Determinants of Financial Inclusion in India

Ayush Pandey[1], Leesha Mahajan[2], Vivek Kumar Gupta[3], Abhishek Singh[4]
[1]*Gokhale Institute of Politics and Economics*
[2,3,4]*Universal AI University*

***Abstract*—This study analyses the determinants of financial inclusion in India over the period 1991–2018. Financial inclusion, recognized by the United Nations as a key enabler of the Sustainable Development Goals (SDGs), has gained increasing policy attention in India, particularly since 2004. While earlier efforts, such as bank nationalization and the establishment of regional rural banks, sought to extend financial services to marginalized communities, their impact remained constrained. This research constructs a composite Financial Inclusion Index (FII) using Principal Component Analysis (PCA) to quantify the level of financial inclusion. Subsequently, employing the Johansen cointegration methodology, the study investigates the long-run equilibrium relationship between the constructed FII and a set of macroeconomic and socio-demographic factors: GDP growth rate, internet penetration rate, inflation, age dependency ratio, and gross secondary school enrolment ratio. The analysis aims to identify the key drivers of financial inclusion in India, contributing to a deeper understanding of its dynamics and informing policy interventions aimed at promoting broader access to financial services. Furthermore, the study explores the evolution of financial inclusion policies in India, contextualizing the quantitative analysis within the broader policy landscape.**

***Index Terms*—financial inclusion, principal component analysis, Johansen cointegration, index, India, SDGs**

I. INTRODUCTION

The provision of financial services to all segments of society, particularly the most vulnerable, is a critical imperative for inclusive development in emerging economies. India, despite experiencing substantial economic growth in recent decades, continues to grapple with the persistent challenges of poverty and income inequality. Financial inclusion has emerged as a crucial strategy for addressing these interconnected issues. As defined by the Rangarajan Committee (2008), financial inclusion is "the process of ensuring access to financial services, including timely and adequate credit, at an affordable cost to vulnerable groups such as weaker sections and low-income groups." Within the context of India's pursuit of sustainable development goals, the broadening of financial participation across all segments of the population is essential. However, limited financial literacy, particularly within rural communities, poses a significant impediment to accessing formal credit channels and consequently hinders broader economic progress (Singh, 2014). While advancements in financial technology have demonstrably transformed urban financial practices, the reach of these innovations has yet to fully extend to the traditionally underserved rural populations.

Existing scholarship suggests a positive correlation between financial inclusion and several desirable macroeconomic outcomes, including stability, sustainable development, employment generation, improved living standards, and inclusive economic growth (Kamath, 2007). A critical distinction must be drawn between voluntary and involuntary financial exclusion, with policy interventions appropriately focused on mitigating the latter. A necessary precursor to effective policy design is the identification of the key determinants of financial inclusion, followed by targeted interventions addressing the most influential factors. Furthermore, the development of a robust and reliable financial inclusion indicator is essential to facilitate both temporal and regional comparisons within countries. In developing regions across Asia, Latin America, and Africa, involuntary exclusion from the formal financial system limits opportunities for savings, investment, and wealth accumulation, thereby underscoring the importance of financial inclusion as a catalyst for stimulating broader economic activity. While banks remain the dominant financial institutions in developing economies, studies typically assess financial inclusion across three primary dimensions: penetration (the availability of banking facilities), accessibility (the ease of access to these facilities), and usage (the actual utilization of available financial services). Limited access due to geographical constraints, coupled with low levels of financial literacy and a lack of acceptable collateral, constitute significant barriers for marginalized communities seeking to participate in the formal financial system.

Financial development is intrinsically linked to broader economic development. While financial deepening, reflecting increased access to capital and credit, has shown improvement in recent years, the breadth and depth of formal finance have not expanded commensurately (Mohan, 2006). Consequently, financial inclusion has become a prominent policy priority in India, particularly since 2005. Although not explicitly articulated within the United Nations' Sustainable Development Goals (SDGs), financial inclusion serves as a crucial catalyst for the achievement of several of these goals (Klapper, El-zoghbi, and Hess, 2016). It is widely recognized as being vital for achieving inclusive and sustainable development outcomes (Thorat, 2007). In India, rural financial service penetration remains comparatively low, often attributable to supply-side constraints such as limited branch networks and a lack of financial products tailored to the specific needs of rural populations, as well as demand-side factors including low incomes and limited financial literacy (Singh, 2014).

Financial exclusion, frequently preceded by social exclusion, initially focused primarily on geographical access to banking facilities (Leyshon and Thrift, 1995). However, the concept has evolved to encompass all individuals with limited or no access to, and usage of, appropriate financial services and products (Kempson and Whyley, 1999; European Commission, 2008). Financial exclusion disproportionately affects vulnerable and marginalized groups, including single parents, the long-term unemployed, and certain ethnic minorities (Kodan & Chhikara, 2013). Broad access to financial services is now widely recognized as essential for both economic and social development (Beck and Torre, 2006; Beck et al., 2000, 2004; Honohan, 2004; Peachey and Roe, 2004). Financial inclusion/exclusion, therefore, occupies a central position in understanding the underlying factors that limit access to the formal financial system (RBI, 2008; Kodan & Chhikara, 2013). Inclusive growth and financial inclusion are no longer considered optional policy choices but rather essential for long-term social and economic stability (Chakrabarty, 2013). Marginalized populations are often excluded from formal finance, lacking access to monetized earnings and the financial services that facilitate economic participation (Bagli & Dutta, 2012). Global data reveals a significant disparity in financial service usage between high-income and developing countries, with account ownership being substantially more prevalent in the former (Demirguc-Kunt and Klapper, 2012). The Global Findex database, initiated in 2011, provides a valuable source of data on the status of financial inclusion across a broad range of countries (Global Findex Methodology, 2011, 2014).

II. RESEARCH QUESTIONS:

- What are the primary factors influencing financial inclusion in India?
- How can financial inclusion be effectively defined and measured?
- What methodologies can be employed to construct a comprehensive financial inclusion index?

A. Research Objectives:

- Construct a unified index to measure financial inclusion in India.
- To empirically analyse the determinants of financial inclusion in India, specifically examining the long-run equilibrium relationship between the constructed index and a set of relevant macroeconomic and socio-demographic variables.

