporting more proactive and informed policy decisions (Warjiyo & Juhro, 2020). Ng and Kwok (2017) highlight the transformative potential of AI and big data in financial regulation and supervision, enhancing the ability of central banks to monitor and manage financial stability.

AI and big data analytics present both opportunities and risks to financial stability analysis. AI enhances decision-making processes through advanced predictive analytics, enabling central banks to forecast economic trends more accurately and make better-informed policy decisions (Ghandour, 2021). Moreover, AI enhances the capability of central banks to analyze large datasets and derive meaningful insights (Liu, 2023). Another use case of AI is for regulatory compliance and anti-money laundering (AML), in which AI assists in monitoring transactions and identifying suspicious activities in real-time (Vp, 2021). However, the integration of AI in banking also heightens vulnerability to cybersecurity threats; inadequately secured AI systems may become targets for cybercriminals, leading to significant breaches and financial losses (Dhashanamoorthi, 2021). Additionally, the extensive use of AI raises significant issues regarding data privacy and security, necessitating robust measures to safeguard against unauthorized access and misuse (Mithra & Duddukuru, 2023).

2.3 Big Data Analytics Application for Economic Activity Analysis

As the amount of text generated data became abundant, approaches, and tools for extracting relevant information became increasingly important. Named Entity Recognition & Classification (NERC) became crucial to identify an entity based on predefined categories such as individuals, organizations, locations, dates/times, and numbers. Named entity extraction was first introduced at the Sixth Message Understanding Conference in 1996. Later, researchers devised other approaches for extracting entity diversity from various languages and text genres. A variety approaches have been developed starting from rule-based, statistical method, machine learning, deep learning, and transformer. However, there is no universal solution in NERC which has resulted in the development of hybrid approaches.

Several studies have demonstrated the application of big data analytics, including AI and machine learning, in analyzing payment system data to gain insights into economic activity and financial system stability. Sulistiawati et al. (2022) utilized big data analytics on RTGS data to identify publicly listed large non-financial corporations to monitor their transactions. The study demonstrates that text mining methodology, particularly text parsing and entity resolution to enhance data quality prior to analyzing corporate sector activity. The indicator derived from payment system can serve as early indicators for tracking economic activity, validated by the findings, which show good correlation between the indicators of RTGS and sectoral GDP, such as construction, wholesale and retail trade, transportation, and business services. Furthermore, another study by Bimantoro et al. (2022) examined big data analytics to categorize retail value payment system data based on user groups and transaction descriptions, then linked them to household consumption and income. This study demonstrates that text mining methodologies can be applied to large datasets from payment systems to gain deeper insights and enhance economic analysis. Their findings showed a strong correlation between the retail payment system indicator and household GDP consumption, highlighting the utility of payment system data in economic analysis.

Based on the literature, we find out that big data analytics (including AI and machine learning) significantly enhance the capability of central banks to derive meaningful insights from large datasets, including but not limited to payment system as data source. However, our research discovered that only a limited number of studies have employed NERC techniques to enhance the quality of payment system data.

In the case of Indonesia, legacy systems such as BI-RTGS were initially developed primarily to ensure uninterrupted payment system activities rather than the needs of economic analysis. In response to the research question and considering on one hand that RTGS serve as a crucial part of financial infrastructure, facilitating high-value economic transactions, and the other hand, the increasing demand for high frequency data to obtain more comprehensive assessment of economic and financial stability, this paper devises three methodologies to enhance the quality of payment system data further by leveraging big data analytics: record linkage, machine learning, and rule-based approaches.

Thus, this study is expected to answer the following research questions:

1. How can central banks further enrich their payment systems data to support policy making for financial stability?
2. How can the resulting indicator from enriched payment systems data complement existing indicators on financial intermediations and financial sector soundness?

3. Data Source & Methodology

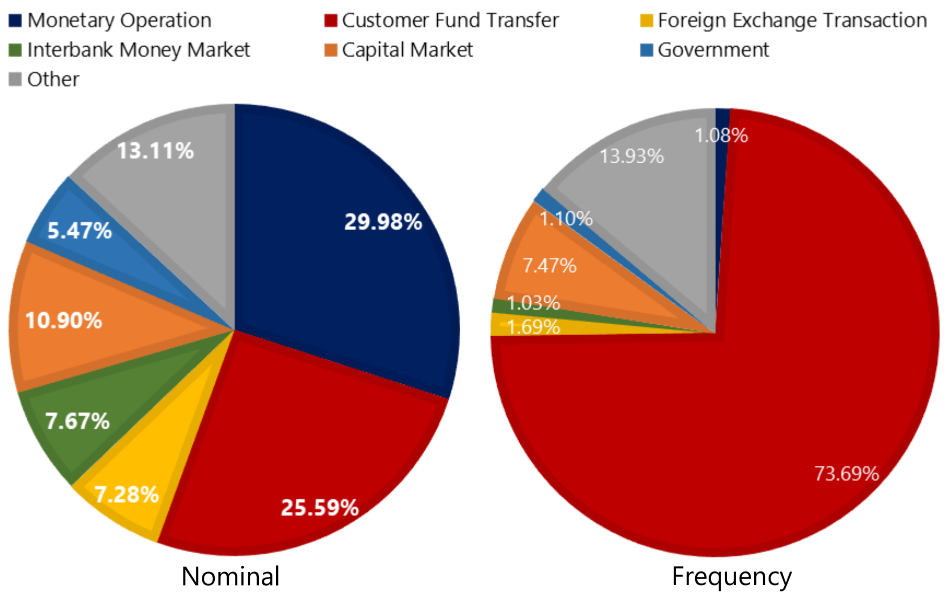
3.1 Data Source

The study utilized data sourced from the Bank Indonesia Real Time Gross Settlement (BI-RTGS) and Bank Reports Data. BI-RTGS is a financial infrastructure managed by Bank Indonesia that facilitates high-value fund transfers, debit clearing, regular payments, and regular billing services (as stated in Regulation of Member of Board of Governors Number 24/5/PADG/2022). Since November 2000, the Real Time Gross Settlement (RTGS) system has had the capability to handle payment transactions exceeding Rp. 100,000,000.00 (one hundred million rupiah) in value.

The data included in this study encompasses RTGS transaction data spanning from January 2016 to May 2024. On average, the value of transactions through this service amounts to approximately 13,000 Trillion Rupiah per month. The Real Time Gross Settlement (RTGS) system encompasses various transaction types. The share for each transaction type is shown in Figure 2, which shows that the largest share of transaction (±29%) is monetary operation, followed by customer fund transfer (±25).

Nominal and Frequency Share of RTGS Transactions Based on Transaction Type in November 2023

Figure 2



Source: BI-RTGS (processed)

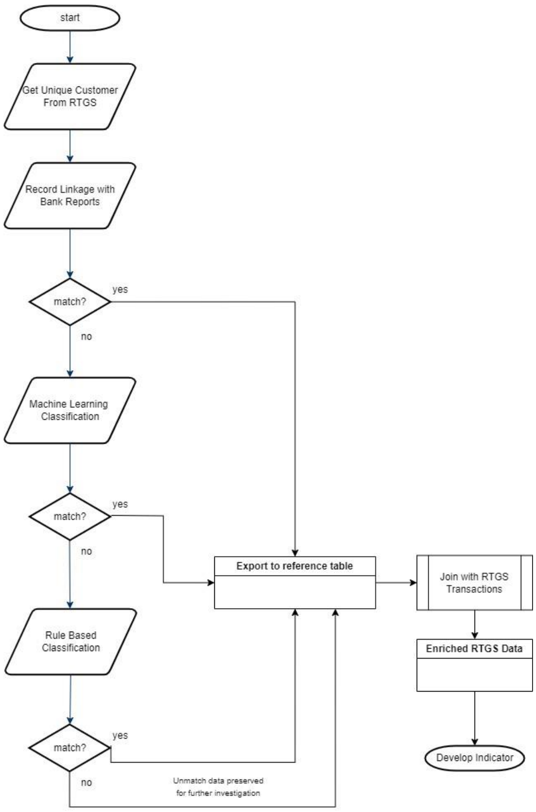
For the purposes of our study, we focus solely on customer fund transfer transactions. To enrich the BI-RTGS data, we integrate it with data from bank reports, which comprise millions of entities already categorized by banks. However, the matching rate between the BI-RTGS data and bank reports are not 100% matched. One of the reasons for this is the varying data governance practices across different banks which leads to inconsistent data. Additionally, some accounts may already be deactivated and thus not included in the bank reports. Therefore, further enhancements using machine learning methods are necessary to improve the matching accuracy and overall reliability of the analysis. To achieve that, we use labeled bank reports dataset as the training data for our machine learning model.

