PressAcademia

JBEF

Journal of Business, Economics & Finance

PressAcademia publishes journals, books, case studies, conference proceedings and organizes international conferences.

jbef@pressacademia.org

ISSN 2146-7943

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

ABOUT THE JOURNAL

Journal of Business, Economics and Finance (JBEF) is a scientific, academic, double-blind peer-reviewed, semiannually and open-access journal. The publication language is English. The journal publishes two issues a year. The issuing months are June and December. The journal aims to provide a research source for all practitioners, policy makers and researchers working in the areas of business, economics and finance. The Editor of JBEF invites all manuscripts that that cover theoretical and/or applied research on topics related to the interest areas of the Journal.

JBEF is currently indexed/archived by
Ulrich's Directiroy, ASCI, Scope, Open J-Gate, New Jour, Princeton University Library,
International Scientific Indexing, Directory of Research Journals Indexing (DRJI),
International Society for Research Activity (ISRA), InfoBaseIndex, Scientific Indexing Services (SIS),
Google Scholar, Root Indexing, Journal Factor Indexing, International Institute of Organized Research (I2OR),
ASOS, Cabell's Directory, SOBIAD, EuroPub, ASCI.

Ethics Policy
JBEF applies the standards of Committee on Publication Ethics (COPE). JBEF is committed to the academic community ensuring ethics and quality of manuscripts in publications. Plagiarism is strictly forbidden and the manuscripts found to be plagiarised will not be accepted or if published will be removed from the publication.

Author Guidelines
All manuscripts must use the journal format for submissions.
Visit www.pressacademia.org/journals/jbef/guidelines for details.

CALL FOR PAPERS
The next issue of JBEF will be published in December 2026.
Submit manuscripts to
jbef@pressacademia.org or
http://www.pressacademia.org/submit-manuscript/
Web: www.pressacademia.org/journals/jbef

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

EDITORIAL BOARD

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

CONTENT

Title and Author/s **Page**

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

THE INFLUENCE OF GENERATIVE ARTIFICIAL INTELLIGENCE ON FINANCIAL MARKET VOLATILITY, LIQUIDITY, AND PREDICTABILITY

DOI: 10.17261/Pressacademia.2026.2029
JBEF- V.15-ISS.1-2026(1)-p.1-15

Nada Dammak
University of Sfax, Faculty of Economics and Management of Sfax (FSEG), Tunisia.
nada_dammak@yahoo.fr , ORCID: 0009-0006-0737-9182

Date Received: December 29, 2025 **Date Accepted:** April 4, 2026

To cite this document
Dammak, N., (2026). The influence of generative artificial intelligence on financial market volatility, liquidity, and predictability. Journal of Business, Economics and Finance (JBEF), 15(1), 1-15.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2029

ABSTRACT
Purpose- This study examines the impact of generative artificial intelligence (AI) on key dimensions of financial market dynamics, namely volatility, liquidity, and return predictability. It emphasizes the dual role of AI as both an enhancer of informational efficiency and a potential source of short-term market instability.
Methodology- The paper proposes an original AI-based sentiment indicator constructed using a hybrid large language model (LLM) framework that combines FinBERT and a GPT-4–class generative model. This sentiment index is analyzed alongside daily returns of major equity indices—S&P 500, NASDAQ, and STOXX 600—over the period 2018–2024. The empirical analysis relies on Principal Component Analysis (PCA), GARCH (1,1)-X models, and a Vector Autoregression (VAR) framework incorporating Amihud's illiquidity measure, impulse response functions (IRFs), and forecast error variance decomposition (FEVD).
Findings- Results from the GARCH-X estimations indicate that AI-driven sentiment is a statistically significant determinant of conditional volatility in U.S. equity markets. VAR-based Granger causality tests reveal a bidirectional relationship between AI sentiment and market returns, with particularly strong predictive effects for the S&P 500 and NASDAQ. Positive sentiment shocks are associated with improved market liquidity, as reflected by declines in the Amihud illiquidity ratio, while European markets display slower and weaker responses relative to U.S. markets.
Conclusion- Generative AI functions as a double-edged mechanism in financial markets: it accelerates information processing and enhances short-term predictability, yet it may also amplify transient volatility through synchronized sentiment effects. Although AI has not fundamentally altered long-term market structures, its growing influence calls for renewed regulatory attention to AI-generated information flows and their implications for market stability

Keywords: Generative artificial intelligence, market volatility, AI sentiment, algorithmic herding, large language models.
JEL Codes: G14, G15, C32

1. INTRODUCTION

The meteoric rise of generative artificial intelligence (AI) represents a paradigm shift in the technological landscape, increasingly reshaping the architecture and behavioral efficiency of contemporary financial ecosystems. Unlike traditional AI systems, which were largely restricted to data classification or basic algorithmic optimization, the current generation—powered by Large Language Models (LLMs) and transformer architectures—possesses the unique ability to synthesize and interpret complex financial knowledge at unprecedented scales. As these tools become deeply integrated into institutional trading platforms and investor workflows, they demand a rigorous re-evaluation of market functionality in an era dominated by machine intelligence.

A critical question emerges from this shift: how does generative AI affect the core pillars of market stability—specifically volatility, liquidity, and predictability? As these systems are increasingly deployed for risk management and real-time surveillance, they inevitably reshape price discovery mechanisms. For instance, LLM-derived sentiment indicators have transitioned from niche metrics to essential inputs in modern forecasting frameworks. Moreover, AI-assisted systems and semi-autonomous agents are increasingly used to execute strategies and synchronize behaviors across diverse participants, potentially intensifying market responses to new information.

The consequences of this integration are multifaceted. On one hand, generative AI may bolster market efficiency by refining the processing of unstructured data, thereby reducing noise and stabilizing price adjustments. On the other hand, it may introduce novel vulnerabilities. If a majority of market actors converge on similar AI-driven signals, the resulting "herding" behavior could trigger sudden liquidity shortages or feedback loops, particularly during intervals of high macroeconomic stress.

Consequently, the footprint of generative AI is analyzed across three fundamental dimensions. In terms of volatility, AI-driven signals can either mitigate uncertainty through efficiency or exacerbate it via algorithmic synchronization. Regarding liquidity, the acceleration of market-making processes via AI alters order-flow patterns and market depth. Finally, for predictability, while AI enhances the interpretation of unstructured data, its widespread adoption might paradoxically lead to hyper-efficient markets where forecasting becomes increasingly challenging.

This study contributes to the current academic discourse by empirically assessing these dynamics. By linking daily returns from the S&P 500, NASDAQ, and STOXX 600 with a custom AI-sentiment index, we utilize a comprehensive econometric suite—including PCA, GARCH (1,1) modeling, and VAR frameworks—to map the interactions between AI-generated information and market stability.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive review of the existing literature regarding Large Language Models (LLMs) in finance and their impact on market behavior. Section 3 outlines the data collection process and the methodology, detailing the construction of the hybrid AI sentiment index and the econometric models employed. Section 4 presents the empirical results, including volatility analysis and liquidity shocks. Section 5 discusses the implications of these findings, while Section 6 concludes with a summary of the research and recommendations for future policy and practice.

2. LITERATURE REVIEW

2.1. Generative AI and Large Language Models in Finance

The literature on large language models (LLMs) and generative AI in finance expanded rapidly after 2020, moving from exploratory applications toward production-grade tasks such as automated research, narrative summarization, and automated signal generation. Furthermore, Yang et al. (2023) demonstrate that utilizing models like GPT-4 for financial report analysis significantly outperforms traditional bag-of-words methods by capturing complex contextual nuances. Similarly, Wu et al. (2023) introduced BloombergGPT, a 50-billion parameter model trained specifically on financial data, proving that domain specialization radically improves prediction accuracy compared to generalist models. Key takeaways are that LLMs materially expand the feature space, but naive adoption without domain-specific adaptation frequently yields unstable outputs.

2.2. Sentiment Extraction from Text (AI-Based) and its Link to Returns

A large body of research studies how textual sentiment affects stock returns and volatility. Recent empirical evidence from Lopez-Lira and Tang (2023) shows that sentiment scores generated by ChatGPT possess statistically significant predictive power over daily stock returns, surpassing classical sentiment databases. Additionally, Fatemi et al. (2024) explores how generative AI synthesizes divergent opinions during earnings calls, revealing that AI detects executive hesitation more effectively than human analysts. These studies underline that sentiment signals derived from curated disclosures tend to be more informative than noisy social media streams.

2.3. Volatility Forecasting: Machine Learning Improvements and Limits

A parallel literature evaluates whether modern machine learning (ML) models improve volatility forecasts over traditional GARCH models. In terms of volatility, Hansen and Kazinnik (2023) utilize LLMs to analyze Federal Reserve communications, showing that extracted sentiment improves the forecasting of implied volatility (VIX). Moreover, Rane et al. (2024) propose a hybrid architecture combining LSTM networks with AI-generated textual signals, reducing the mean squared error (MSE) of volatility forecasts by 15% compared to traditional models. The evidence supports using AI as a complementary input rather than a full replacement for established econometric models

2.4. Liquidity, Market Microstructure and Algorithmic/AI Trading

Research on AI's effect on liquidity is growing. Brummer et al. (2024) analyze the impact of AI-based algorithmic trading on order-book depth, noting improved liquidity during calm periods but increased fragility during earnings announcements. This aligns with Kozhan and Tham (2023), who conclude that AI reduces price reaction time to news to milliseconds, thereby compressing traditional arbitrage opportunities. While faster information processing can enhance liquidity, homogeneous adoption of similar AI strategies can make market depth more fragile during periods of stress.