B. Research Gap:

While a general consensus exists regarding the broad concept of financial inclusion, variations persist in its precise definition and measurement across different studies (Kodan & Chhikara, 2013). Prior research frequently suffers from a lack of consistent and reliable macro-level data and standardized measurement approaches. Although existing studies have explored cross-country analyses of the determinants of financial inclusion and its impact on poverty reduction and income inequality, country-specific studies, particularly those focusing on developing economies like India, remain relatively limited. This paper aims to address this important research gap by constructing a novel financial inclusion index specifically for India, utilizing relevant indicators that capture the unique characteristics of financial inclusion within the Indian context, and subsequently analysing the key determinants of financial inclusion using this newly developed index.

III. LITERATURE REVIEW

Financial inclusion, a multifaceted concept, has become a critical area of focus in development economics, recognized as a crucial driver of socioeconomic progress. This review synthesizes existing literature, exploring its determinants, measurement methodologies, and multifaceted impacts, with a particular emphasis on developing economies. A comprehensive understanding of financial inclusion necessitates moving beyond the simplistic notion of mere access to financial services. It encompasses the actual usage of these services and, crucially, the quality-of-service delivery (Tiriki & Faye, 2013; G-20; Alliance for Financial Inclusion, 2013; Stein et al., 2011). Financial inclusion initiatives aim to provide accessible and affordable financial products and services, particularly to vulnerable and marginalized groups traditionally excluded from the formal financial sector. This exclusion can perpetuate cycles of poverty and limit economic opportunities.

A significant body of research has concentrated on developing robust methodologies for measuring financial inclusion, recognizing its multi-dimensional nature. Composite indices have become a popular tool for capturing this complexity. Sarma (2008) pioneered a widely used method based on the normalized inverse of Euclidean distance, offering a relatively simple yet effective approach. Amidžić, Massara, and Mialou (2014) employed a weighted geometric mean, explicitly incorporating dimensions of outreach (geographic and demographic penetration), usage (deposit and lending activity), and quality (cost, disclosure, and dispute resolution mechanisms). This approach allows for a more nuanced assessment of financial inclusion. Other studies have utilized parametric techniques like Principal Component Analysis (PCA) and factor analysis (Cámara & Tuesta, 2015; Amidi et al., 2014), which statistically derive weights based on the underlying data. However, these methods have limitations. While statistically rigorous, they can potentially overlook crucial dimensions by making simplifying assumptions (Steiger, 1979). Conversely, non-parametric approaches (Chakravarty & Pal, 2013; Sarma, 2012), while simpler to implement, can be susceptible to subjective weight assignment by researchers (Lockwood, 2004). The choice of methodology often depends on the specific context and data availability. Lenka, Barik, and Parida (2022), for instance, employed PCA to construct a financial inclusion index specifically for Indian states, highlighting the regional variations in financial access.

A substantial body of literature has explored the determinants of financial inclusion, identifying a range of macroeconomic, institutional, demographic, and socioeconomic factors. At the macroeconomic level, per capita GDP and internet penetration consistently emerge as strong positive drivers (Omar & Inaba, 2020; Park & Mercado, 2015). Higher levels of economic development and increased connectivity facilitate greater access to and usage of financial services. Conversely, inflation, age dependency ratio, and income inequality are often negatively correlated with financial inclusion (Omar & Inaba, 2020). High inflation can erode purchasing power and discourage savings, while a high dependency ratio can strain household budgets. Income inequality can further marginalize certain segments of the population, limiting their access to financial services. Demographic factors, such as population size and age structure, also play a role (Park & Mercado, 2015). Institutional quality, particularly the rule of law and contract enforcement, is crucial for fostering trust and confidence in the financial system, thereby promoting financial inclusion (Omar & Inaba, 2020; Park & Mercado, 2015; Allen et al., 2016). Other important determinants include bank-specific characteristics like bank size, deposit rates, and the cost-to-income ratio (Uddin et al., 2017). Socioeconomic factors, such as literacy rates, education levels, and financial literacy, are also significant (Uddin et al., 2017; Bhanot et al., 2012; Kohli, 2013; Gashaw & Gebe, 2017). Individuals with higher levels of education and financial knowledge are more likely to understand and utilize financial products and services. Infrastructural factors, particularly the density of bank branches and ATMs, are critical for ensuring accessibility, especially in geographically dispersed and rural areas (Gupte et al., 2012; Kumar, 2013; Chakravarty & Pal, 2013). Rafiq and Adewale (2019), in their study of India, highlighted the positive influence of industrialization (proxied by the number of factories) and outstanding credit on financial inclusion, suggesting a link between economic activity and access to finance. Gashaw and Gebe (2017) explored the determinants of financial inclusion in Ethiopia, finding that urban residence, religious affiliation, financial awareness, and marital status all influence account ownership and usage.

The relationship between financial inclusion and socioeconomic development is complex and has been extensively investigated, with varying conclusions. While a substantial body of research supports the positive association between financial development and economic growth (Bagehot, 1873; Schumpeter, 1934; Goldsmith, 1959; Gurley & Shaw, 1955; Levine, 1997), other studies focus on the distributional implications of financial inclusion. Pal and Pal (2012) and Singh and Naik (2017) highlighted the persistent challenge of financial exclusion across income groups and regions in India, emphasizing the need for targeted interventions. The role of microfinance in promoting financial inclusion, particularly among low-income households and in remote areas, has also been studied extensively (Mendoza, 2009). Furthermore, research suggests that financial inclusion can contribute to human development by empowering individuals, improving access to education and healthcare, and reducing vulnerability (Kohli, 2013). Cámara and Tuesta (2015) found that per capita GDP, education level, financial system performance, and financial strength are key determinants of financial inclusion. Yadav and Sharma (2016) developed a multi-dimensional index to rank Indian states and union territories, identifying literacy, infrastructure development, and agricultural factors as influential.