3.2 Methodology

We propose a hybrid method using record linkage, machine learning model, and rule-based classification to classify individuals, financial institutions, non-financial corporation, and governments in the payment system. Beginning with unique customer data from the RTGS system, we linked it to bank report categorizations using record linkage method using account number. Direct matches were followed by machine learning and rule-based categorization ensuring comprehensive data integration. This hierarchical approach classified all data and saved mismatched data for further investigation. The enriched transaction data, which now shows more entity classification, will be used to develop more financial indicators, improving our financial analysis and decision-making. The entity classification workflow is shown in Figure 3.

Entity Classification Workflow

Figure 3



The first step in our machine learning process involves text processing from bank reports data, which is crucial for preparing textual data for machine learning models and rule-based classification. The primary steps in text processing include word tokenization, vectorization using Count Vectorizer, and string labeling using String Indexer. The feature generated then used to develop machine learning model and applied for rule-based classification. The process is as follows:

1. Word Tokenization

The main objective of tokenization is to encode text in a way that is significant for machines while preserving its context. Through the process of tokenization, algorithms can efficiently detect and analyze patterns in text. Pattern recognition is essential as it enables machines to comprehend and react to human input. This technique dissects text into separate words. To further explore the mechanics, let us examine the line, "Chatbots are beneficial." Upon tokenizing this sentence, it undergoes a transformation into an array consisting of separate words like this ["Chatbots", "are", "helpful"].

2. Count Vectorizer

Machines lack the ability to comprehend characters and words. When working with text data, it is necessary to convert information into numerical form for the machine to comprehend it. Count Vectorizer is a technique used to transform text into numerical representations where the text is converted into a sparse matrix and the row corresponds to the number of words.

3. String Labelling

To handle categorical labels, we use String Indexer to convert string labels into numerical indices. This step is necessary for classifiers that require numerical input for the target variable.

4. Machine Learning Model

We utilize machine learning techniques to categorize the entity based on the features gathered during text processing. This categorization is carried out using logistic regression, SVM, and Naive Bayes algorithm. We take two approaches: one-versus-rest and direct multinomial categorization. One-versus-rest is utilized as an estimator, with the base classifier accepting Classifier objects and producing binary classification problems for each of the k classes. The classifier for class i is trained to predict whether the label is i or not, hence separating it from the other classes. Predictions are made by evaluating each binary classifier and using the index of the most confident classifier as a label. As for direct multinomial approach, we employed two multiclass algorithms directly: Naive Bayes and Multinomial Logistic Regression.

5. Rule-based Method

The rule-based method uses specified rules to create classifications. In this scenario, we use tokenized text and a database of top words associated with specific categories to divide text into categories. Every token in the text is compared to a predefined lexicon of top words. There are keywords linked with each category as shown in Table 1. For example, given the text "PT Manulife BCA Finance", tokenizing the text yields tokens such as "Finance", "BCA", and "Manulife". These tokens match keywords in the "Financial Institution" category, leading to the classification of the text under "Financial Institution".

Rule-based Keywords

Table 1

Category	Keywords
Financial Institution	finan, life, asuransi, insurance, bank, bpd, bpr, reksadana, sekuritas, dana, reksa, bank, asuransi, bpr, syariah, dplk, fund, jiwa, sekuritas, finance, life, pensiun, saham, money , danamas, kas, likuid, terproteksi, koperasi, equity, financial, insurance, rupiah,
Non-Financial Corporation	restaurant, automotive, otomotif, logistikc, product, industry, industries, pharma, prod, dairy, hotel, green, consulting, textile, rsia, rsij, rs, rumah, sakit, research, distribusi, medical, astra, husada, trade, agriculture, agrikultur, food, coffee, operation, maintenance, steel, supplies, jaya, abadi, karya, nusantara, perkasa, prima, lestari, makmur, persada, graha
Government	otoritas, badan, dinas, kab, kabupaten, kec, kecamatan, kel, kelurahan, kpu, pemerintah, pempov, pemkab, pemkot, provinsi, prov, rkud, depdikbud, rkun, embassy, ditjen, direktorat, jenderal, menteri, kementerian, kemendikbud, disdikbud, camat, pdam, kas, umum, kasda, rpkbun, span, pajak, setda, sekda, sekretariat, polres, polda, polsek, kepolisian, pamong, praja, dewan, perwakilan, rakyat, dpr, mpr, bendahara, umum, daerah, bendum, rek, daerah, penerimaan, pendapatan, desa, kab, daerah, pemerintah, dinas, kpknl, kasda, polres, kota, kabupaten, rekening, bendahara, rkud, aceh, palembang, bpp, pemkab, prov, polda, provinsi
Individual	Selected top keywords from tokens that represent individual. (Due to compliance with data protection regulations, the names of individuals are not displayed)

3.3 Model Evaluation

To assess the performance the algorithm that we use, first we compare their accuracy, precision, recall, and F-1 score. The performance indicators demonstrate no major differences amongst the algorithms. However, the Multinomial Logistic Regression slightly surpasses the others in terms of accuracy, precision, recall, and F-1 score, as shown in Table 2. Next, we thoroughly examine each categorization precision to confirm its strength and reliability.

Classification Model Overall Error Metrics

Table 2

	Overall Error Metrics			
Algorithm	**Accuration**	**Precision**	**Recall**	**F-1 Score**
Support Vector Machine	**0.82**	**0.81**	**0.82**	**0.81**
Binomial Logistic Regression	**0.82**	**0.81**	**0.82**	**0.81**
Multinomial Logistic Regression	**0.83**	**0.82**	**0.83**	**0.82**
Naïve Bayes	**0.82**	**0.81**	**0.82**	**0.81**

Table 3 demonstrates that all models excel at identifying individual entities, but there are rooms for improvement in other categories to enhance its precision. For example, we can add more stages to the classification process. First, we build a model to divide the entity into individual and non-individual groups. And then build another model to classify non-individual entities as financial institutions, non-financial corporations or government. This two-step method could leverage model's precision in predicting individual entities while also improving classification precision in other categories by focusing on more detailed distinctions in the second stage.

Precission Comparison Between Algorithm

Table 3

	Precision for Multinomial Logistic Regression			
Algorithm	**Individual**	**Non-Financial Corp.**	**Financial Inst.**	**Government**
Support Vector Machine	92.94	77.59	70.03	78.41
Binomial Logistic Regression	92.27	77.27	70.79	78.67
Multinomial Logistic Regression	92.94	77.60	69.03	78.42
Naïve Bayes	92.94	77.60	69.03	78.42

4. Result and Analysis

Through entity classification, we generate statistics on RTGS transactions, which can be monitored on a monthly basis. These statistics illustrate the RTGS flows undertaken by various economic agents, including financial institutions, non-financial corporations, households, government, and others, thereby reflecting the economic activity of these agents. Furthermore, RTGS statistics reveal the flow of money and liquidity across different sectors.

During the observation period (2016-2023), RTGS transactions exhibited volatile growth, corresponding with the broader economy. In 2023, RTGS transactions reached IDR 39,368 trillion, approximately double Indonesia's GDP. Overall, RTGS transactions experienced strong growth from 2017 to 2019 (ranging from 9.4% to 17.4%), a contraction in 2020 (-8.9%) due to the economic slump caused by the COVID-19 pandemic, and a robust recovery in 2021-2022 (15.3%-19.1%), followed by limited growth in 2023 (2.8%).

Non-financial corporations consistently contributed the most to RTGS transactions, accounting for approximately 58.0%-62.3% of transactions during the observation period. This was followed by financial institutions (approximately 16.6%-24.8%) and households (approximately 9.8%-16.0%). Although the government's share of RTGS transactions was limited (<5.0%), its contribution increased significantly in the past four years, particularly during the COVID-19 period.Table 1. Total RTGS transactions by economic agents (2016-2023)

Total RTGS Transactions (2016-2023)

Table 4

RTGS Transactions (by sender)	2016	2017	2018	2019	2020	2021	2022	2023
Total Amount (IDR Trillion, current prices)								
Financial Institutions	3,517	4,844	5,389	6,289	6,692	7,645	9,490	9,388
Non-Financial Corporations	12,618	14,770	17,432	18,479	16,171	19,323	23,433	24,313
Households	3,072	3,584	4,390	4,893	3,647	3,446	3,760	3,933
Government	134	294	388	491	962	1,316	1,121	1,252
Others	1,790	1,308	390	458	418	438	501	482
Total	21,131	24,800	27,990	30,610	27,890	32,167	38,304	39,368
RTGS Trx Composition (as a percent)								
Financial Institutions	16.6	19.5	19.3	20.5	24.0	23.8	24.8	23.8
Non-Financial Corporations	59.7	59.6	62.3	60.4	58.0	60.1	61.2	61.8
Households	14.5	14.4	15.7	16.0	13.1	10.7	9.8	10.0
Government	0.6	1.2	1.4	1.6	3.5	4.1	2.9	3.2
Others	8.47	5.28	1.39	1.50	1.50	1.36	1.31	1.22
Annual growth (year on year percentage change)								
Financial Institutions		37.7	11.3	16.7	6.4	14.2	24.1	-1.1
Non-Financial Corporations		17.1	18.0	6.0	-12.5	19.5	21.3	3.8
Households		16.6	22.5	11.5	-25.5	-5.5	9.1	4.6
Government		120.1	31.7	26.5	96.1	36.8	-14.9	11.7
Others		-26.9	-70.2	17.4	-8.7	4.7	14.4	-3.7
Total		17.4	12.9	9.4	-8.9	15.3	19.1	2.8