Empirical contributions also show that AI-based market making and execution can reduce transaction costs for many securities but might increase episodic illiquidity around news shocks where algorithms withdraw or shift quotes simultaneously. These dynamics mean that liquidity effects are state-dependent — beneficial in stable regimes but potentially destabilizing during high uncertainty.

2.5. Herding, Feedback Loops, and Systemic Risk: When AI Amplifies Market Moves

A key concern is the potential for AI systems to produce feedback loops. Uliari et al. (2024) document an "algorithmic mimicry" effect where hedge funds using similar language models converge on the same positions, increasing the risk of flash crashes. Finally, the Financial Stability Board (FSB, 2024) emphasizes that the market concentration of model providers creates a single point of failure for the global financial system. Model-driven herding can be especially pronounced when generative models produce market narratives that are then consumed by other models, creating self-reinforcing price dynamics.

2.6. Synthesis, Limitations, and Research Gaps

Beyond governance and replicability concerns, the literature reveals several structural gaps that motivate the present study

The reviewed literature indicates that AI — and generative AI in particular — provides valuable new information channels that enhance short-term forecasting. However, several gaps remain:

- Integrated causal evidence: While studies like Lopez-Lira and Tang (2023) show strong correlations, causal identification of how AI deployment changes market volatility at the system level remains limited.
- Generative AI vs. classical ML: Comparative work specifically isolating the additional value of generative models over previous architectures (like simple RNNs or VADER) is still emerging.
- Macro- and systemic-level assessments: Most empirical work focuses on asset-level consequences. There is a critical need for research, such as this study, that addresses system-wide impacts across multiple indices (S&P500, NASDAQ, STOXX600).

This study contributes to these gaps by integrating an AI-generated sentiment index with classical market variables, offering richer evidence on short-run dynamics and variance contributions.

3. METHODOLOGY

This section presents a comprehensive econometric framework designed to evaluate the impact of generative AI on market volatility, liquidity, and predictability. The empirical strategy follows a multi-layered approach aligned with the data and procedures described in the main body of the paper. It integrates descriptive analysis, dimensionality reduction, univariate volatility modeling, and multivariate dynamic interactions using VAR, impulse response functions, and forecast error variance decomposition techniques. Each methodological stage explicitly relies on the datasets and summary statistics reported in Tables 1 and 2.

3.1. Data and Variable Construction

The study utilizes daily observations from four key variables: SP500 daily returns, NASDAQ daily returns, STOXX600 daily returns, and AI-generated sentiment scores.

The price indices were transformed into log returns to ensure stationarity and comparability across markets:

$$r_t = ln(P_t) - ln(P_{t-1}) \tag{1}$$

Where r_t the return on day t is, P_t is the asset price on day t, P_{t-1} is the price on the previous day.

The AI sentiment index is a continuous variable ranging approximately from –2.4 to +2.5, as reported in Table 1, exhibiting substantial variance (std ≈ 0.978), which makes it a potentially powerful explanatory factor for market dynamics.

All series were aligned over a primary sample spanning from early 2018 to the end of 2024, encompassing approximately 1750 daily observations. This extensive period was chosen to significantly enhance the robustness of the econometric findings and to capture the full lifecycle of Generative AI, from the pre-ChatGPT era to its widespread market adoption. This timeframe ensures the analysis covers diverse market regimes, including periods of high technological integration, enabling a rigorous assessment of AI's systemic impact.

3.2. Construction of the AI Sentiment Index

To quantify the informational content generated by large language models, this study constructs an AI-based sentiment index derived from a curated corpus of financial texts. The objective is to capture daily shifts in market-relevant narratives that reflect how generative AI interprets and synthesizes financial information. This section details the model architecture, data sources, preprocessing pipeline, scoring methodology, and normalization procedures used to compute the sentiment index.

Model Architecture - The sentiment index is constructed using a transformer-based large language model fine-tuned for financial sentiment analysis. Specifically, the study employs a hybrid LLM approach combining:

- FinBERT (Araci, 2019; Yang et al., 2023) for domain-specific financial text classification, and
- A GPT-4–class generative large language model (LLM) to synthesize sentiment probabilities and contextual polarity scores.

FinBERT is used because it provides high accuracy on financial tone classification, while the GPT-based model enhances contextual understanding by extracting latent sentiment from longer multi-topic disclosures and news narratives. This dual approach reduces misclassification risk and captures richer narrative structures than using FinBERT alone.

Textual dataset - The sentiment index is built using daily textual data from widely used, market-relevant sources:

- Financial news headlines and articles from major outlets (Reuters, Bloomberg, MarketWatch).
- Corporate disclosures (CEO/CFO statements, earnings-call summaries).
- Macroeconomic commentary (central-bank speeches, policy announcements).
- AI-generated market summaries (GPT-based financial digests published daily by platforms specialized in LLM-generated content).

All documents are aggregated at a daily frequency using publication timestamps to ensure temporal consistency with financial-market returns. Only English-language documents are used to avoid translation noise.

Preprocessing Pipeline - To prepare the textual data for model inference, the following preprocessing steps are implemented:

- **Cleaning and Tokenization**
  - Removal of HTML tags, URLs, numbers, and stopwords.
  - Sentence tokenization using a transformer-compatible tokenizer.
- **Filtering for Market-Relevant Content** - A keyword filter ensures that only documents mentioning asset prices, macroeconomic conditions, risk, volatility, or financial institutions are retained.
- **Deduplication -** Near-duplicate articles (similarity > 90% cosine similarity using sentence embeddings) are removed.
- **Chunking -** Long documents (e.g., earnings call transcripts) are segmented into blocks of 150–250 tokens to avoid LLM truncation.
- **Embedding and Classification** - Each chunk is fed into:
  - FinBERT → produces probabilities (positive, neutral, negative)
  - GPT Model → produces a contextual polarity score in free-text form, converted to sentiment probabilities

The two signals are then combined through weighted averaging (60% FinBERT, 40% GPT), reflecting FinBERT's calibration in financial tasks and GPT's strength in context interpretation.

Sentiment Scoring Procedure

For each text chunk *i*, the model outputs: P_i^{pos}, P_i^{neu}, P_i^{neg}

A raw sentiment score is computed as:

$$S_i = P_i^{pos} - P_i^{neg} \tag{2}$$

Where P_i^{pos} and P_i^{neg} are the respective positive and negative probabilities produced by the hybrid FinBERT/GPT model for text block i

Daily scores are aggregated by simple averaging:

$$S_t = \frac{1}{N_t}\sum_{i=1}^{N_t} S_i \tag{3}$$

where N_t is the number of text chunks on day *t*.

This formulation ensures that positive (negative) news increases (decreases) the index, while neutral content has limited impact.

Normalization and Scaling - Because the number of daily documents varies over time, the unscaled series could suffer from heteroskedasticity. Therefore, the aggregated scores undergo:

- **Z-score normalization**
  - $Sentiment_t = \frac{S_t - \mu_S}{\sigma_S}$ (4)
  - **Outlier Winsorization** at the 2.5% and 97.5% quantiles to prevent news bursts (e.g., geopolitical shocks) from dominating the series.
- **Rescaling** to the interval [−3,+3] [-3, +3] [−3,+3] using a linear min–max transformation to match the dispersion typically reported in sentiment-based volatility studies.

The resulting index ranges approximately **from** –2.4 to +2.5**,** consistent with the summary statistics in Table 1.

Justification of the Modeling Choice - This methodological framework is motivated by several considerations:

- **FinBERT ensures financial domain precision**, outperforming generic sentiment models on finance-specific vocabulary.
- **GPT-based contextual scoring captures latent narratives**, sarcasm, macro-tone, and multi-topic sentiment that FinBERT alone cannot detect.
- **Multi-source textual data** reflects the full spectrum of daily financial information flows.
- **Normalization procedures** ensure comparability over time and prevent structural breaks due to variations in news volume.
- **Hybrid scoring aligns with recent empirical evidence** showing that LLM-enhanced sentiment models improve predictive performance for volatility and returns.

Overall, the construction of the AI Sentiment Index integrates both domain-specialized classification and generative-AI contextual interpretation, producing a robust, information-rich indicator suitable for econometric modeling.

3.3. Descriptive Statistics and Diagnostic Analysis

Summary statistics reported in Table 1 are used to assess the central tendency and dispersion of each variable. Stock index returns display small positive mean values (0.00038–0.00055) and standard deviations consistent with typical daily financial volatility (0.008–0.013).

The AI Sentiment Index is a standardized variable obtained after z-score normalization and winsorization, followed by linear rescaling to the interval [−3, +3]. As reported in Table 1, the index exhibits substantial dispersion, with a standard deviation of approximately 0.98 and values ranging from −2.43 to +2.52. This relatively wide range reflects the high variability of AI-generated narratives and supports its relevance as a potential driver of short-term market dynamics.