Despite significant progress in expanding financial inclusion, several challenges remain. Collins et al. (2009), through their financial diaries study, highlighted the continued reliance on informal financial instruments by low-income households, suggesting persistent barriers to accessing formal financial services. These barriers can include lack of documentation, high transaction costs, and a lack of trust in formal institutions. Beck, Demirguc-Kunt, and Levine (2007) observed variations in account ownership and balances across income levels, emphasizing the need for targeted interventions to address the specific needs of different segments of the population. Fungácová and Weill (2015) examined financial inclusion in China, identifying wealth, education, gender, age, and trust as important influencing factors. They also highlighted the different concerns of various demographic groups, with poorer individuals often citing lack of funds as a major constraint, while more educated individuals are more concerned about cost and trust in the banking system. Future research should focus on gaining a deeper understanding of the specific needs and constraints of different population segments, developing tailored financial products and services, and leveraging technology to enhance accessibility and affordability. Further investigation into the complex interplay between financial inclusion, poverty reduction, income inequality, and other development outcomes is also warranted. Finally, exploring the transformative potential of digital financial services and fintech innovations for advancing financial inclusion is a crucial area for future research.

IV. DATA AND METHODOLOGY

This section details the data sources and methodological approach employed in this study. The research has two primary objectives: first, to construct a composite Financial Inclusion Index (FII) for India, and second, to analyse the influence of specific macroeconomic and socio-demographic factors on financial inclusion over the period 1991–2018.

To achieve the first objective, Principal Component Analysis (PCA) is employed. PCA is a statistical technique used to reduce the dimensionality of a dataset by creating new, uncorrelated variables (principal components) that capture the maximum variance in the original data. In this study, PCA is used to aggregate five key indicators into a single FII. These indicators, sourced from the Reserve Bank of India's (RBI) Basic Statistical Returns (BSR) for Scheduled Commercial Banks (SCBs), are: the number of bank employees, the number of bank branches, the total amount of credit outstanding, the total amount of deposits, and the number of savings bank accounts. These variables are selected to represent different facets of financial inclusion, including access points (branches, employees), credit availability, deposit mobilization, and account penetration. The data spans the period 1991–2018.

The second objective, analysing the determinants of financial inclusion, utilizes five independent variables: GDP growth rate, inflation rate (measured by the Consumer Price Index, CPI), gross secondary school enrolment ratio, age dependency ratio, and the number of internet users. Data for these variables are sourced from the World Development Indicators (WDI) database, covering the same period, 1991–2018. GDP growth rate is used as a proxy for economic development, CPI reflects macroeconomic stability, the gross secondary school enrolment ratio captures human capital development, the age dependency ratio represents demographic structure, and the number of internet users' proxies for technological access and information dissemination. These variables are chosen based on existing literature and their hypothesized relationship with financial inclusion. The subsequent analysis will employ econometric techniques to examine the long-run relationship between these determinants and the constructed FII.

Table 1: Objective and Data source of variables used in the study

Variables	Objective	Data Source
Number of saving bank accounts	To construct FII	Basic Statistical Returns (BSR), RBI
Number of bank employees	To construct FII	Basic Statistical Returns (BSR), RBI
Number of bank branches	To construct FII	Basic Statistical Returns (BSR), RBI
Amount of deposit	To construct FII	Basic Statistical Returns (BSR), RBI
Amount of credit	To construct FII	Basic Statistical Returns (BSR), RBI
Age dependency ratio	To analyse the impact on financial inclusion	World Development Indicator database
Number of internet users	To analyse the impact on financial inclusion	World Development Indicator database
Inflation rate (CPI)	To analyse the impact on financial inclusion	World Development Indicator database
GDP per capita	To analyse the impact on financial inclusion	World Development Indicator database
Gross Secondary School Enrolment ratio	To analyse the impact on financial inclusion	World Development Indicator database

V. DESCRIPTION OF THE VARIABLES

1. Variables used to create Financial Inclusion Index

a) Number of Savings Bank Accounts: Access to basic banking services is a fundamental dimension of financial inclusion, enabling individuals to participate more fully in the formal financial system. The number of savings bank accounts serves as a crucial indicator of this dimension, reflecting the extent to which individuals have access to and utilize formal financial services for saving and transaction purposes. Individuals lacking savings bank accounts often face significant barriers to saving, investing, and accessing credit, thereby limiting their economic opportunities and financial well-being. Savings accounts not only facilitate secure and convenient financial transactions but also play a critical role in establishing a financial history, which can be essential for accessing other financial products and services, such as loans, insurance, and investment options. Furthermore, savings accounts promote financial literacy and empower individuals to manage their finances more effectively, contributing to greater financial resilience. Including the number of savings bank accounts in the FII provides a valuable measure of access to essential financial services and helps identify potential gaps in financial inclusion across different segments of the population. It also reflects the overall reach and penetration of the formal banking sector.

b) Number of Bank Branches: The number of bank branches is a significant indicator of the penetration or accessibility dimension of financial inclusion. It reflects the physical presence and geographic reach of formal financial institutions within a given area or region. The availability of bank branches can significantly improve access to financial services, particularly for individuals and communities with limited digital literacy or restricted access to internet banking facilities. Physical branches often serve as vital points of contact for customers, providing personalized services, building trust, and fostering stronger relationships with financial institutions. Moreover, bank branches play an important role in promoting financial literacy by offering financial education, guidance, and support to customers. They also act as crucial distribution channels for a wide array of financial products and services, including loans, insurance, investment products, and other financial instruments. Including the number of bank branches in the FII offers insights into the extent of access to formal financial institutions and services across different geographic areas and helps identify regions where physical access may be limited, highlighting the need for targeted interventions to enhance financial inclusion. It also reflects the investment in physical infrastructure by the banking sector.

c) Number of Bank Employees: The number of bank employees can be considered an indicator of the capacity and outreach of the banking sector. A larger workforce within the banking industry may suggest a greater ability to provide personalized services to customers, potentially leading to increased customer satisfaction, trust, and confidence in formal financial institutions. Banks with a larger staff may also be better equipped to serve customers with more complex financial needs, such as small business owners or individuals with diverse financial portfolios, who may require specialized advice and assistance. Furthermore, the number of bank employees can reflect employment trends and growth within the financial sector, potentially indicating expansion and increased economic activity related to financial services. Including this variable in the FII provides insights into the service capacity and human capital of formal financial institutions and offers a broader perspective on the economic and social implications of financial inclusion. It can also reflect the ability of the banking sector to absorb and serve a larger population.