Total RTGS Transactions, by Receiver (2016-2023)

Table 5

RTGS Transactions (by sender)	2016	2017	2018	2019	2020	2021	2022	2023
Transfer to Financial Institutions (IDR Trillion, current prices)								
Financial Institutions	1,522	2,166	2,730	3,247	3,411	3,678	4,863	4,662
Non-Financial Corporations	2,264	2,651	3,143	2,921	2,166	2,402	2,586	2,436
Households	555	644	777	1,110	959	895	795	848
Government	33	58	129	150	296	355	303	284
Others	295	246	96	203	137	123	135	133
Total	4,669	5,765	6,875	7,632	6,968	7,453	8,682	8,363
Annual growth (year on year percentage change)								
Financial Institutions		42.3	26.0	18.9	5.0	7.8	32.2	-4.1
Non-Financial Corporations		17.1	18.5	-7.1	-25.8	10.9	7.7	-5.8
Households		16.0	20.7	42.9	-13.7	-6.7	-11.2	6.7
Total		23.5	19.3	11.0	-8.7	7.0	16.5	-3.7
Transfer to Non-Financial Corporations (IDR Trillion, current prices)								
Financial Institutions	1,204	1,763	1,762	1,986	2,256	2,685	3,093	2,898
Non-Financial Corporations	7,123	9,515	11,948	13,091	12,255	14,955	18,513	19,394
Households	1,378	1,663	2,304	2,268	1,282	1,237	1,405	1,543
Government	38	104	145	176	229	364	357	382
Others	529	473	150	154	166	193	250	249
Total	10,273	13,517	16,309	17,676	16,188	19,435	23,618	24,466
Annual growth (year on year percentage change)								
Financial Institutions		46.4	0.0	12.7	13.6	19.0	15.2	-6.3
Non-Financial Corporations		33.6	25.6	9.6	-6.4	22.0	23.8	4.8
Households		20.6	38.6	-1.6	-43.5	-3.5	13.5	9.9
Total		31.6	20.7	8.4	-8.4	20.1	21.5	3.6
Transfer to Households (IDR Trillion, current prices)								
Financial Institutions	340	467	590	672	599	708	846	1,045
Non-Financial Corporations	821	1,192	1,588	1,645	1,150	1,392	1,644	1,698
Households	497	763	1,029	1,184	1,201	1,113	1,283	1,237
Government	14	29	38	57	94	133	116	135
Others	147	223	53	59	66	77	67	53
Total	1,820	2,674	3,298	3,617	3,110	3,423	3,955	4,168
Annual growth (year on year percentage change)								
Financial Institutions		37.3	26.3	14.0	-10.9	18.2	19.4	23.5
Non-Financial Corporations		45.2	33.2	3.6	-30.1	21.1	18.1	3.3
Households		53.3	34.9	15.1	1.5	-7.3	15.2	-3.6
Total		46.9	23.3	9.7	-14.0	10.1	15.6	5.4

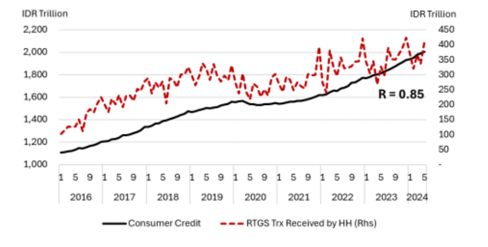
To explore the usefulness of RTGS data in inferring the condition of financial system stability, this study examines the interrelatedness of RTGS transactions with intermediation and financial soundness indicators. Regarding intermediation, the study investigates the relationship between consumer credit and RTGS transfers received by households. Additionally, it explores the interrelation between working capital credit and RTGS transactions received by Non-Financial Corporations (NFCs).

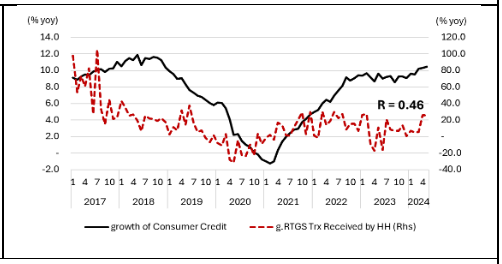
On the aspect of financial sector soundness, the analysis delves into the connection between RTGS transfers and credit risk for both households and NFCs. By examining these relationships, the study aims to provide insights into how RTGS data can serve as a valuable tool for monitoring and understanding financial system stability. This approach underscores the potential of RTGS transactions to reflect the broader dynamics within the financial sector, thereby offering a comprehensive view of economic health and risks faced by different market participants.

RTGS Transactions and Intermediation Indicator

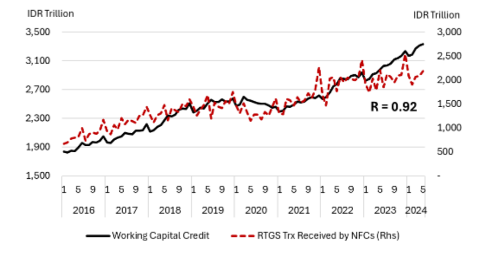
Figure 4

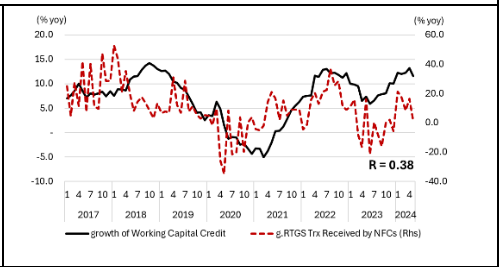
Panel A. Interrelation between RTGS transactions and consumer credit





Panel B. Interrelation between RTGS transactions and working capital credit



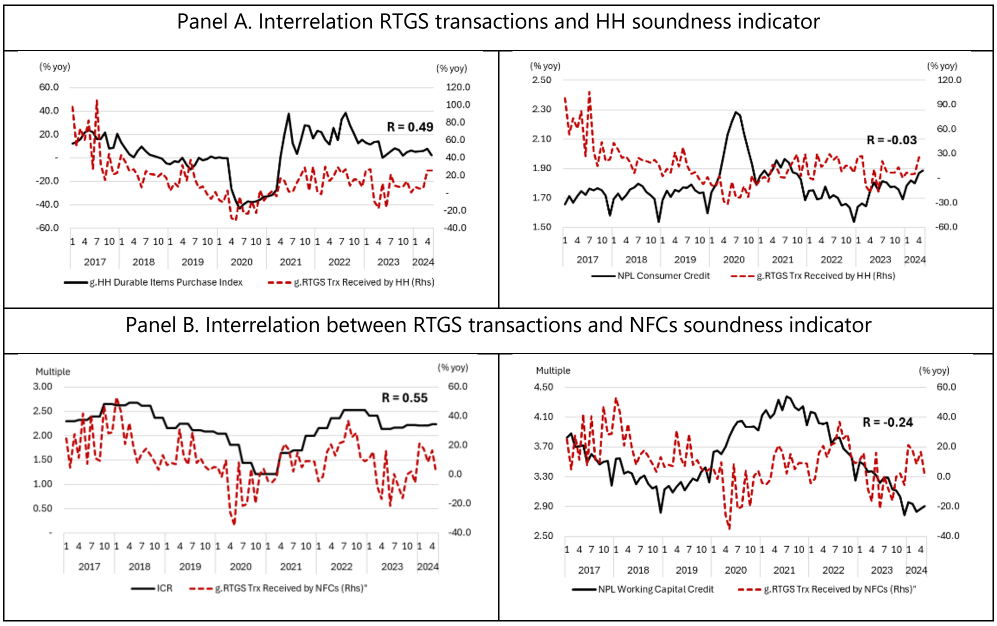


RTGS transactions exhibit a high correlation with both consumer and working capital credit. Specifically, the correlation coefficients are 0.85 for consumer credit and 0.92 for working capital credit (Figure 4). This positive correlation suggests that RTGS transactions, serving as a proxy for economic activity, align closely with the demand for credit by households and non-financial corporations. A low volume of RTGS transactions received indicates reduced economic activity, reflecting a decreased need for credit among households and NFCs, and vice versa.