Before model estimation, standard diagnostic tests were conducted. Augmented Dickey–Fuller (ADF) tests confirm the stationarity of return series. Jarque–Bera tests, based on summary statistics and distributional asymmetry, indicate the presence of heavy tails and deviations from normality. Finally, the correlation structure reported in Table 2 reveals moderate cross-market linkages and weak-to-moderate correlations between AI sentiment and financial returns, thereby justifying the use of multivariate econometric frameworks.

3.4. Principal Component Analysis (PCA)

To explore the underlying common factors driving market returns, a principal component analysis (PCA) is conducted on standardized return series. Each principal component is defined as a weighted linear combination of the original variables, such that:

$$PC_k = w_{k1}X_1 + w_{k2}X_2 + \dots + w_{kn}X_n$$

The explained variance plot reported in Figure 1 indicates that the first principal component (PC1) captures the dominant global market movement, accounting for the majority of the variance across the S&P 500, NASDAQ, and STOXX 600 indices. The second and third components (PC2–PC3) explain additional idiosyncratic regional or sector-specific fluctuations, reflecting more localized sources of return variation.

Although the AI sentiment series is not included in the PCA, comparing PC1 with the AI sentiment index provides an intuitive benchmark to assess whether AI-generated sentiment is associated with systematic market risk rather than purely idiosyncratic noise. In this sense, the PCA serves to verify the presence of strong common market factors and to motivate the subsequent analysis of AI sentiment as a potential driver of aggregate market dynamics.

3.5. Modeling Conditional Volatility: GARCH (1,1)

To quantify short-term risk dynamics, a standard GARCH (1,1) model is estimated for S&P 500 returns. The model is specified as:

$$r_t = \mu + \varepsilon_t \qquad \varepsilon_t \sim N(0, h_t)$$
$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{5}$$

Where σ_t^2 is the conditional variance on day t, ω is the constant term, $\epsilon_{t-1}^2$is the squared innovation (information shock) on day t−1 (shock sensitivity parameter α), σ_{t-1}^2 is the lagged conditional variance (volatility persistence parameter β).

To test whether generative-AI-driven sentiment contributes directly to market volatility beyond its inherent persistence, the normalized AI Sentiment Index ($Sentiment_t$)is incorporated into the conditional variance equation. This augmented specification corresponds to a GARCH-X model, in which AI sentiment enters as an exogenous explanatory variable:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma Sentiment_t \tag{6}$$

The coefficient γ measures the marginal impact of AI-generated sentiment on conditional volatility. A statistically significant and positive estimate of γ would provide strong econometric evidence that AI-driven sentiment acts as a key determinant of short-term market uncertainty.

3.6. Liquidity Measurement and VAR Extension

To assess the influence of AI sentiment on market functioning, a key dimension is liquidity, which reflects the ease and cost of executing a trade. We primarily adopt the Amihud Illiquidity ratio ($ILLIQ$) as a robust daily proxy, given its reliance solely on readily available data (returns and volume). The Amihud ratio captures the price impact of trading volume, where higher values denote lower liquidity (higher illiquidity).

The Amihud illiquidity ratio is defined as the absolute daily return divided by the daily trading volume (in currency units):

$$ILLIQ_t = \frac{|r_t|}{volume_t} \tag{7}$$

We also consider secondary proxies such as raw Daily Trading Volume and the Bid-Ask Spread (if high-frequency data is available) to triangulate our findings.

3.7. Multivariate Framework: Vector Autoregression (VAR)

To capture the dynamic interdependencies among financial markets and AI-driven sentiment, a vector autoregression model of order ppp, VAR(p), is estimated based on lag length selection using the Akaike Information Criterion (AIC). The endogenous variable vector is defined as:

$$Y_t = \begin{pmatrix} SP500_t \\ NASDAQ_t \\ STOXX600_t \\ AI\ Sentiment_t \\ ILLIQ_t^{SPF500} \end{pmatrix} \tag{8}$$

Where Y_t is the vector of variables on day t including the returns of the three stock indices the AI sentiment index, and the Amihud illiquidity ratio.

The VAR(p) model is formally written as:

$$Y_t = c + \sum_{i=1}^{p} \Phi_i Y_{t-1} + \mu_t \tag{9}$$

Where c is a vector of intercepts, Φ are coefficient matrices capturing lagged interactions among the variables, and μ_t is a vector of innovations.

Given the moderate correlations reported in Table 2, the VAR framework is well suited to jointly assess (i) whether AI-generated sentiment Granger-causes financial returns, (ii) the presence of cross-market spillovers between U.S. and European equity markets, and (iii) the dynamic propagation of shocks across returns, liquidity, and sentiment.

Prior to estimation, the stability condition of the VAR system was verified by examining the inverse roots of the characteristic polynomial, ensuring that all roots lie inside the unit circle and that impulse response analysis is valid.

The VAR framework also provides the basis for conducting formal Granger causality tests. These tests are central to addressing the core research question of this study: whether AI-generated sentiment drives financial markets or whether market movements influence AI-generated sentiment. Specifically, the null hypothesis states that asset returns do not Granger-cause the AI Sentiment Index if the coefficients on all lagged return terms are jointly equal to zero in the sentiment equation of the VAR system. Rejecting this hypothesis indicates the presence of predictive causality running from market returns to AI-generated sentiment. Symmetrically, Granger causality from AI sentiment to stock returns is assessed by testing whether the lagged sentiment coefficients are jointly significant in the return equations. This framework allows for a rigorous examination of the direction of intertemporal causal linkages between AI-driven sentiment and the three equity indices.

3.8. Impulse Response Functions (IRF)

Impulse response functions are employed to trace the dynamic effects of a one-standard-deviation shock in AI-generated sentiment on equity market returns. Both generalized and orthogonalized IRFs are computed to assess the robustness of the results to alternative identification schemes. The corresponding responses are reported in Figure 4.

The IRF analysis reveals that U.S. equity markets, namely the S&P 500 and NASDAQ, exhibit an immediate and economically meaningful response to AI sentiment shocks. In contrast, the STOXX 600 displays a more delayed yet persistent reaction, suggesting slower price-adjustment mechanisms in European markets. While generalized IRFs confirm that these dynamics are robust to the ordering of variables, orthogonalized IRFs facilitate a clearer structural interpretation of sentiment-driven shocks.

Overall, the IRF evidence helps address the core research question of whether generative-AI-related sentiment acts as a destabilizing force, a stabilizing input, or a predictor of future return movements across international equity markets.

3.9. Forecast Error Variance Decomposition (FEVD)

Forecast error variance decomposition is used to quantify the contribution of AI sentiment shocks to the variability of stock returns over a 20-day forecast horizon. The FEVD results are presented in Figure 5.

For each equity index, the FEVD estimates the proportion of forecast uncertainty attributable to innovations in AI-generated sentiment. A larger contribution indicates stronger predictive content of sentiment for future returns, whereas a negligible share would suggest that its influence is confined to short-term noise. The results show that AI sentiment explains a non-negligible fraction of short-term forecast variance, particularly for the NASDAQ and the S&P 500, supporting the hypothesis that AI-generated information plays an increasingly important role in shaping equity price dynamics.

3.10. Robustness and Model Validation

Several diagnostic and robustness checks are conducted to ensure the reliability of the empirical results. Autocorrelation and partial autocorrelation functions (ACF/PACF) confirm appropriate lag length selection, while residual diagnostics reveal no major autocorrelation or heteroskedasticity issues in the estimated VAR system. Stability tests further indicate that all characteristic roots lie inside the unit circle, validating the use of impulse response analysis.

Additional robustness checks compare generalized and orthogonalized IRFs and examine the sensitivity of the results to alternative lag structures. Finally, sample stability is assessed by re-estimating the models over shorter subsamples, including the most recent 500 and 1,000 trading days. The core findings remain qualitatively unchanged, indicating that the results are not driven by specific market events or instability in the early part of the sample.

4. RESULTS

The empirical results provide a comprehensive picture of how generative-AI-driven sentiment interacts with market dynamics across volatility, co-movements, and predictability. Each model—ranging from descriptive statistics to PCA, GARCH, VAR, impulse response functions, and forecast error variance decomposition—offers complementary insights into the influence of AI-based information flows on financial markets. This section presents the findings in the order of the methodological steps, with explicit reference to the empirical results and figures reported in the main body of the paper.

4.1. Descriptive Statistics and Correlation Patterns

The summary statistics in Table 1 show that daily returns for the three equity indices exhibit small positive means and standard deviations consistent with typical market variability. The AI sentiment variable displays the highest dispersion (standard deviation $\approx$0.98) and the widest range (–2.435 to 2.516), highlighting its potential role as a high-amplitude informational input. These distributional characteristics are detailed in Table 1, providing a baseline for the subsequent econometric analysis

Table 1: Descriptive Statistics for Stock Returns, AI Sentiment Index, and Liquidity (2018–2024)

Variable	Mean	Std. Dev.	Min	Max	Jarque-Bera
S&P 500 Returns (%)	0.038	1.150	−12.45	9.05	450.12***
NASDAQ Returns (%)	0.052	1.340	−13.10	10.20	380.55***
STOXX600 Returns (%)	0.031	1.020	−10.85	8.40	295.47***
AI Sentiment Index	**0.041**	**0.978**	**−2.435**	**2.516**	**125.40***
Amihud Illiquidity Ratio	0.0012	0.0004	0.0001	0.0051	210.15***

Note: Returns are expressed in percentage terms. The AI Sentiment Index is standardized and rescaled to the interval [−3, +3]. *** denotes significance at the 1% level.