d) Amount of Deposits: The total amount of deposits mobilized by the formal financial system is a crucial indicator of financial inclusion, reflecting the extent to which individuals and businesses utilize these institutions for saving, wealth accumulation, and financial planning. Higher deposit volumes can signal increased financial security, greater trust in formal financial institutions, and a greater propensity to save. Access to deposit accounts and the ability to make regular deposits can significantly improve financial stability, particularly for low-income individuals and households who may be more vulnerable to unforeseen financial expenses or economic shocks. The number of deposits can also provide valuable information about the overall level of economic activity and financial deepening within a given area or region. Higher levels of deposits may indicate increased economic activity, investment, and confidence in the financial system. Including the number of deposits in the FII provides valuable insights into the extent of savings mobilization through formal channels and its broader economic impact. It reflects the degree to which individuals are actively participating in the formal financial system.

e) Amount of Credit: The total amount of credit extended by formal financial institutions is a critical measure of financial inclusion, reflecting access to formal credit markets for individuals, businesses, and other economic actors. Access to credit can play a vital role in facilitating economic growth and development by enabling individuals to invest in education, housing, or entrepreneurial ventures. It can also help individuals and businesses manage cash flow, smooth consumption patterns, and address unforeseen expenses or financial emergencies. Furthermore, access to credit can assist individuals in building credit histories, which can be essential for accessing future financial services and improving their overall financial standing. The amount of credit disbursed can also provide insights into the level of economic activity and investment within a region. Higher levels of credit may be associated with increased economic growth, business expansion, and job creation. Including the amount of credit in the FII helps assess the extent of access to formal credit channels and its broader impact on economic development and financial inclusion. It reflects the ability of the financial system to provide resources for productive investments and consumption smoothing.

2. Description of the variables used for analysing the determinants of financial inclusion

a) GDP per capita (constant LCU): Gross Domestic Product (GDP) per capita, calculated as GDP divided by midyear population, reflects the average economic output per person in a country. GDP at purchaser's prices is the sum of gross value added by all resident producers, plus any product taxes and minus any subsidies. It is measured without accounting for natural resource depletion or asset depreciation and is expressed in local currency units (LCU). Theoretically, GDP per capita is expected to have a positive relationship with financial inclusion. Higher GDP per capita generally correlates with higher individual incomes, providing individuals with greater capacity to engage with financial services. Furthermore, in economies with higher GDP per capita, financial institutions tend to offer more diverse financial products and services and are incentivized to invest in infrastructure development, including expanding ATM and branch networks and adopting digital banking technologies. Conversely, in countries with lower GDP per capita, financial institutions may be less inclined to invest in infrastructure and new product development due to higher costs and lower potential returns, potentially creating barriers to financial inclusion. While GDP per capita is not the sole determinant of financial inclusion, it plays a significant role in shaping the incentives for financial institutions to expand access to financial services.

b) Age Dependency Ratio: The age dependency ratio measures the proportion of dependents (individuals aged 0-14 and 65 and older) to the working-age population (15-64). It is expressed as the number of dependents per 100 working-age individuals. This ratio indicates the burden placed on the working-age population to support those who are economically dependent. While it reflects demographic structure, it does not directly measure economic dependence, as some working-age individuals may be unemployed, and some dependents may be economically active. Theoretically, a high age dependency ratio can negatively influence financial inclusion. A larger dependent population can strain the resources of the working-age population, potentially reducing their capacity to save, invest, and engage with financial services. Conversely, a lower age dependency ratio may create more opportunities for financial inclusion, as a smaller dependent population may free up resources for saving and investment. However, the impact of the age dependency ratio on financial inclusion can be influenced by other factors, such as the level of economic development, the quality of financial infrastructure, and the availability of financial services. These factors can either amplify or mitigate the relationship between the age dependency ratio and financial inclusion.

c) Gross Secondary School Enrolment Ratio: The gross secondary school enrolment ratio represents the total enrolment in secondary education, regardless of age, as a percentage of the population in the age group that officially corresponds to that level of education. Secondary education aims to build a foundation for lifelong learning and human development by providing more specialized training and instruction. Theoretically, the gross secondary school enrolment ratio is expected to have a positive impact on financial inclusion. Education, particularly secondary education, plays a vital role in enhancing financial literacy and numeracy skills, which are essential for making informed financial decisions. Educated individuals are better equipped to understand different financial products and services, evaluate their suitability, and manage their finances effectively. Furthermore, secondary education can improve employment opportunities and income levels, increasing individuals' capacity to save, invest, and access credit. It can also foster a culture of saving and encourage the use of financial services. Moreover, secondary education can facilitate access to financial education programs and financial services offered by educational institutions, governments, or NGOs. Therefore, higher levels of gross secondary school enrolment can contribute to increased financial inclusion and improved financial well-being.

d) Inflation: Inflation refers to a general increase in the price level of goods and services over time. The Consumer Price Index (CPI) measures inflation by tracking changes in the cost of a basket of goods and services consumed by a typical household. High inflation can negatively affect financial inclusion. When inflation is high, individuals tend to allocate a larger proportion of their income to essential expenses like food, housing, and healthcare, reducing their disposable income and their capacity to save. High inflation also erodes the purchasing power of savings and investments due to currency devaluation. Furthermore, high inflation can increase the operating costs for financial institutions, potentially leading to higher fees or lower interest rates on loans and deposits, thus affecting the availability, affordability, and accessibility of financial services.

e) Internet Users: The number of internet users, expressed as a proportion of the total population, reflects access to and utilization of internet facilities. The growth of internet access has significantly transformed the financial landscape, facilitating greater access to and engagement with financial products and services through the development of digital banking technologies. Online banking, mobile banking applications, and other fintech innovations have overcome traditional barriers to financial inclusion, such as geographical distance, high transaction costs, and limited financial literacy. These technologies enable individuals to conduct various financial transactions remotely, including payments, balance checks, fund transfers, and loan applications. This is particularly beneficial for individuals and households in rural areas with limited access to traditional banking infrastructure. Furthermore, the internet has fostered the emergence of new financial products and services, such as peer-to-peer lending platforms and crowdfunding platforms, which can cater to the needs of underserved or marginalized populations, providing access to credit and investment opportunities that may have been difficult to obtain through traditional banking channels. In summary, the growth in internet users has a profound positive impact on financial inclusion by expanding access, reducing costs, and promoting innovation in the financial sector.