Furthermore, the growth trends in RTGS transactions can provide early indicators of subsequent movements in both consumer and working capital credit growth. A decline in RTGS transactions received by households is typically followed by a slower growth rate in consumer credit, and vice versa. Similarly, fluctuations in RTGS transactions received by NFCs are followed by corresponding changes in the growth of working capital credit. This relationship indicates the potential of RTGS transaction data as a predictive component for credit demand trends, highlighting its importance in economic analysis and policymaking.

We examined the relationship between RTGS transactions and several financial sector soundness indicators. For households, we utilized the durable items purchase index and the non-performing loans (NPL) of consumer credit. For NFCs, we employed the interest coverage ratio (ICR) and the NPL of working capital credit.

Figure 5



Our findings indicate that RTGS transactions are more effective in reflecting the soundness of NFCs compared to households. As shown in Figure 5, RTGS transactions exhibit a relatively low correlation with household durable item purchases (0.49) and a slight negative correlation with the NPL of consumer credit (-0.03). The low correlation with household NPL could be influenced by the COVID-19 restructuring programs, which disrupt NPL values. In contrast, RTGS transactions show a moderate correlation with the ICR of NFCs (0.55) and a negative correlation with the NPL of working capital credit (-0.24), as anticipated. These correlations indicate that RTGS transaction data is a more dependable indicator for NFCs financial health, than it is of household financial soundness. This disparity may be attributed to households' greater reliance on retail payment system infrastructure rather than RTGS.

5. Conclusion and Future Work

5.1 Conclusion

In this study, we have proposed a novel approach to enrich payment system data by leveraging big data analytics and machine learning methodology to categorize entity into individuals, financial institutions, non-financial corporations, and government. We applied multinomial logistic regression model as our pilot implementation of the enriched RTGS data.

To explore the potential indicator produced by the data, we examined its interrelationship with intermediation and financial sector soundness indicators. We found that RTGS transactions hold significant potential as a proxy of credit demand, as evidenced by their strong relationship with consumer credit and working capital credit. In the context of financial sector soundness, RTGS transactions provide strong correlation with NFCs soundness indicator. However, we found limited potential of RTGS transactions data to indicate household soundness. This limitation can arise due to household transactions predominantly relying on retail payment system infrastructure.

Another potential application derived from our enhanced payment system data is to identify the interconnectedness between banks and customers within the payment network. By employing network analysis and clustering techniques, we attempt to explore the patterns of interbank relationships that occurred from 2018 to 2023, including the pandemic period. In this use case, we focus on transactions related to financial institutions and non-financial companies, which comprise the majority of RTGS transactions. However, these preliminary results must be validated to assess their robustness. Therefore, we will continue to examine this in our future research.

5.2 Future Work

There are several improvements in the methodology that can be applied for future works.

1. Develop separate model to initially divides the entity into individuals and non-individuals and create another model to classify non-individuals as financial institutions, non-financial corporations, and government.
2. Enhance the data source to encompass the whole interbank payment system in Indonesia, including but not limited to retail payment system such as fast payment and card-based transaction.
3. Incorporate additional economic indicators into the BI-RTGS transactions to expand validation data for the developed data.
4. Identify the interconnectedness within the payment system network to understand the behavioral analytics of banks as intermediaries of business activity. This will be achieved by leveraging network analysis and clustering methods to examine further our enhanced payment system data.

References

Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, Matthew Richardson. (2017). Measuring systemic risk. *The Review of Financial Studies, 30*(1).

Adrian, T., & Liang, N. (2018). Monetary policy, financial conditions, and financial stability. *International Journal of Central Banking, 14*(1), 73-131.

Bimantoro, R. A., Hidayat, M. K., Jabbar, M. A., & Zulen, A. A. (2022). Big data analytics on payment system data for measuring household consumption in Indonesia. 11th Biennial IFC Conference on "Post-pandemic landscape for central bank statistics.

BIS. (2018). BIS Annual Economic Report 2018. Bank for International Settlements.

Borio, Claudio. (2020). The Covid-19 economic crisis: dangerously unique. *Business Economics, Palgrave Macmillan; National Association for Business Economics, 55*(4), 181-190.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives, 23*(1), 77-100.

Brunnermeier, M. K., Eisenbach, T. M., & Sannikov, Y. (2013). Macroeconomics with Financial Frictions: A Survey. *Advances in Economics and Econometrics, 2*, 3-94.

Carstens, Agustín. (2020). Shaping the future of payments. *BIS Quarterly Review*, Bank for International Settlements, March.

Claessens, S., Ghosh, S. R., & Mihet, R. (2013). Macro-Prudential Policies to Mitigate Financial System Vulnerabilities. *Journal of International Money and Finance*, 39, 153-185.

Dhashanamoorthi, Balaji. (2021). Artificial intelligence in combating cyber threats in Banking and Financial services. *International Journal of Science and Research Archive*, 210-216.

Ek, T., Kirkegaard, C., Jonsson, H., & Nugues, P. (2011). Named entity recognition for short text messages. *Procedia - Social and Behavioral Sciences, 27*, 178-187.

Ettrich, O., Stahlmann, S., Leopold, H., & Barrot, C. (2024). Automatically identifying customer needs in user-generated content using token classification. Decision Support Systems.

Freixas, Xavier, Luc Laeven, and José-Luis Peydró. (2015). Systemic risk, crises, and macroprudential regulation. The MIT Press.

Ghandour, A. (2021). Opportunities and challenges of artificial intelligence in banking: Systematic literature review. *TEM Journal, 10*(4), 1581-1587.

Goyal, A., Gupta, V., & Kumar, M. (2018). Recent named entity recognition and classification techniques: A systematic review. *Computer Science Review*.

Juhro, S. M., & Warjiyo, P. (2020). Central Banking Policy: Theory and Practice. Bank Indonesia.

Liu, X. (2023). New characteristics of AI analysis of financial risks and their transmission traceability paths. *International Journal of Big Data Intelligent Technology*. https://doi.org/10.38007/ijbdit.2023.040105.

Mithra, A., Duddukuru, V., & S., M. (2023). How artificial intelligence is revolutionizing the banking sector: The applications and challenges. *Asian Journal of Management*. https://doi.org/10.52711/2321-5763.2023.00028.

Ng, Artie & Kwok, Benny. (2017). Emergence of Fintech and cybersecurity in a global financial centre: Strategic approach by a regulator. *Journal of Financial Regulation and Compliance*.

Ng, A. W., & Kwok, B. K. (2017). The impact of big data and AI on financial regulation. *Journal of Financial Regulation and Compliance, 25*(2), 253-270.

Philippon, Thomas. (2016). The Fintech opportunity. *NBER Working Papers 22476*, National Bureau of Economic Research, Inc.

Schinasi, G. J. (2019). Understanding financial stability: Towards a practical framework. International Monetary Fund.

Vp, R. (2021). Role of AI in Banking. *Banking & Insurance eJournal*. https://doi.org/10.2139/ssrn.3803749.

Warjiyo, P., & Juhro, S. M. (2019). Introduction to Central Bank Policy: Theory and Practice. Emerald Publishing Limited: Leeds.

Zhang, Y., Xiao, G. (2024). Named entity recognition datasets: A classification framework. *International Journal Computer Intelligence Syststem, 17*.

DIDD

DEPARTEMEN INOVASI & DIGITALISASI DATA

Data at your fingertips.