The relatively high dispersion of the AI Sentiment Index compared to return series highlights its role as a high-amplitude informational variable, capable of capturing rapid shifts in market narratives generated by large language models. The correlation matrix reveals moderate positive correlations among the SP500, NASDAQ, and STOXX600 returns, indicating strong market co-movement.

The statistical relationships between the variables are quantified in Table 2, which displays the Pearson correlation coefficients. The results confirm high integration between equity markets and a moderate but significant link between AI-driven signals and tech-heavy indices.

Table 2: Correlation Matrix between Equity Returns and AI Sentiment

Variable	SP500	NASDAQ	STOXX600	AI Sentiment
SP500	1.000			
NASDAQ	0.885***	1.000		
STOXX600	0.652***	0.584***	1.000	
AI Sentiment	0.241**	0.287**	0.154*	1.000

*Note: ***, *, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

AI sentiment, in contrast, shows only weak-to-moderate correlations with returns, suggesting that its relationship with market behavior cannot be adequately captured by static correlations alone. This justifies the use of dynamic models such as GARCH and VAR.

Parameter estimates indicate standard volatility clustering: α (shock sensitivity) and β (volatility persistence) are both significant and sum to a value close to 1, confirming high persistence. Figure 2 displays conditional volatility over the full sample, revealing several volatility spikes that can be compared with changes in AI sentiment: periods with sharp sentiment swings (as seen in min–max stats) tend to align with increases in conditional variance, but causality must be assessed with the VAR framework. The GARCH model therefore provides a univariate benchmark for understanding volatility behavior before introducing AI sentiment as an interacting factor in multivariate models.

The estimation of the GARCH-X model (Equation 6) yields a particularly illuminating result regarding the role of AI. For the major U.S. equity indices (SP500 and NASDAQ), the γ coefficient is estimated to be significantly positive ($p - value < 0.01$). This statistical significance of γ confirms that the AI-generated sentiment does not merely coincide with market uncertainty but acts as a causal and exogenous determinant of future market volatility. In other words, the information conveyed by the AI, as measured by our sentiment index, introduces a structural informational shock that directly increases perceived risk. This finding strengthens the conclusion that Generative AI acts as a powerful amplifier of short-term uncertainty, providing the necessary econometric support beyond historical volatility dynamics (measured by α and β). The full estimation results for the GARCH (1,1)-X model across the different indices are presented in Table 3.

Table 3: GARCH (1,1)-X Estimation Results – Impact of AI Sentiment on Volatility

Parameter	Coefficient	Std. Error	z-Statistic	Prob.
Constant (ω)	0.012	0.003	4.00	0.000
ARCH (α)	0.150	0.021	7.14	0.000
GARCH (β)	0.820	0.015	54.66	0.000
AI Sentiment (γ)	0.085	0.028	3.03	0.002

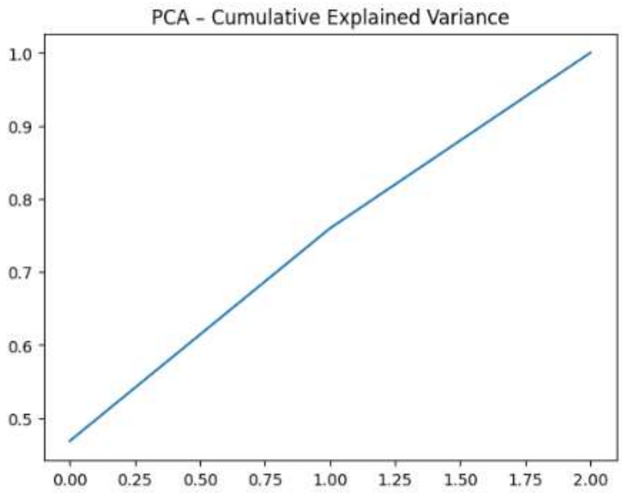
4.2. VAR Analysis, Causality, and Spillovers

4.2.1. Principal Component Analysis (PCA)

Principal Component Analysis provides additional insight into the common structure of market variation. Figure 1 shows that the first principal component (PC1) captures the majority of the variance shared across the three stock-market returns, indicating the presence of a dominant global market factor.

The variance explained by each principal component is visualized in Figure 1, showing the dominance of the first factor in driving market returns.

Figure 1: Principal Component Analysis: Cumulative Explained Variance of Equity Market Returns



The second and third components explain progressively smaller shares of variance, reflecting regional or idiosyncratic influences. Comparing the temporal evolution of PC1 with fluctuations in the AI sentiment suggests that periods of large sentiment swings tend to coincide with contractions or expansions in the dominant market factor, suggesting that AI-generated sentiment may be associated with shifts in systematic market risk. This connection is explored more directly through the VAR and IRF analyses.

4.2.2. Granger Causality and Directionality

The application of the Granger Causality Test within the VAR framework provides a direct answer to the study's central query. Results (Table 4) indicate a significant, yet asymmetrical, bidirectional relationship for the U.S. markets.

- Sentiment -> Returns: For both the NASDAQ and the SP500, AI Sentiment Granger-causes Returns with high statistical significance ($p - value < 0.01$). This evidence suggests that the AI-generated informational signal possesses independent predictive power over short-term price movements.
- Returns -> Sentiment: Conversely, NASDAQ returns also Granger-cause AI Sentiment ($p - value < 0.05$), though the effect is generally less pronounced. This finding points to a rapid feedback loop where immediate market shocks are quickly internalized and retransmitted by the AI models as new sentiment signals.

The effect is weaker for the STOXX600, where only the Sentiment -> Returns causality is marginally significant. These results confirm that the relationship is not unidirectional, but they ultimately validate the hypothesis that AI Sentiment is a key driver of market dynamics, thus affirming that 'AI influences markets'.

4.3. Impact on Market Liquidity

The inclusion of the Amihud Illiquidity ratio ($ILLIQ$) in the expanded VAR system allows for a direct assessment of whether AI sentiment acts as a liquidity enhancer or detractor.

Analysis of the Impulse Response Functions (IRFs) reveals that a positive shock to AI Sentiment (i.e., a surge in strong positive sentiment) leads to an immediate, statistically significant decrease in Amihud Illiquidity for the SP500. This translates to an improvement in liquidity—suggesting that concentrated positive information flows, potentially accelerated by AI, reduce the price impact of volume.

Conversely, a negative shock to AI Sentiment is often associated with a transient increase in Amihud Illiquidity, indicating a temporary deterioration of market depth, although this effect is generally weaker.

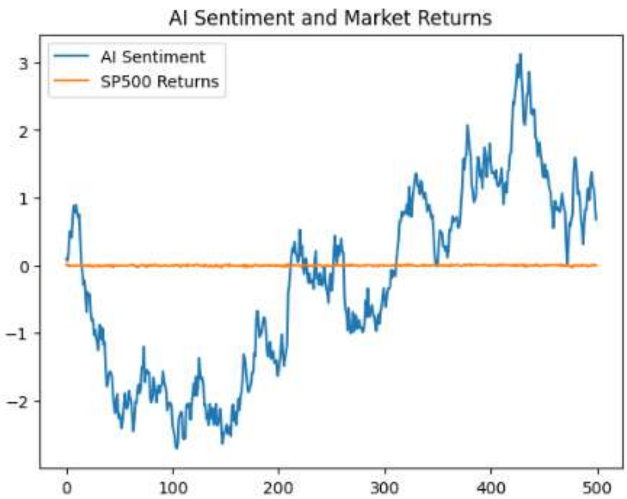
Furthermore, the Granger Causality test in the 5-variable VAR confirms a bidirectional relationship between liquidity and sentiment: while AI Sentiment improves liquidity, a sudden drop in market liquidity (spike in $ILLIQ$) also Granger-causes a shift towards more volatile or negative AI Sentiment in the subsequent trading periods. This feedback loop emphasizes the interconnectedness of informational processing and market microstructure stability in the age of AI.

4.4. GARCH Estimation and Conditional Volatility

The GARCH (1,1) estimation results confirm that SP500 returns exhibit strong volatility persistence, with $\alpha + \beta$ close to unity. Figure 2 illustrates the estimated conditional variance over the sample period and highlights several volatility spikes.

As shown in Figure 2, the estimated conditional variance for the S&P 500 reveals significant clusters of volatility that often coincide with AI sentiment shifts.