Principal Component Analysis

Previous research on financial inclusion index construction has frequently relied on analytical and hierarchical procedures or axiomatic approaches. However, these methods have been subject to criticism due to their potentially flawed and restrictive weighting processes, despite sometimes appearing superficially appealing. This study addresses these limitations by employing Principal Component Analysis (PCA). PCA offers a distinct and objective method for assigning weights to various indicators, while also accommodating a larger number of variables in the index construction. For this study, five indicators related to Scheduled Commercial Banks are utilized to construct the financial inclusion index. PCA derives factor scores (weights) based on the eigenvalues of the indicators. Each variable's factor score is then multiplied by the corresponding standardized original variable. A key advantage of PCA is that it avoids imposing subjective weights and allows the data to determine the relative importance of each indicator (Kendall, 1939; Lenka & Barik, 2018). The composite financial inclusion (FI) index value for a given state *i* at time *t* is then calculated by summing these weighted variables. Formally, following the PCA procedure, FI can be expressed as:

$FII_i = W_{1i}\ SBA_i + W_{2i}\ BB_i + W_{3i}\ BE_i + W_{4i}\ DEP_i + W_{5i} CRE_i$

Where FII_i is the index of financial inclusion of i^{th} year and W_1, $W_2 \ldots W_5$ are the respective weights (factor scores) of different factors.

Table 2: Principal Component selection table

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	4.48049	3.99268	0.8961	0.8961
Comp2	0.487806	0.459869	0.0976	0.9937
Comp3	0.0279376	0.024908	0.0056	0.9992
Comp4	0.00302956	0.0022934	0.0006	0.9999
Comp5	0.0007362	.	0.0001	1

Principal component analysis was performed using E-views. The number of principal components to retain is determined by considering both cumulative variance and eigenvalues. A common criterion based on cumulative variance is to select the number of components that explain at least 70% of the total variance. The cumulative variance represents the total proportion of variance explained by all retained principal components up to a given component. Another widely used rule of thumb, based on eigenvalues, is to retain components with eigenvalues greater than 1. In this study, the results indicate that the first principal component (pc1) explains 89.61% of the total variance. Furthermore, only the first principal component has an eigenvalue exceeding 1. Thus, both the cumulative variance and eigenvalue criteria suggest retaining only the first principal component for further analysis.

Table 3: Loadings table

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Unexplained
Number of saving bank account	0.467	-0.125	-0.769	0.418	0.0104	0
Number of bank branches	0.469	-0.156	-0.153	-0.848	-0.117	0
Number of bank employees	0.359	0.9301	0.0734	0.011	0.0212	0
Amount of deposit	0.465	-0.23	0.3975	0.1255	0.7465	0
Amount of credit	0.466	-0.206	0.4705	0.3009	-0.655	0

The table generated by E-views provides the loadings for each principal component. These loadings represent the weights assigned to each original variable within each principal component. As only the first principal component is being retained, the weights for each variable are derived from the "comp1" column of the loadings table. These weights are subsequently multiplied by their corresponding standardized original variable observations and summed, as previously described in the index construction equation. This process yields the desired financial inclusion index values.

Unit root test

To analyse the determinants of financial inclusion, this study utilizes time series data for four independent variables and the constructed financial inclusion index (the dependent variable). A common challenge in time series analysis is the potential for non-stationarity. A time series is considered stationary if it exhibits a constant mean and variance over time. Conversely, a non-stationary time series displays changes in its mean, variance, or both. Employing non-stationary data in regression analysis can lead to spurious regression, also known as nonsense regression. This occurs when a regression analysis suggests a statistically significant relationship between variables when, in fact, no such relationship exists. Unit root tests, such as the Dickey-Fuller, Augmented Dickey-Fuller (ADF), or Phillips-Perron tests, are employed to determine the presence of a unit root, which indicates non-stationarity. In this study, the Augmented Dickey-Fuller (ADF) test is used to assess the stationarity of the variables.

Null Hypothesis: AGE_DEPENDENCY_RATIO has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.305347	0.0002
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

*MacKinnon (1996) one-sided p-values.

The stationarity of the age dependency ratio data was assessed using the Augmented Dickey-Fuller (ADF) test within E-views. The ADF test results indicate that the age dependency ratio is stationary at its level, meaning its mean and variance are constant over the time period examined.

The null hypothesis in the ADF test posits that the variable has a unit root (i.e., is non-stationary). The decision to reject or fail to reject the null hypothesis is based on the t-statistic and the associated p-value. If the p-value is less than the chosen significance level (typically 0.05), the null hypothesis is rejected, implying that the variable is stationary.

In this case, the p-value for the age dependency ratio is 0.0002, which is less than 0.05. Therefore, the null hypothesis is rejected, and it is concluded that the age dependency ratio is stationary at its level.

Null Hypothesis: GDP_GROWTH_RATE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.179388	0.0033
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

*MacKinnon (1996) one-sided p-values.

The absolute levels of GDP per capita were found to be non-stationary, even after second-differencing. Therefore, the GDP growth rate was selected as the relevant independent variable for the analysis. The Augmented Dickey-Fuller (ADF) test was applied to the GDP growth rate at its level. The resulting p-value of 0.0033 is less than the conventional significance level of 0.05. This result indicates that the GDP growth rate is stationary at its level.