BANK INDONESIA

BANK SENTRAL REPUBLIK INDONESIA

High-Frequency Indicators for Financial Stability – A Methodology Utilizing Big Data in Payment System Data

23 August 2024

Renardi Ardiya Bimantoro, Charvin Lim, Mohammad Khoyrul Hidayat

Department of Data Innovation and Digitalization, Department of Macroprudential Policy, Department of Statistics – Bank Indonesia

Email: renardi.ardiya@bi.go.id; charvin_lim@bi.go.id; moh.khoyrul@bi.go.id

BACKGROUND

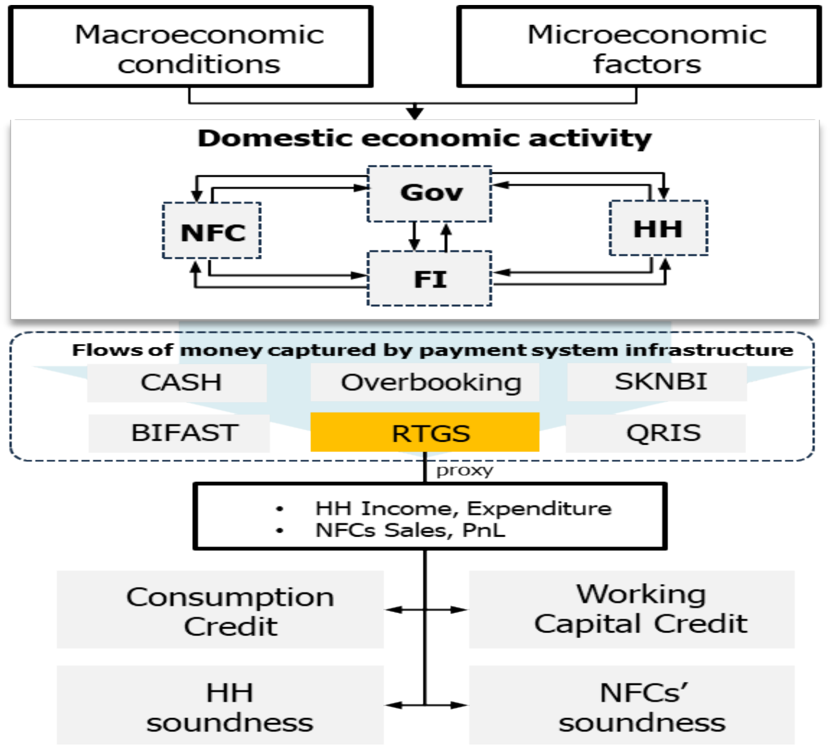
STRATEGIC ENVIRONMENT

- The 2008 GFC highlighted the need for central banks to expand their mandates to include financial system stability.
- The COVID-19 pandemic showed that traditional, lagged data sources are inadequate for capturing rapid economic changes.
- There is increasing recognition of the importance of granular and high-frequency data for timely economic and financial insights, improving risk management (Borio, 2020; Carstens, 2020).
- Digitalization of financial services has introduced new prompt and accurate data, potentially enhancing the analysis of economic and financial system stability.
- Advanced methodologies like AI and machine learning are essential for analyzing high-frequency transactional data (Philippon, 2016).

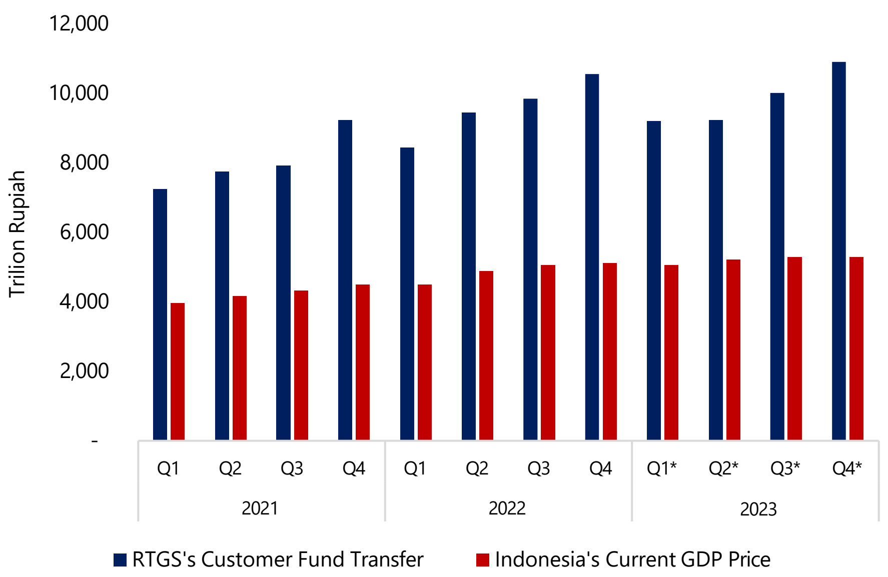
RESEARCH QUESTIONS

1. How can central banks further enrich their payment systems data to support policy making for financial stability?
2. How can the resulting indicator from enriched payment systems data complement existing indicators on financial intermediations and financial sector soundness?

RTGS AS PROXY OF ECONOMIC ACTIVITY



The large size of RTGS's customer fund transfers highlights its significant role in Indonesia's economy.



LITERATURE REVIEW

Role of Central Banks in Financial Stability

- Central banks are uniquely positioned to contribute to maintaining financial stability, due to their role as monetary authority and financial sector regulator.
- Systemic risk in the financial system manifests in two principal forms: time series risk (procyclicality) and cross section risk (interconnection and common exposure)
- The flexibility and real-time nature of big data enable the extraction of timely economic signals, application of new statistical methodologies, and enhancement of financial stability assessments (Tissot, 2019).

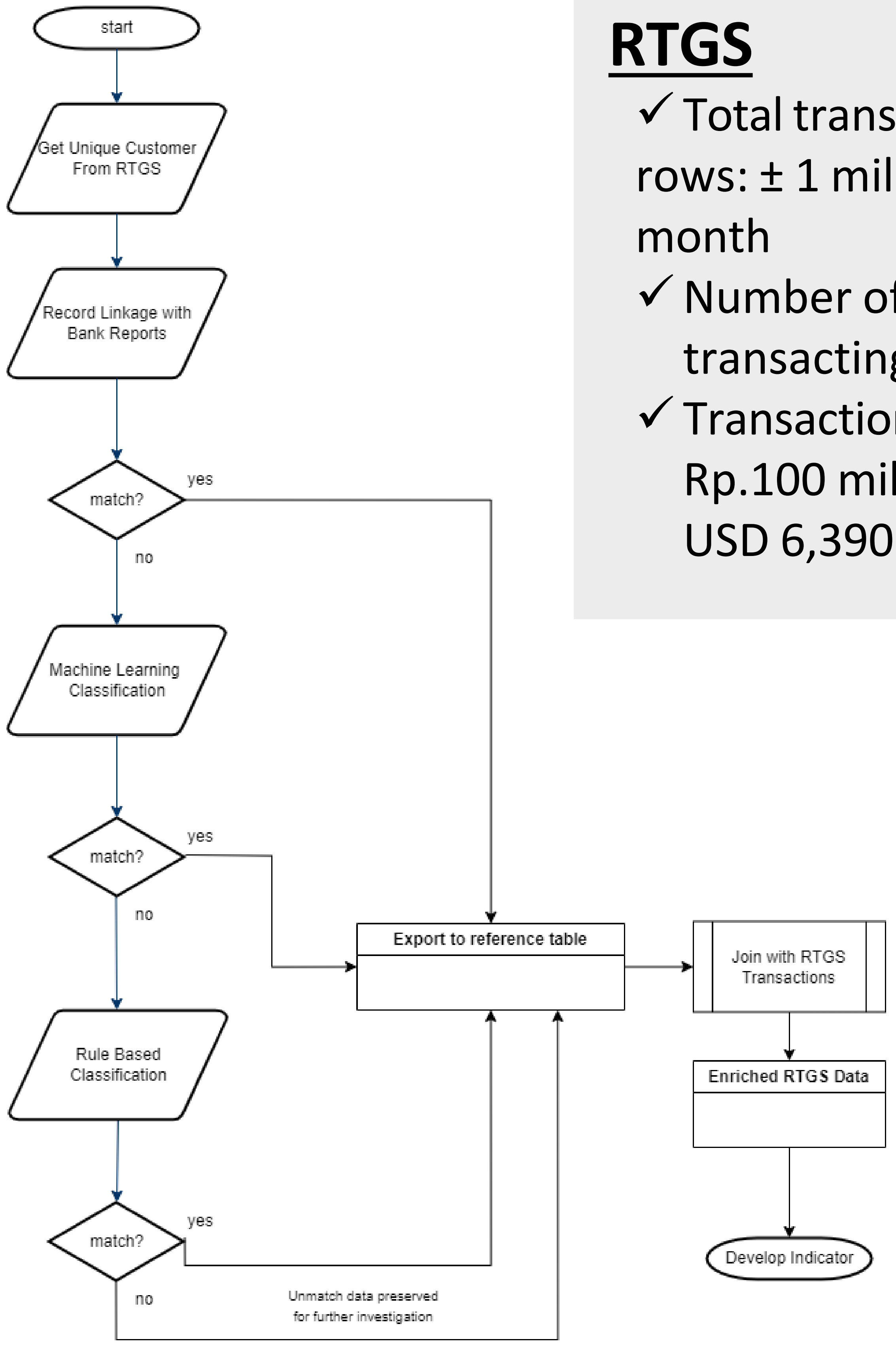
Application of AI and Big Data in Financial Stability Analysis

- AI enhances decision-making processes through advanced predictive analytics, enabling central banks to forecast economic trends more accurately and make better-informed policy decisions (Ghandour, 2021).
- Sulistiawati et al. (2022) utilized big data analytics on payment system data to identify publicly listed large non-financial corporations to monitor their transactions.