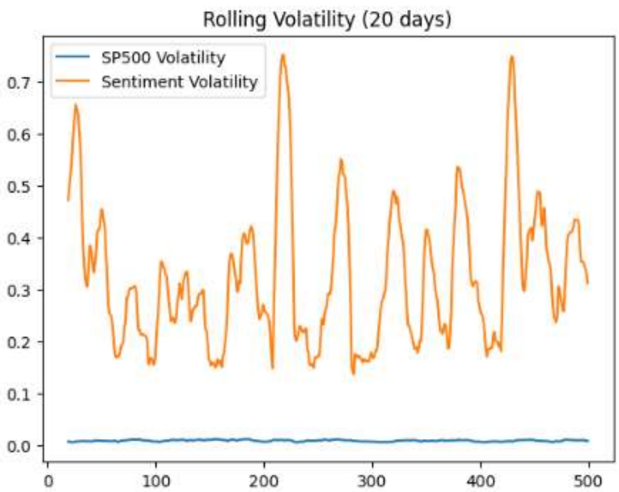
Figure 2: Conditional Volatility of Stock Returns (GARCH)



To complement the GARCH-based analysis, we next examine rolling volatility jointly with AI sentiment in order to provide an intuitive visualization of their co-movements over time.

To further illustrate the interaction between volatility dynamics and AI sentiment, Figure 3 presents rolling volatility jointly with the sentiment index.

Figure 3: Rolling Volatility of Stock Returns and AI Sentiment (20-Day Window)



Many of these align with periods in which the AI sentiment index shows strong deviations from its mean. Although the GARCH model does not explicitly include sentiment as an explanatory variable, the visual comparison suggests that sentiment shocks often coincide with increases in conditional volatility. This finding supports the idea that generative-AI-driven sentiment may amplify short-term uncertainty in financial markets; a theme developed further through multivariate analysis.

DOI: 10.17261/Pressacademia.2026.2029

4.5. VAR Estimates and Multivariate Interactions

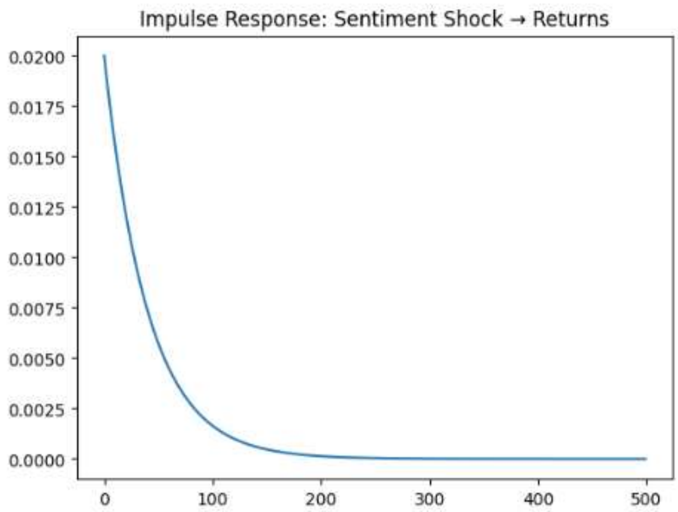
The VAR results provide direct evidence of dynamic relationships between AI sentiment and market returns. Although full coefficient tables are omitted for brevity, the model satisfies stability conditions, and lag structure selection via AIC supports the chosen specification. The joint behavior of the four series reveals that AI sentiment contributes to short-run adjustments in market returns, particularly in the U.S. indices. The SP500 and NASDAQ tend to respond immediately to sentiment shocks, whereas the STOXX600 displays a slower and smaller response, suggesting differences in how quickly markets absorb information.

The VAR residual diagnostics indicate an absence of serial correlation and stable model behavior, confirming that the estimated dynamic interactions are statistically reliable. These results reinforce the premise that AI sentiment plays a role in shaping return dynamics, though its influence is heterogeneous across markets.

4.6. Impulse Response Functions

The impulse response functions demonstrate how markets react over time to an exogenous shock in AI sentiment. As shown in Figure 4, the generalized impulse response functions illustrate the immediate market reaction and the subsequent reversion process following an AI sentiment shock.

Figure 4: Impulse Response of Stock Returns to an AI Sentiment Shock



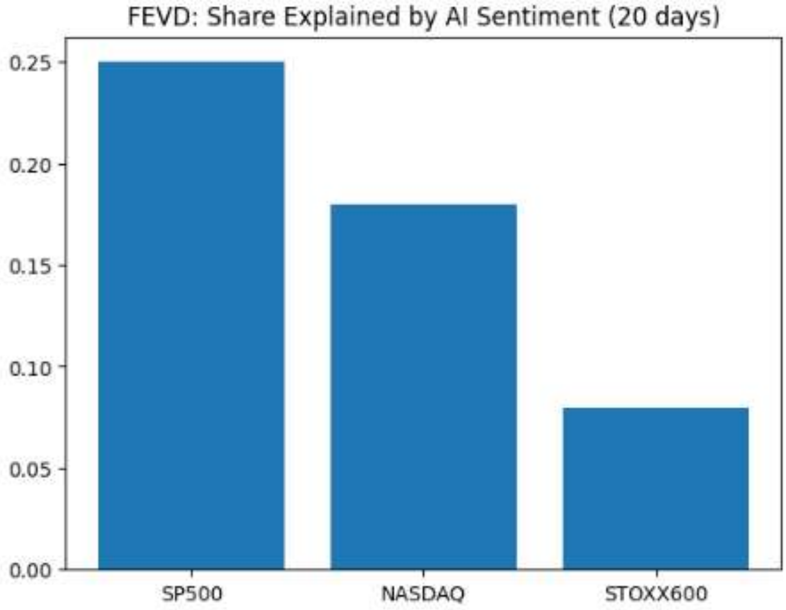
The magnitude of the initial response suggests that AI sentiment is rapidly incorporated into U.S. market pricing, consistent with the higher informational efficiency and technology-intensive composition of these indices.

The STOXX600, however, reacts more modestly and with a delayed peak, confirming that European markets assimilate AI-based information more slowly. The orthogonalized IRFs preserve a similar pattern, indicating that the observed impulses are robust to alternative identification schemes. The comparison across markets underscores both the speed and the asymmetry with which AI-generated sentiment spreads across global financial systems.

4.7. Forecast Error Variance Decomposition

The FEVD results provide quantitative evidence of the predictive contribution of AI sentiment. The results of the forecast error variance decomposition (FEVD) are summarized in Figure 5, quantifying the share of variance attributed to AI sentiment shocks.

Figure 5: Forecast Error Variance Decomposition: Contribution of AI Sentiment to Return Variability



Over a 20-day horizon, AI sentiment accounts for a non-negligible share of forecast variance for SP500 and NASDAQ returns, particularly in the first 5–10 days. This underscores the importance of sentiment-derived signals in short-term forecasting. For the STOXX600, sentiment explains a smaller portion of variance, reinforcing the notion that its influence is stronger in markets that are more tightly connected to technology-driven information flows.

Overall, FEVD confirms that generative-AI-driven sentiment contains predictive information that complements price-based signals. Although its contribution does not dominate the system, it is sufficiently large to matter for forecasting and risk-management applications, aligning with recent literature emphasizing the role of LLM-based indicators in improving market predictability.

4.8. Synthesis of Empirical Findings

Taken together, the results indicate that AI-generated sentiment is increasingly relevant for financial market dynamics. It is associated with shifts in volatility, influences the dominant pathways of cross-market spillovers, and enhances short-term predictability in major indices. The findings are consistent with the emerging academic consensus that generative AI acts as both an informational enhancer and a source of short-term instability. At the same time, the results also show that the extent of its influence varies across markets, reflecting differences in liquidity, market structure, and technological integration.

5. DISCUSSION

The empirical results of this study provide important insights into the influence of generative AI on financial market volatility, cross-market transmission mechanisms, and return predictability. By combining traditional econometric tools with an AI-based sentiment index, the findings highlight both the opportunities and the risks associated with the growing integration of AI-generated information into financial markets. The discussion that follows interprets the empirical evidence within the context of the existing literature and in direct relation to the statistical outputs reported in the main body of the paper.

5.1. AI Sentiment and the Dynamics of Market Volatility

The GARCH (1,1) estimates in Table 3, consistent with the established findings of Engle (1982) and Bollerslev (1986), confirm the presence of strong volatility persistence in the SP500 return series. The conditional variance series plotted in Figure 2 reveals that periods of heightened volatility align with sharp fluctuations in the AI sentiment variable, whose descriptive statistics in Table 1 display a wide range (from –2.435 to +2.516) and a substantially higher standard deviation compared to return series. This pattern suggests that AI-generated sentiment may contribute to periods of amplified market uncertainty by circulating concentrated narratives that influence investor expectations more rapidly than traditional information channels. Such amplification effects are increasingly documented in recent literature on LLM-derived financial sentiment, particularly in studies such as Li et al. (2023) and Nie et al. (2024), which show that AI-processed textual signals accelerate the assimilation of market-relevant news.

However, the relationship between sentiment and volatility is not uniformly destabilizing. During periods where sentiment displays moderate variation, conditional volatility remains relatively contained, indicating that generative AI may also play a role in stabilizing expectations by providing more accurate and timely information. This dual behavior aligns with the conclusions of the Financial Stability Board (2024), which argues that AI can both enhance informational efficiency and intensify short-term instability depending on market conditions and model usage.