Null Hypothesis: D(GROSS_SECONDARY_SCHOOL_ENROLLMENT_RATIO) has a unit ...
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.794595	0.0007
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

The Augmented Dickey-Fuller (ADF) test results for the gross secondary school enrolment ratio at its level indicated non-stationarity, with a p-value of 0.99. Consequently, the variable was tested for a unit root at its first difference. The results show that the gross secondary school enrolment ratio becomes stationary after first-differencing, with a p-value of 0.0007, which is less than the significance level of 0.05. Therefore, the first-differenced gross secondary school enrolment ratio, calculated as the difference between the current year's value and the previous year's value, will be used as the independent variable in the subsequent analysis.

Null Hypothesis: D(INFLATION__CPI___IN___) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-6.677398	0.0000
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

The inflation rate, as measured by the CPI, was found to have a unit root at its level, with a p-value of 0.06. Similar to the gross secondary school enrolment ratio, the inflation rate becomes stationary after first-differencing. At the first difference level, the ADF test yields a p-value of 0.000007. This result allows us to reject the null hypothesis of a unit root at the first difference level, confirming that the first-differenced inflation rate will be used in the analysis.

Null Hypothesis: INT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=6)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.113163	0.0003
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

The final independent variable considered is the number of internet users. A logarithmic transformation was applied to this variable before testing for a unit root using the Augmented Dickey-Fuller (ADF) test. The resulting p-value of 0.0003, which is less than 0.05, indicates that the log-transformed number of internet users is stationary at its level.

Cointegration

Historically, time series analysis was conducted using linear regression before the development of cointegration tests. However, Granger and Newbold critiqued this method, highlighting that applying linear regression to time series data could generate misleading correlations. In 1987, Engle and Granger introduced the concept of the cointegrating vector approach, providing a statistical framework to determine whether two or more non-stationary time series variables exhibit cointegration. This signifies a long-term equilibrium relationship that remains stable despite short-term fluctuations. In econometrics and time series analysis, cointegration is a fundamental concept, as it helps identify enduring associations between economic variables, aiding in forecasting and policy formulation.

When considering two time series variables, denoted as and, both integrated of order one (I(1)), if there exists a coefficient such that the linear combination results in an I(0) (stationary) series, then the variables are said to be cointegrated. In simpler terms, cointegration implies that these variables share a common stochastic trend, which can be eliminated through a specific transformation, rendering the resulting series stationary.

Approaches to Cointegration Analysis

There are two primary approaches for analysing cointegration: the Engle-Granger test and the Johansen test.

Engle-Granger Test

The Engle-Granger two-step methodology initially involves estimating residuals using a static regression model, followed by testing these residuals for unit roots. This approach employs the Augmented Dickey-Fuller (ADF) test or other similar methods to assess stationarity. If the residuals are found to be stationary, cointegration is confirmed. However, the Engle-Granger method has limitations—it is designed for single-equation models, making it less effective when multiple variables exhibit multiple cointegrating relationships.

Johansen Test

The Johansen cointegration test determines the presence of long-run relationships among multiple non-stationary time series variables. This test estimates a Vector Error Correction Model (VECM) and determines the number of cointegrating relationships by analysing the eigenvalues and eigenvectors of the residual matrix. Unlike the Engle-Granger test, the Johansen test accommodates more than two non-stationary time series variables and identifies multiple cointegrating relationships. However, it requires a large sample size due to its reliance on asymptotic properties.

The Johansen test consists of two primary approaches:

Trace Test

The trace test assesses the number of cointegrating relationships among variables. The null hypothesis is: and the alternative hypothesis is: where represents the number of cointegrating relationships, and is a pre-specified value, typically set to zero. The trace test evaluates the sum of eigenvalues and compares them to a set of critical values. If the computed sum exceeds the critical value, the null hypothesis is rejected, indicating the presence of at least one cointegrating relationship.

Maximum Eigenvalue Test

The maximum eigenvalue test, another approach within the Johansen framework, determines the number of cointegrating relationships by comparing the largest eigenvalue of the residual matrix to critical values. If the highest eigenvalue surpasses the critical threshold, the null hypothesis of no cointegration is rejected, confirming the presence of at least one cointegrating relationship. The total number of cointegrating relationships is determined by identifying how many eigenvalues exceed the threshold.

Comparing the Tests

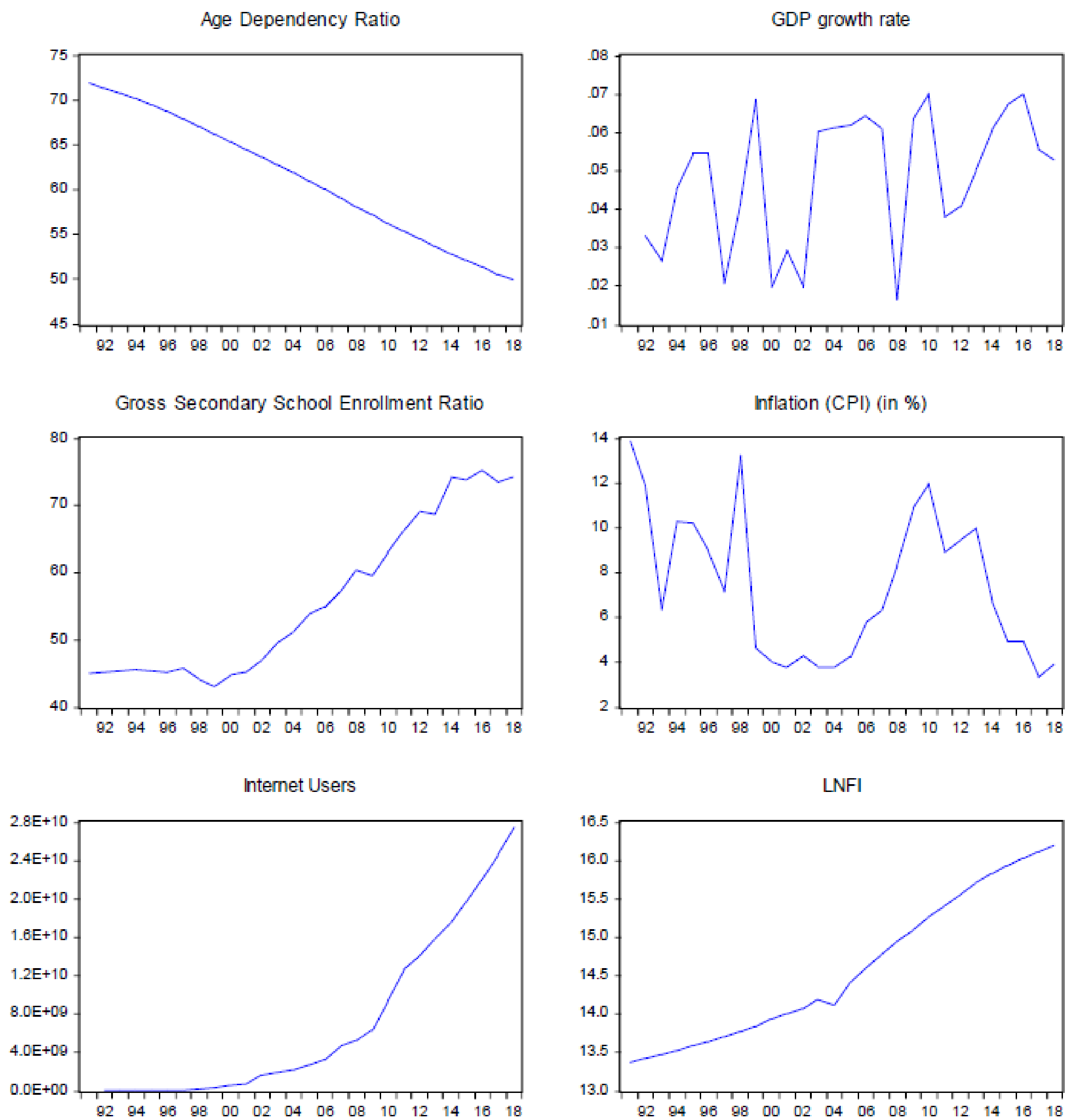
When comparing the trace test and the maximum eigenvalue test within the Johansen framework, the maximum eigenvalue test tends to be more robust when the number of time series variables is large relative to the sample size. Conversely, when the sample size is sufficiently large, the trace test often provides more reliable results.