STATE OF RTGS DATA IN INDONESIA

Legacy systems like BI-RTGS in Indonesia were **originally developed** to **ensure uninterrupted payment system** activities rather than to meet the **needs of economic analysis**. Therefore, **further process** are needed to **enrich** the granular information in RTGS Transaction.

DATA SOURCES & METHODOLOGY



RTGS

- ✓ Total transaction rows: ± 1 million rows per month
- ✓ Number of banks transacting: 117 banks
- ✓ Transaction amount ≥ Rp.100 million or ≥ USD 6,390

BANK REPORTS

- ✓ 10 millions Accounts
- ✓ 50 Banks
- ✓ Customer Information*
  - ✓ Entity Classification
  - ✓ Gender
  - ✓ Domicile
  - ✓ Birth Year
  - ✓ CASA
  - ✓ etc

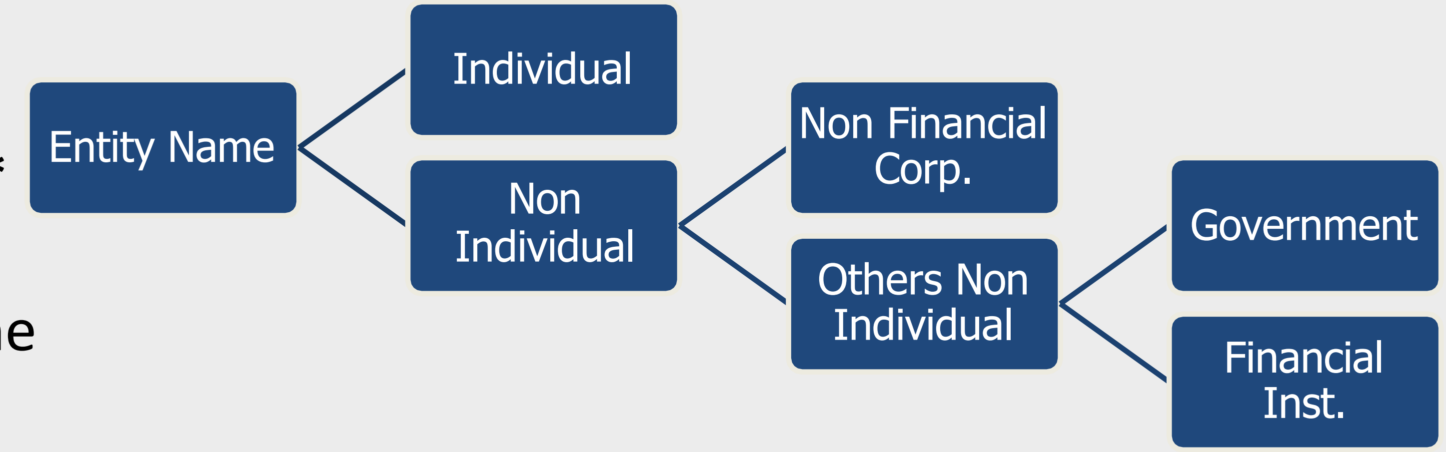
Rule Based

- ✓ Remainders unmatch entity then processed with rulebased classification based on majority countvectorizer from each entity classification

MACHINE LEARNING

Spark MLlib

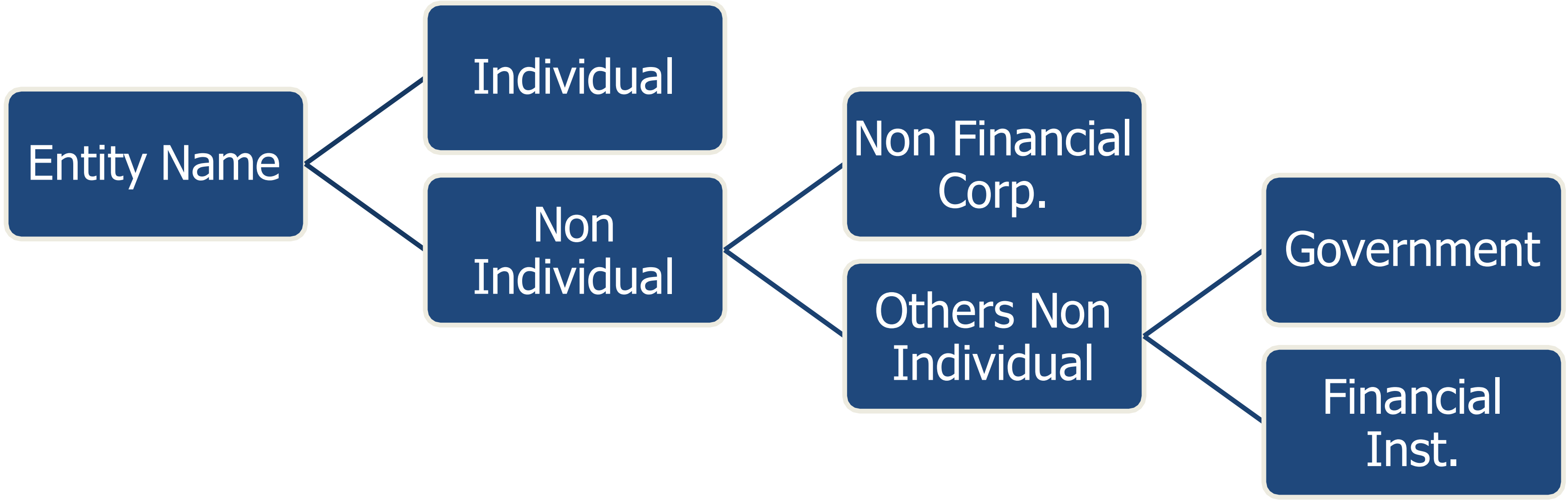
- ✓ Word Tokenization
- ✓ Count Vectorizer
- ✓ Supervised Classification**
  - ✓ **Logistic Regression**
  - ✓ Support Vector Machine
  - ✓ Naïve Bayes



Notes:
***) Only consistent classification between banks are used to train the model**
****) We build 3 models, to classify entity name**