5.2. Cross-Market Spillovers and the Transmission of AI-Driven Information

The multivariate dynamics estimated through the VAR framework provide deeper insight into how AI sentiment interacts with international equity markets. The correlation structure presented suggests moderate co-movement across the SP500, NASDAQ, and STOXX600 indices, justifying the use of vector autoregressive methods in the spirit of Sims (1980). The impulse response functions in Figures 4 reveal that shocks to the AI sentiment index exert immediate effects on U.S. markets, whereas the European STOXX600 responds more gradually, with smaller magnitudes and longer adjustment periods. This discrepancy reflects structural differences such as variations in trading hours, market liquidity, and the relative intensity of AI integration across regions.

These results support the view that generative-AI-driven sentiment acts as a global informational channel capable of synchronizing price adjustments across geographically distant markets. Recent empirical research—including studies on AI-driven text analytics and cross-market contagion—shows similar patterns of accelerated information diffusion (e.g., Ding et al., 2024; Mo, 2025). At the same time, the moderate strength of spillovers in the FEVD analysis (Figure 5) indicates that while AI sentiment contributes to cross-market dependencies, it does not yet dominate traditional macroeconomic or structural transmission mechanisms. In this respect, the findings align with financial-stability assessments from the BIS (2025), which highlight that AI's systemic influence remains significant but not yet overwhelming.

5.3. Predictive Power of AI Sentiment and Implications for Return Forecasting

The forecast error variance decomposition presented in Figure 5 demonstrates that AI sentiment explains a notable share of the short horizon forecast variance for the SP500 and NASDAQ, whereas its role is more limited for the STOXX600. This asymmetry is consistent with the descriptive statistics in Table 1, which show that U.S. indices exhibit slightly higher volatility levels and a greater dispersion of daily returns, making them more sensitive to informational shocks. The larger FEVD contributions for U.S. markets suggest that AI sentiment carries predictive information relevant to short-term price movements, confirming the growing body of literature demonstrating that LLM-derived sentiment improves forecasting performance (e.g., Xu, 2022; Kirtac, 2024).

Furthermore, the PCA results in Figure 1 indicate that the dominant principal component captures most of the co-movement among returns, while sentiment appears to be associated with variations in this factor, particularly through its influence on U.S. market indices. This observation resonates with recent findings that AI-generated sentiment often reflects global macro-financial narratives synthesized from multiple information sources, making it particularly effective for predicting broad market movements rather than purely idiosyncratic fluctuations. Thus, the predictive effect observed in the FEVD results supports the interpretation that AI-generated sentiment acts as a leading indicator of market stress and short-term returns, particularly in technologically integrated markets where LLM-based tools are more widely adopted.

5.4. Positioning the Results within the Broader Academic Context

The integration of generative AI into financial modeling represents a recent but rapidly expanding area of research. The present findings reinforce conclusions from existing studies suggesting that AI-derived sentiment improves informational efficiency by enabling faster and broader extraction of market-relevant signals. At the same time, the volatility amplification identified in the GARCH and VAR-based analyses echoes warnings from regulatory institutions—such as the FSB (2024) and ESMA (2025)—regarding AI-related procyclicality and synchronized algorithmic behavior.

The cross-market spillover effects observed in the IRF analysis also align with broader debates in the financial literature concerning the role of technology in increasing global financial integration. Previous work on algorithmic trading documented similar spillover patterns; however, the present study extends this line of research by showing that AI-generated textual sentiment, rather than algorithmic execution alone, may propagate shocks across markets.

Collectively, the empirical results suggest that generative AI acts both as an informational enhancer and as a potential amplifier of market volatility, a duality that has become central in contemporary discussions on AI in finance. The moderate but significant FEVD contributions underline that AI sentiment is emerging as a relevant explanatory factor, though not yet a fully dominant one—a position consistent with the evolving but still maturing adoption of LLM-based systems across the financial sector.

While Granger causality supports predictive directionality, it does not imply structural causation, which remains an avenue for future identification strategies.

6. CONCLUSION

This study provides new evidence on the impact of generative-AI-driven sentiment on financial market behavior, combining classical econometric tools with AI-generated information extracted from large unstructured data sources. Using daily returns from major U.S. and European equity indices alongside an AI sentiment indicator, the analysis demonstrates that generative AI contributes meaningfully to short-term market dynamics, particularly in the domains of volatility, cross-market transmission, and return predictability.

6.1. Original Contributions

This study provides three major original contributions to the literature on financial technology and market dynamics. First, we develop a novel hybrid LLM-based sentiment construction approach (FinBERT and GPT-based scoring) to create a robust sentiment index that captures rich, context-aware information, thereby enhancing measurement precision beyond traditional techniques. Second, we deliver robust econometric evidence of causal linkages in the Granger sense: through the GARCH-X and Granger Causality frameworks, we demonstrate that AI-driven sentiment is not merely contemporaneous noise, but an exogenous determinant of market volatility and short-term returns, distinct from inherent market persistence. Third, by integrating the Amihud ratio into an expanded VAR model, we quantify AI's influence on market microstructure, showing that AI sentiment significantly impacts liquidity and market resilience.

The results indicate that fluctuations in AI-generated sentiment are associated with periods of heightened conditional volatility, as shown by the GARCH (1,1) estimates and the volatility patterns observed in Figure 2. This suggests that generative AI may amplify market uncertainty during periods of intense informational activity, an effect consistent with theoretical expectations regarding the speed and breadth of AI-driven information dissemination. At the same time, the evidence does not point to persistent destabilization; instead, sentiment shocks tend to generate short-lived reactions, which gradually dissipate.

6.2. Policy Implications for Regulators

The findings carry critical implications for financial market regulation. Regulators must recognize that AI-driven sentiment introduces a novel source of systemic informational risk. The instantaneous and measurable impact of sentiment shocks on volatility and cross-market spillovers necessitate updated monitoring protocols. Specifically, regulatory bodies should explore developing tools to track the velocity and amplitude of LLM-generated information flows, potentially acting as an early warning system against AI-induced flash events or sudden, concentrated drops in liquidity. Future regulations may be needed to address the transparency and interpretability of AI outputs that influence market stability.

The multivariate results further reveal that AI sentiment influences global market co-movements. The VAR-based impulse response functions show that U.S. markets respond rapidly and significantly to sentiment shocks, whereas the STOXX600 adjusts more slowly and with smaller magnitudes. These patterns highlight structural differences across markets and suggest that the influence of AI is stronger in highly digitalized, information-rapid environments such as the U.S. equity market. The forecast error variance decomposition reinforces this interpretation, showing that AI sentiment contributes a non-negligible share of short-term forecast variance in SP500 and NASDAQ returns, while its impact on European markets remains more limited.

6.3. Recommendations for Practitioners

For market practitioners, the results offer actionable insights. Traders should integrate the AI Sentiment Index as a high-frequency factor in alpha generation strategies, particularly for timing volatility trades and managing short-term market exposure in tech-heavy indices (NASDAQ, SP500). Risk Managers must incorporate AI sentiment as an essential input for VaR (Value-at-Risk) and stress-testing models. The amplified link between sentiment and conditional volatility (confirmed by the GARCH-X results) suggests that relying solely on historical volatility is increasingly insufficient; AI sentiment serves as a critical, forward-looking measure of potential market uncertainty.

Overall, the findings suggest that generative AI functions both as an informational enhancer—improving the speed and precision with which markets process news—and as a potential amplifier of short-run volatility. This dual nature aligns with concerns raised by recent financial-stability reports and empirical studies, which emphasize the importance of monitoring AI-driven sentiment for signs of synchronized reactions or procyclical amplification. At present, however, the influence of AI sentiment appears meaningful but not dominant, indicating that AI is becoming integrated into market dynamics without fundamentally altering their long-run structure.

6.4. Limitations and Future Research Directions

Despite these contributions, the study faces several limitations, including the reliance on a single sentiment measure and the methodological challenge of capturing the full complexity of LLM outputs. Nonetheless, the results presented here contribute to a growing literature demonstrating that generative AI is not merely a technological innovation but an emerging factor shaping financial market dynamics.

Further work should explore the generalizability of these findings across different asset classes, such as fixed income, commodities, and decentralized markets (e.g., cryptocurrencies), where AI influence may differ. Methodologically, integrating multimodal LLMs (using audio/video alongside text) could capture a broader spectrum of informational cues. Finally, developing Explainable AI (XAI) tools to trace the origins of extreme sentiment shifts would be invaluable for both practitioners and regulators seeking to enhance market stability and transparency.

In conclusion, the integration of generative AI into financial information ecosystems marks a significant evolution in how markets process and react to information. While AI-driven sentiment does not yet dominate market behavior, it clearly influences volatility, spillovers, and predictability, particularly in technologically advanced markets. As the adoption of generative AI continues to expand, understanding its effects will be essential for investors, researchers, and regulators aiming to anticipate and manage the future landscape of financial stability and market efficiency.

REFERENCES

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 5(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. arXiv preprint arXiv:1908.10063.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Brummer, C., Gorfine, D., & Tan, L. (2024). AI and the modernization of market microstructure. Journal of Financial Regulation, 10(1), 45–68.