In conclusion, selecting between the Engle-Granger and Johansen tests depends on the complexity of the dataset. The Engle-Granger method is more straightforward and suitable for two-variable cases, whereas the Johansen test is preferable for analysing multiple variables, as it identifies multiple cointegrating relationships and offers a comprehensive analysis of long-term equilibrium dynamics.

Cointegration Equation:

$$\log FII = \alpha_1 \text{GDP growth rate} + \alpha_2 \text{Inflation rate} + \alpha_3 \text{Age dependency ratio} + \alpha_4 \log \text{Internet users} + \alpha_5 \text{GSSER}$$

Graphical Representation



The graphical representations of the time series variables used in this study reveal several trends. The age dependency ratio has generally decreased over the period from 1991 to 2018, while the gross secondary school enrolment ratio, the number of internet users, and the financial inclusion index have shown increasing trends. Inflation, as measured by the CPI, has fluctuated considerably but appears to be at a lower level in 2018 compared to 1991. The GDP growth rate exhibits a generally positive trend over the period. These visualizations provide a preliminary overview of the variables' behaviour over time, which will be further investigated through quantitative analysis.

Optimal lag length selection

VAR Lag Order Selection Criteria
Endogenous variables: AGE DEPENDENCY RATIO GDP GROWTH RATE G...
Exogenous variables: C

Sample: 1991 2018
Included observations: 25

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-671.0359	NA	1.34e+16	54.16287	54.45540	54.24401
1	-438.2602	335.1971	2.14e+09	38.42081	40.46853	38.98876
2	-366.8527	68.55115*	2.26e+08*	35.58822*	39.39111*	36.64298*

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Cointegration analysis, followed by the Vector Error Correction Model (VECM), was employed to investigate the determinants of financial inclusion in India. The optimal lag length for the cointegration analysis was determined using the Akaike Information Criterion (AIC), which suggested two lags. However, due to constraints in the Johansen cointegration test, a single lag was ultimately used in the analysis.

VI. RESULTS AND INTERPRETATION

PCA Result

Table 4: Financial Inclusion Index

Year	FII (in Millions)		Year	FII (in Millions)
1991	647,050		1992	674,144
1993	705,669		1994	743,389
1995	793,471		1996	842,664
1997	887,981		1998	957,531
1999	1,027,316		2000	1,127,889
2001	1,201,058		2002	1,296,804
2003	1,447,741		2004	1,359,215
2005	1,834,646		2006	2,199,354
2007	2,645,755		2008	3,124,728
2009	3,626,965		2010	4,235,455
2011	4,970,217		2012	5,693,569
2013	6,610,207		2014	7,503,023
2015	8,236,830		2016	9,131,056
2017	9,931,902		2018	10,710,902

The table presents the calculated values of the financial inclusion index for India from 1991 to 2018. These values were derived using the following weights for the respective components: 0.467 for the number of savings bank accounts, 0.469 for the number of bank branches, 0.359 for the number of bank employees, 0.465 for the amount of deposits, and 0.466 for the amount of credit. These weights, obtained through principal component analysis, reflect the relative importance of each component in explaining the overall variation in financial inclusion.

Unit Root Test Result

Table 5: Unit root test results

Variables	ADF (Levels)	p-value	ADF (first difference)	p-value
Age dependency ratio	-5.3053	0.0002		
GDP growth rate	-4.1793	0.0033		
Inflation rate	-2.8832	0.0605	-6.6773	0.0000
Log of internet users	-5.1131	0.0003		
Gross secondary school enrolment ratio	0.7168	0.9903	-4.7945	0.0007
Log of financial inclusion index	1.6609	0.9993	-4.0967	0.0040

The results indicate that age dependency ratio, GDP growth rate, and log values of number of internet users are stationary at levels. While inflation rate, gross secondary school enrolment ratio, and log values of financial inclusion index are stationary at first difference levels. Thus, some variables are integrated of the first order while the rest have zero order of integration.