MODEL EVALUATION

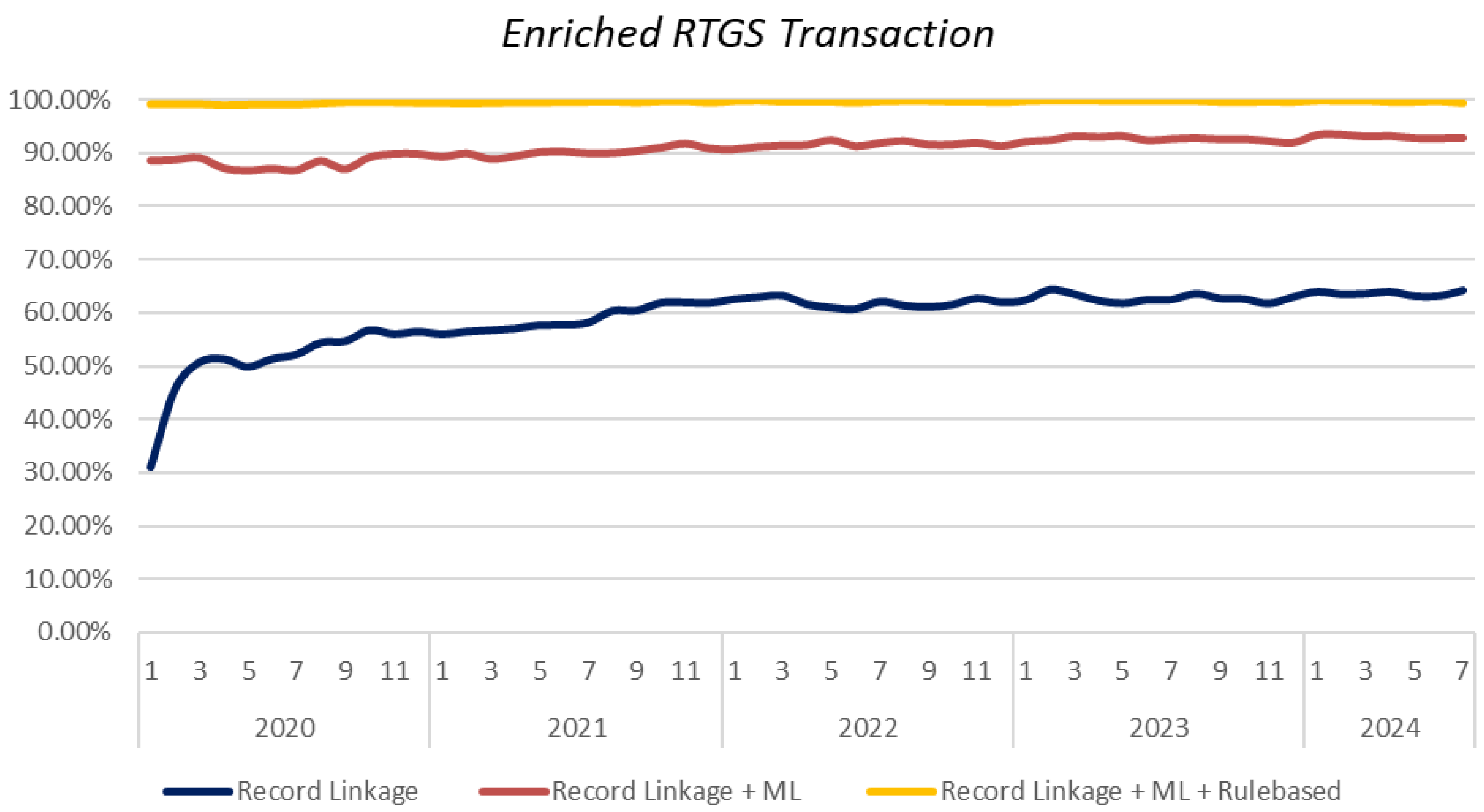
MODEL EVALUATION



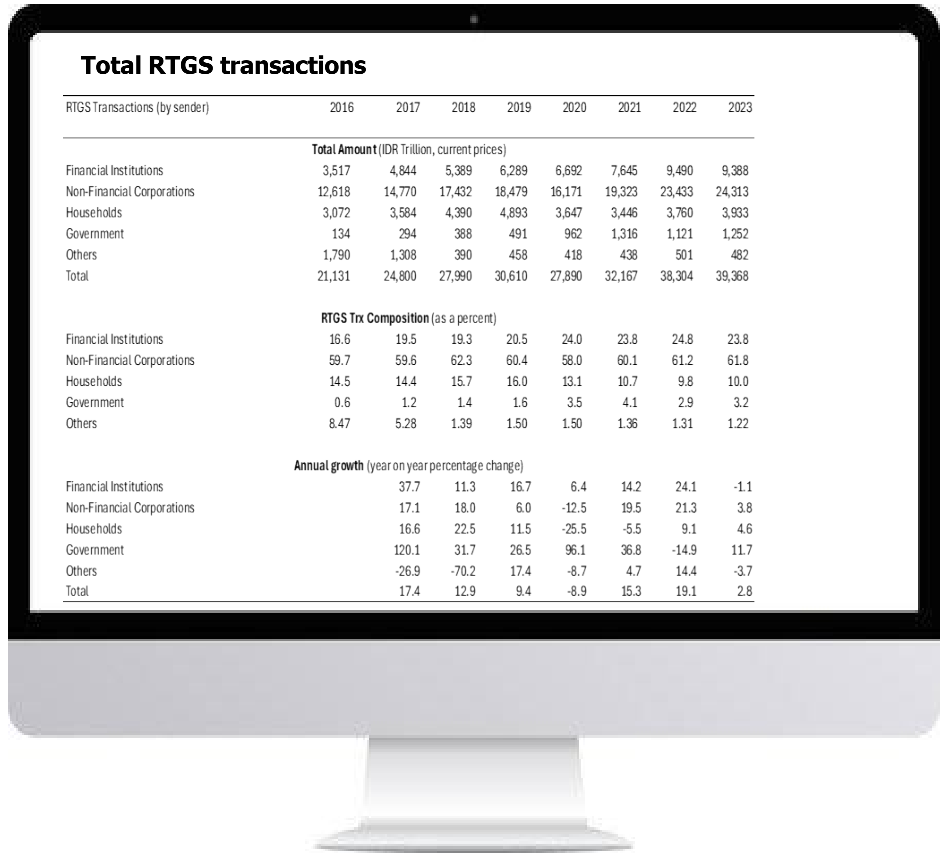
Algorithm	Klasifikasi	Train*				Test*			
		F1-Score	Precision	Recall	AUC	F1-Score	Precision	Recall	AUC
Logistic Regression	Perorangan	0.97	0.97	0.97	0.99	0.97	0.97	0.97	0.97
	Perusahaan Non Finansial	0.72	0.82	0.76	0.83	0.66	0.79	0.72	0.72
	Institusi Keuangan / Pemerintah	0.98	0.98	0.98	0.99	0.96	0.96	0.96	0.98
SVM	Perorangan	0.99	0.98	0.98	0.98	0.98	0.96	0.96	0.96
	Perusahaan Non Finansial	0.83	0.77	0.66	0.62	0.77	0.76	0.65	0.61
	Institusi Keuangan / Pemerintah	0.97	0.92	0.92	0.02	0.97	0.91	0.91	0.91
Naïve Bayes	Perorangan	0.57	0.98	0.98	0.98	0.59	0.96	0.96	0.96
	Perusahaan Non Finansial	0.48	0.76	0.71	0.68	0.47	0.76	0.7	0.68
	Institusi Keuangan / Pemerintah	0.69	0.89	0.89	0.89	069	0.89	0.88	0.88

HOW ML CLOSING THE GAP FOR UNMATCHED ACCOUNTS

Method	Matching Percentage
Record Linkage	64%
Record Linkage + ML	92%
Record Linkage + ML + Rulebased	99%

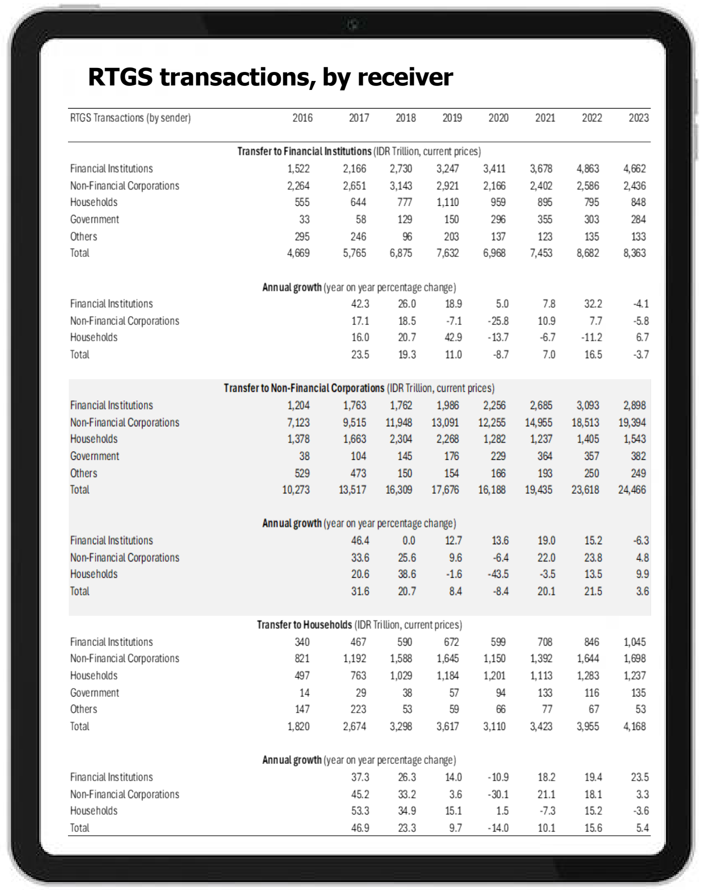


Output #1. Enriched Statistics of RTGS transactions



Total RTGS transactions

RTGS Transactions (by sender)	2016	2017	2018	2019	2020	2021	2022	2023
Total Amount (IDR Trillion, current prices)								
Financial Institutions	3,517	4,844	5,389	6,289	6,692	7,645	9,490	9,388
Non-Financial Corporations	12,618	14,770	17,432	18,479	16,171	19,323	23,433	24,313
Households	3,072	3,584	4,390	4,893	3,647	3,446	3,760	3,933
Government	134	294	388	491	962	1,316	1,121	1,252
Others	1,790	1,308	390	458	418	438	501	482
Total	21,131	24,800	27,990	30,610	27,890	32,167	38,304	39,368
RTGS Trx Composition (as a percent)								
Financial Institutions	16.6	19.5	19.3	20.5	24.0	23.8	24.8	23.8
Non-Financial Corporations	59.7	59.6	62.3	60.4	58.0	60.1	61.2	61.8
Households	14.5	14.4	15.7	16.0	13.1	10.7	9.8	10.0
Government	0.6	1.2	1.4	1.6	3.5	4.1	2.9	3.2
Others	8.47	5.28	1.39	1.50	1.50	1.36	1.31	1.22
Annual growth (year on year percentage change)								
Financial Institutions		37.7	11.3	16.7	6.4	14.2	24.1	-1.1
Non-Financial Corporations		17.1	18.0	6.0	-12.5	19.5	21.3	3.8
Households		16.6	22.5	11.5	-25.5	-5.5	9.1	4.6
Government		120.1	31.7	26.5	96.1	36.8	-14.9	11.7
Others		-26.9	-70.2	17.4	-8.7	4.7	14.4	-3.7
Total		17.4	12.9	9.4	-8.9	15.3	19.1	2.8