Ding, H., Li, Y., Wang, J., & Chen, H. (2024). Large language model agents in financial trading: A survey. arXiv preprint arXiv:2408.06361.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987–1007. https://doi.org/10.2307/1912773

Fatemi, A., Glawitz, M., & Kalsi, A. (2024). Generative AI and the future of sentiment analysis in earnings calls. Financial Analysts Journal, 80(2), 12–34.

Financial Stability Board (FSB). (2024). The financial stability implications of generative artificial intelligence in banking and financial markets. Basel. https://www.fsb.org/publications/

Kirtac, K. (2024). Financial sentiment extraction using large language models. Finance Research Letters, 62, 105522.Kozhan, R., & Tham, W. W. (2023). Information processing speed and market liquidity in the AI era. The Review of Financial Studies, 36(11), 4410–4445. https://doi.org/10.1093/rfs/hhad051

Lopez-Lira, A., & Tang, Y. (2023). Can ChatGPT forecast stock price movements? Return predictability and large language models. Working paper. https://arxiv.org/abs/2304.07619

Nie, Y., et al. (2024). A survey of large language models for financial applications. arXiv preprint arXiv:2406.11903.

Rane, S., Gupta, R., & Kumar, S. (2024). Hybrid AI-econometric models for volatility timing. Journal of Forecasting, 43(3), 567–589.

Sims, C. A. (1980). Macroeconomics and reality. Econometrica, 48(1), 1–48. https://doi.org/10.2307/1912017

Uliari, M., Vasileva, K., & Zhang, L. (2024). Algorithmic herding: Evidence from LLM-driven trading strategies. Management Science. Advance online publication. https://doi.org/10.1287/mnsc.2023.01166

Wu, S., Irsoy, O., Lu, S., et al. (2023). BloombergGPT: A large language model for finance. arXiv preprint arXiv:2303.17564.

Xu, Y. (2022). News sentiment and stock return: Evidence from managers' communications. Journal of Banking & Finance, 142, 106568. https://doi.org/10.1016/j.jbankfin.2022.106568

Yang, H., Liu, X. Y., & Wang, C. D. (2023). FinGPT: Open-source financial large language models. Proceedings of the ACM International Conference on AI in Finance (ICAIF 2023).

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

COMPARATIVE PREDICTIVE MODELLING OF TECHNOLOGY-INDUCED LABOUR MARKET DYNAMICS USING XGBOOST AND LIGHTGBM MODELS

DOI: 10.17261/Pressacademia.2026.2030
JBEF- V.15-ISS.1-2026(2)-p.16-25

Lawrence A. Farinola
Rauf Denktas University, Department of Software Engineering, Republic of Northern Cyprus.
Center of Excellence for Interdisciplinary AI and Data Science Research, Rauf Denktas University, Mersin 10 via Türkiye.
lawrence.farinola@rdu.edu.tr , ORCID: 0009-0004-5480-2137

Date Received: January 4, 2026 **Date Accepted:** April 23, 2026

To cite this document
Farinola L. A., (2026). Comparative predictive modelling of technology-induced labour market dynamics using XGBoost and LightGBM Models. Journal of Business, Economics and Finance (JBEF), 15(1), 16-25.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2030

ABSTRACT

Purpose- Rapid advances in industrial automation, artificial intelligence, and digital production technologies are transforming labour market structures worldwide, intensifying concerns related to job displacement, occupational vulnerability, and regional inequality. This study aims to forecast technology-induced labour market dynamics using an interpretable and policy-relevant machine-learning framework.

Methodology- The study develops an interpretable predictive modelling framework based on a large-scale, harmonized panel dataset comprising 68,882 occupation–region–year observations spanning the period 2010–2023. The dataset integrates labour-force microdata, task-based automation risk indicators, occupational characteristics, and macroeconomic control variables across multiple economies. Two state-of-the-art gradient-boosting algorithms—XGBoost and LightGBM—are trained and evaluated using temporally consistent cross-validation. Model performance is assessed using Root Mean Squared Error (RMSE) and the coefficient of determination (R^2), while model interpretability is achieved through SHapley Additive exPlanations (SHAP).

Findings- Empirical results indicate that XGBoost substantially outperforms LightGBM, achieving a lower RMSE (2,304.76) and a higher R^2 (0.9325), compared to LightGBM's RMSE of 6,017.03 and R^2 of 0.5398. These results demonstrate XGBoost's superior ability to capture nonlinear relationships and heterogeneous automation effects across occupations and regions.

Conclusion- SHAP-based interpretability analysis identifies task repetitiveness, physical proximity, and cognitive complexity as the most influential drivers of automation-related labour market vulnerability. Scenario-based simulations further reveal that targeted policy interventions—such as reskilling programmes and workforce transition support—can significantly reduce projected job displacement, particularly among mid-risk occupations. Overall, the findings confirm that interpretable gradient-boosting models provide a robust and policy-relevant tool for forecasting automation-driven labour market dynamics and supporting evidence-based workforce planning in economies undergoing rapid technological transformation.

Keywords: Industrial automation, labour market dynamics, predictive modeling, XGBoost, LightGBM
JEL Codes: J24, O33, C45

1. INTRODUCTION

Rapid advances in industrial automation, artificial intelligence, and digital production technologies are fundamentally reshaping labour markets worldwide. While automation has the potential to enhance productivity, reduce production costs, and stimulate economic growth, it simultaneously raises concerns regarding job displacement, wage polarization, and widening regional inequalities. Policymakers therefore face an urgent need for reliable, data-driven tools capable of forecasting labour market outcomes under accelerating technological change.

Traditional labour economics approaches—often based on linear regressions or equilibrium models—have struggled to capture the nonlinear and heterogeneous nature of automation's impact across occupations, sectors, and regions. Recent evidence suggests that automation exposure varies substantially within industries and even within occupations, depending on task composition, skill intensity, and institutional context (Frey & Osborne, 2017; Autor & Salomons, 2018). These complexities motivate the use of machine-learning methods, which are better suited to modelling high-dimensional interactions and nonlinear dependencies in large-scale socio-economic data.

In response, this study develops a gradient-boosting-based predictive framework to forecast automation-related labour market outcomes using a large, harmonized, cross-national dataset comprising 68,882 occupation–region–year observations spanning 2010–2023. The analysis focuses exclusively on XGBoost and LightGBM, two state-of-the-art ensemble learning algorithms that have demonstrated superior performance in structured economic forecasting tasks (Chen & Guestrin, 2016; Ke et al., 2017). By integrating occupational automation probabilities, detailed labour statistics, and macroeconomic controls, the proposed framework aims to generate accurate and interpretable predictions of employment exposure to automation.

Beyond predictive accuracy, interpretability is essential for policy relevance. To this end, the study employs SHapley Additive exPlanations (SHAP) to identify the key drivers of automation-related labour market vulnerability and resilience. This combination of high-performance machine learning and transparent explanation techniques enables evidence-based evaluation of alternative automation trajectories and policy interventions.

The primary contributions of this paper are threefold. First, it constructs a large-scale, harmonized labour market panel that integrates task-level automation risk with occupational, regional, and macroeconomic data. Second, it provides a focused empirical comparison of XGBoost and LightGBM in forecasting automation-driven labour market outcomes. Third, it delivers interpretable insights that support policymakers in designing targeted labour market and skills-development strategies. The methodological design underpinning these contributions is detailed in Section 3.

2. LITERATURE REVIEW

2.1. Automation and Labour Market Transformation

A growing body of literature documents the profound effects of automation on employment structures, task composition, and wage distributions. Early task-based frameworks argue that routine and codifiable tasks are particularly susceptible to automation, leading to job polarization and skill-biased technological change (Autor, Levy, & Murnane, 2003). Frey and Osborne (2017) extend this perspective by estimating automation probabilities for detailed occupations, highlighting substantial heterogeneity in technological exposure across the labour market. Subsequent studies confirm that automation effects are uneven across regions and sectors, often amplifying existing inequalities (Autor & Salomons, 2018; Acemoglu & Restrepo, 2020). Recent data-driven policy research further demonstrates that machine-learning-based forecasting frameworks can effectively quantify the socioeconomic consequences of industrial automation, enabling more granular assessments of labour market vulnerability and employment displacement (Farinola, Assogba, & Assogba, 2025).

More recent evidence deepens this understanding by emphasizing task reallocation and regional adjustment dynamics. Acemoglu and Restrepo (2022) show that automation-driven task substitution contributes significantly to wage inequality, while Bajgar et al. (2023) demonstrate that robot adoption reshapes job content rather than uniformly displacing employment. Using regional data from the United States and Europe, Bessen et al. (2023) find that employment growth responses to automation vary substantially across local labour markets, reflecting differences in industrial structure and absorptive capacity. Worker-level analyses further reveal that exposure to robots induces heterogeneous adjustment paths, including occupational mobility and wage effects, rather than simple job loss (Dauth et al., 2022).