Cointegration results

Sample (adjusted): 1994 2018
Included observations: 25 after adjustments
Trend assumption: Quadratic deterministic trend
Series: LNFI AGE DEPENDENCY RATIO GDP GROWTH RATE GROSS...
Lags interval (in first differences): 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.893445	158.0466	107.3466	0.0000
At most 1 *	0.819650	102.0692	79.34145	0.0004
At most 2 *	0.681624	59.24776	55.24578	0.0213
At most 3	0.578554	30.63470	35.01090	0.1364
At most 4	0.294320	9.033103	18.39771	0.5767
At most 5	0.012650	0.318259	3.841465	0.5727

Trace test indicates 3 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.893445	55.97738	43.41977	0.0014
At most 1 *	0.819650	42.82143	37.16359	0.0101
At most 2	0.681624	28.61306	30.81507	0.0908
At most 3	0.578554	21.60160	24.25202	0.1079
At most 4	0.294320	8.714843	17.14769	0.5255
At most 5	0.012650	0.318259	3.841465	0.5727

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Cointegration analysis was performed to examine the long-run relationship between the variables: log of financial inclusion index, age dependency ratio, GDP growth rate, inflation rate (CPI), number of internet users, and gross secondary school enrolment ratio. The Trace test indicated three cointegrating equations, while the Maximum Eigenvalue test suggested two cointegrating equations at the 0.05 significance level. As both tests confirm the presence of at least two cointegrating equations, the study concludes that a long-run relationship exists between the dependent variable (financial inclusion index) and the independent variables. This finding implies that these variables tend to move together in the long term, even though they may experience short-term deviations.

1 Cointegrating Equation(s): Log likelihood 97.55790

Normalized cointegrating coefficients (standard error in parentheses)

LNFI	AGE DEPE...	GDP GRO...	GROSS S...	INFLATION ...	LNINT
1.000000	0.750879	-4.871731	0.002120	-0.006945	-0.071788
	(0.11172)	(0.71716)	(0.00648)	(0.00463)	(0.01995)

The normalized cointegrating coefficients table shows that there is a significant negative relationship in the long run between age dependency ratio and financial inclusion index. GDP growth rate and number of internet users positively affects the financial inclusion index in the long run. Inflation also positively affects financial inclusion but the relationship is not significant. While gross secondary school enrolment ratio has a non-significant negative relationship with financial inclusion.

VII. CONCLUSION

This study investigates the determinants of financial inclusion in India from 1991 to 2018. A composite financial inclusion index was constructed using five indicators. The analysis explores the impact of GDP growth rate, inflation (measured by the CPI), internet penetration, the age dependency ratio, and the gross secondary school enrolment ratio on financial inclusion. Given the time-series nature of the data, stationarity tests were conducted for all variables. The age dependency ratio and GDP growth rate were found to be stationary at their original levels, while the remaining variables, including the financial inclusion index, achieved stationarity after first-differencing. The Johansen cointegration test was employed to examine the long-run equilibrium relationship between the dependent variable (financial inclusion index) and the independent variables.

VIII. LIMITATION AND SCOPE FOR FURTHER RESULT

The absence of a globally standardized measure for financial inclusion presents a significant challenge for researchers. A consistent metric is essential for cross-national comparisons of financial inclusion levels and for tracking individual countries' progress over time. Recognizing the importance of financial inclusion for inclusive economic growth, development, and improved living standards, this study contributes to the literature. While the methodology employed for constructing the financial inclusion index has been utilized by other researchers with varying indicator sets, a definitive and universally accepted measure of financial inclusion remains an open area for development.

Another persistent challenge in financial inclusion research is data availability. Data on key indicators, such as the number of ATMs, usage of informal financial services, and other relevant metrics, are often limited or unavailable for extended periods. Analysing the determinants of financial inclusion, as well as its impact on poverty, inequality, and economic growth, can provide valuable insights for policymakers. Therefore, further research on financial inclusion, particularly at the microeconomic level (district, state, regional, and country levels), is warranted and holds considerable promise for informing policy and practice.

REFERENCES

[1] Allen, F., Demirguc-Kunt, A., Klapper, L., & Peria, M. S. M. (2016). The foundations of financial inclusion: Understanding ownership and use of formal accounts. *Journal of Financial Intermediation*, *27*, 1–30. https://doi.org/10.1016/j.jfi.2015.02.004

[2] Arif, B. D., & Ahmed, F. (2021). Financial inclusion determinants and impediments in India: Insights from the global financial inclusion index. *Journal of Financial Economic Policy*, *13*(3), 391–408. https://doi.org/10.1108/JFEP-11-2019-0227

[3] Beck, T., Demirgüç-Kunt, A., & Honohan, P. (2009). Access to financial services: Measurement, impact, and policies. *The World Bank Research Observer*, *24*(1), 119–145. https://doi.org/10.1093/wbro/lkn008

[4] Chakravarty, S. R., & Pal, R. (2013). Financial inclusion in India: An axiomatic approach. *Journal of Policy Modeling*, *35*(5), 813–837. https://doi.org/10.1016/j.jpolmod.2012.12.007

[5] Demirgüç-Kunt, A., & Klapper, L. (2013). Measuring financial inclusion: Explaining variation in use of financial services across and within countries. *Brookings Papers on Economic Activity*, 279–321.

[6] Kumar, N. (2013). Financial inclusion and its determinants: Evidence from India. *Journal of Financial Economic Policy*, *5*(1), 4–19. https://doi.org/10.1108/17576381311317754

[7] Omar, M. A., & Inaba, K. (2020). Does financial inclusion reduce poverty and income inequality in developing countries? A panel data analysis. *Journal of Economic Structures*, *9*(1). https://doi.org/10.1186/s40008-020-00214-4

[8] Park, C.-Y., & Mercado, R. J. (2015). Financial inclusion, poverty, and income inequality in

developing Asia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2558936
[9] Simon, T. D. (2020). Financial inclusion in India: An analysis of its pattern. *Artha Journal of Social Sciences*, *19*(2), 37–51. https://doi.org/10.12724/ajss.53.3
[10] Yadav, P., & Sharma, A. K. (2016). Financial inclusion in India: An application of topsis. *Humanomics*, *32*(3), 328–351. https://doi.org/10.1108/h-09-2015-0061