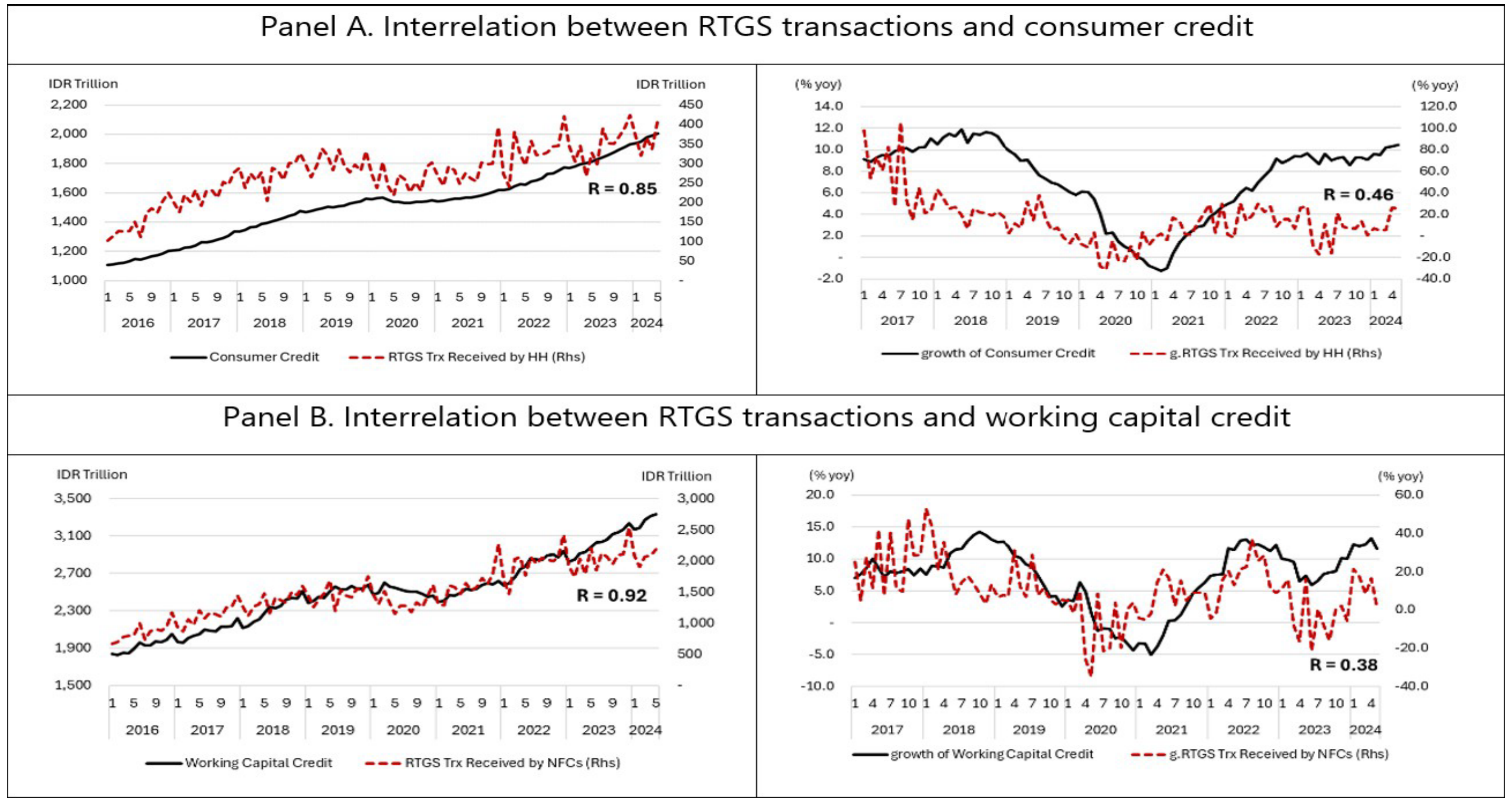
RTGS transactions, by receiver

RTGS Transactions (by sender)	2016	2017	2018	2019	2020	2021	2022	2023
Transfer to Financial Institutions (IDR Trillion, current prices)								
Financial Institutions	1,522	2,166	2,730	3,247	3,411	3,678	4,863	4,662
Non-Financial Corporations	2,264	2,651	3,143	2,921	2,166	2,402	2,586	2,436
Households	555	644	777	1,110	959	895	795	848
Government	33	58	129	150	296	355	303	284
Others	295	246	96	203	137	123	135	133
Total	4,669	5,765	6,875	7,632	6,968	7,453	8,682	8,363
Annual growth (year on year percentage change)								
Financial Institutions		42.3	26.0	18.9	5.0	7.8	32.2	-4.1
Non-Financial Corporations		17.1	18.5	-7.1	-25.8	10.9	7.7	-5.8
Households		16.0	20.7	42.9	-13.7	-6.7	-11.2	6.7
Total		23.5	19.3	11.0	-8.7	7.0	16.5	-3.7
Transfer to Non-Financial Corporations (IDR Trillion, current prices)								
Financial Institutions	1,204	1,763	1,762	1,986	2,256	2,685	3,093	2,898
Non-Financial Corporations	7,123	9,515	11,948	13,091	12,255	14,955	18,513	19,394
Households	1,378	1,663	2,304	2,268	1,282	1,237	1,405	1,543
Government	38	104	145	176	229	364	357	382
Others	529	473	150	154	166	193	250	249
Total	10,273	13,517	16,309	17,676	16,188	19,435	23,618	24,466
Annual growth (year on year percentage change)								
Financial Institutions		46.4	0.0	12.7	13.6	19.0	15.2	-6.3
Non-Financial Corporations		33.6	25.6	9.6	-6.4	22.0	23.8	4.8
Households		20.6	38.6	-1.6	-43.5	-3.5	13.5	9.9
Total		31.6	20.7	8.4	-8.4	20.1	21.5	3.6
Transfer to Households (IDR Trillion, current prices)								
Financial Institutions	340	467	590	672	599	708	846	1,045
Non-Financial Corporations	821	1,192	1,588	1,645	1,150	1,392	1,644	1,698
Households	497	763	1,029	1,184	1,201	1,113	1,283	1,237
Government	14	29	38	57	94	133	116	135
Others	147	223	53	59	66	77	67	53
Total	1,820	2,674	3,298	3,617	3,110	3,423	3,955	4,168
Annual growth (year on year percentage change)								
Financial Institutions		37.3	26.3	14.0	-10.9	18.2	19.4	23.5
Non-Financial Corporations		45.2	33.2	3.6	-30.1	21.1	18.1	3.3
Households		53.3	34.9	15.1	1.5	-7.3	15.2	-3.6
Total		46.9	23.3	9.7	-14.0	10.1	15.6	5.4

Usage:

- Generate statistics on RTGS transactions, which can be monitored on a monthly basis, as proxy of economic activities.
- Illustrate the RTGS flows undertaken by various economic agents.
- reveal the flow of money and liquidity across different sectors.

Output #2. Analysis on RTGS data as proxy for Intermediation and Soundness



Analysis:

- Examines the interrelatedness of RTGS transactions with intermediation and financial soundness indicators
- RTGS transactions exhibit a high correlation with both consumer and working capital credit
- RTGS transactions are more effective in reflecting the soundness of NFCs compared to households

KEY FINDINGS & FUTURE WORKS

KEY FINDINGS:

1. The use of big data analytics and machine learning methodology can enrich payment system data, categorizing entity into individuals, financial institutions, non-financial corporations, and government, providing potentials for economic and financial analytics.

2. RTGS transactions hold significant potential as a proxy of credit demand. In the context of financial sector soundness, RTGS transactions provide strong correlation with NFCs soundness indicator. However, we found limited potential of RTGS transactions data to indicate household soundness.

FUTURE WORK:

- Expanding data sources to include more payment systems
- Exploring interconnectedness within the payment system network
- Exploring more financial or economics variable to validate other enriched information such as regional analysis, entity interconnection, spending habit by demographics, etc

DIDD

DEPARTEMEN INOVASI & DIGITALISASI DATA

Data at your fingertips.

BANK INDONESIA

BANK SENTRAL REPUBLIK INDONESIA

THANK YOU

23 August 2024

Renardi Ardiya Bimantoro, Charvin Lim, Mohammad Khoyrul Hidayat

Department of Data Innovation and Digitalization, Department of Macroprudential Policy, Department of Statistics – Bank Indonesia

Email: renardi.ardiya@bi.go.id; charvin_lim@bi.go.id; moh.khoyrul@bi.go.id