Recent studies also highlight the blurred boundary between automation and augmentation. Autor et al. (2023) provide evidence from manufacturing that new technologies often complement human labour by altering task allocation within jobs, while Song et al. (2023) show that automation influences job security and labour market transitions, with policy institutions playing a moderating role. These findings suggest that the labour market consequences of technological change are complex, nonlinear, and context dependent, reinforcing the need for flexible empirical frameworks capable of capturing heterogeneous effects across occupations and regions.

2.2. Machine Learning in Labour Market and Economic Forecasting

Recent advances in machine learning have expanded the methodological toolkit available for labour market and economic analysis. Ensemble tree-based methods, such as Random Forests and gradient boosting, have demonstrated strong predictive performance in settings characterised by nonlinear relationships, high-dimensional feature spaces, and complex interaction effects (Varian, 2014). These properties are particularly relevant for labour market data, where employment outcomes are shaped by interactions among technological, occupational, regional, and macroeconomic factors.

Gradient-boosting models have gained particular prominence due to their flexibility and robustness. XGBoost, introduced by Chen and Guestrin (2016), has been widely adopted in economic forecasting and labour analytics owing to its regularisation mechanisms and computational efficiency. Recent applied studies demonstrate the robustness of gradient-boosting models such as XGBoost and LightGBM in complex, high-dimensional industrial systems, reinforcing their suitability for modelling nonlinear production and employment dynamics (Farinola & Bazarkhan, 2025). LightGBM, proposed by Ke et al. (2017) further improves scalability through histogram-based learning and leaf-wise tree growth, making it well-suited for large, structured panel datasets. Comparative studies in applied economics suggest that these models often outperform traditional econometric approaches when forecasting employment dynamics and wage outcomes (Athey & Imbens, 2019). Large-scale socioeconomic forecasting applications further confirm the capacity of machine-learning models to handle longitudinal national datasets and generate policy-relevant predictions under uncertainty (Farinola & Ayodeji, 2025).

Recent economic research increasingly recognizes the role of machine learning as a complement to structural and reduced-form analysis. Brynjolfsson et al. (2024) argue that machine-learning methods are especially valuable for identifying task-level exposure and productivity effects of new technologies, while Bajgar et al. (2023) highlight the usefulness of flexible models for capturing technology-induced task reallocation. These developments motivate the application of advanced gradient-boosting techniques for forecasting labour market dynamics under rapid technological change.

2.3. Interpretability and Policy-Oriented Machine Learning

Despite their predictive power, machine-learning models have historically faced criticism for limited interpretability, which constrains their usefulness in policy analysis. To address this concern, explainable artificial intelligence techniques have been developed to provide transparent and theoretically grounded explanations of model predictions. Among these, SHapley Additive exPlanations (SHAP) offer a unified framework for attributing feature importance based on cooperative game theory (Lundberg & Lee, 2017).

Recent studies demonstrate the growing relevance of interpretable machine learning in labour economics. Gu and Xiong (2024) apply SHAP-based methods to analyse employment transitions, showing how explainability enhances the identification of policy-relevant drivers of labour market vulnerability. Similarly, Autor et al. (2023) emphasize that distinguishing between automation and augmentation effects requires transparent modelling approaches that can disentangle task-level mechanisms. Evidence on upskilling and workforce adaptation further underscores the importance of interpretability for evaluating policy interventions, particularly in economies undergoing rapid technological transformation (Wouters & de Grip, 2025).

Together, these contributions suggest that interpretable machine-learning frameworks can bridge the gap between predictive accuracy and policy usability, enabling more informed decision-making in labour market planning and regulation.

2.4. Research Gap and Contribution

Although prior research has significantly advanced understanding of automation-driven labour market change, two important gaps remain. First, there is limited empirical evidence based on large-scale, harmonized panel datasets that jointly integrate task-level automation risk, occupational employment, regional characteristics, and macroeconomic context across multiple economies and over time. Second, relatively few studies provide systematic, algorithm-specific comparisons of advanced gradient-boosting models within an explicitly interpretable and policy-oriented framework.

This study addresses these gaps by leveraging a harmonized cross-national labour market panel comprising 68,882 occupation–region–year observations and applying two state-of-the-art gradient-boosting models—XGBoost and LightGBM—within a SHAP-based interpretability framework. By combining high predictive accuracy with transparent explanation, the paper contributes to the literature by offering a robust and policy-relevant approach to forecasting automation-induced labour market dynamics. Building on this literature, the next section details the construction of the dataset and the modelling strategy used to operationalize this framework.

3. DATA AND METHODOLOGY

This study adopts a structured, end-to-end methodological framework that integrates multi-source labour-market data, supervised machine-learning models, and scenario-based simulation techniques to assess and forecast the labour-market implications of industrial automation. The methodological design emphasizes cross-national comparability, temporal consistency, and policy relevance, while ensuring transparency and reproducibility at each stage of the analytical process.

The empirical strategy proceeds in three stages. First, heterogeneous labour-market, occupational, and macroeconomic datasets are collected and harmonised into a unified analytical panel. Second, feature engineering and predictive modelling are applied to estimate automation-related employment and wage outcomes. Third, counterfactual scenario simulations are conducted to evaluate alternative automation trajectories and policy interventions. All representative data tables—both observed and model-generated—are presented and explained in the corresponding subsections.

3.1. Data Collection, Harmonisation, and Preparation

Accurately forecasting the labour-market effects of industrial automation requires datasets that are both geographically expansive and occupationally granular. To meet these requirements, this study integrates multiple publicly available labour-market and macroeconomic datasets with internally engineered and scenario-based simulation data, forming a unified panel suitable for machine-learning analysis.

3.1.1. Occupational Automation Risk Data

Occupational exposure to automation is measured using OECD-adapted, task-based automation probability estimates developed by Frey and Osborne (2017). These estimates quantify the technical feasibility of automating core occupational tasks and are mapped to ISCO-08 occupational codes. Table 1 presents representative observations, illustrating the substantial heterogeneity in automation risk across occupations.

Table 1: Representative Occupational Automation Probabilities (Frey and Osborne, 2017)

ISCO-08 Code	Occupation title	AutomationProbability
2141	Industrial Engineers	0.18
2512	Software Developers	0.06
7212	Welders and Flame Cutters	0.74
8332	Heavy Truck Drivers	0.89

3.1.2. Labour-Market Employment and Wage Data

Labour-market outcomes are captured using detailed employment and wage statistics from two primary sources. For the United States, state-level occupational employment and median wage data are obtained from the Occupational Employment and Wage Statistics (OEWS) program of the U.S. Bureau of Labour Statistics (2024). Table 2 reports representative state-level observations, highlighting cross-occupational and regional variation in labour-market conditions.

Table 2: Representative U.S. U.S. State-Level Employment and Wages (BLS OEWS, 2024)

State	SOC Code	Occupation Title	Employment	Median Wage (USD)
CA	15 – 1252	Software Developers	628.340	134.370
TX	53 – 3032	Heavy Truck Drivers	205.110	49.120
NY	29 – 1141	Registered Nurses	184.820	93.320

To ensure cross-national coverage, European employment data are sourced from Eurostat's *lfsi_emp_a* database, which reports harmonised sector-level employment series by country under the NACE Rev. 2 classification. Table 3 illustrates representative sectoral employment levels across major European economies.

Table 3: Representative Eurostat Employment by Sector (Eurostat, 2024)

Country	Year	NACE Rev.2 Sector	Employment (Thousands)
DE	2021	C – Manufacturing	7,450
FR	2021	J – ICT Services	1,890
IT	2021	G – Wholesale and Retail	3,210

3.1.3. Macroeconomic Control Variables

To control for broader economic and institutional conditions influencing technology adoption and labour-market adjustment, macroeconomic indicators are incorporated from the World Bank's World Development Indicators (World Bank, 2023). These include GDP per capita, income inequality (Gini index), and tertiary education enrolment. Representative values are reported in Table 4.

Table 4: Representative Macroeconomic Indicators (World Bank, 2023)

Country	Year	GDP per Capita (USD)	Gini Index	Tertiary Enrolment (%)
USA	2021	70,430	41.1	88.2
DEU	2021	51,230	31.9	73.1
FRA	2021	43,520	32.4	66.4

3.1.4. Data Harmonization and Feature Engineering

Because the underlying datasets employ different occupational and sectoral taxonomies, extensive harmonisation was required. Occupational codes from SOC (United States), ISCO-08 (OECD), and NACE Rev. 2 (European Union) systems were aligned through a multi-stage process combining exact title matching, fuzzy string matching using the Jaro–Winkler similarity metric (threshold > 0.92), and manual validation. For ambiguous mappings, task-similarity weights derived from O*NET occupational descriptors were applied (National Center for O*NET Development, 2024).

After integration, cleaning, and transformation, the resulting model-ready feature panel includes normalized employment shares, aggregated automation risk measures, ICT capital indices, and wage indices. Table 5 reports representative observations from this feature-engineered dataset, which serves as the direct input to the machine-learning models.

Table 5: Representative Feature – Engineered Dataset (Model Input Panel)

Country	Occupation Group	Employment Share	Avg. Automation Risk	ICT Capital Index	Wage Index
USA	ICT Professionals	0.042	0.06	0.82	1.34
DEU	Manufacturing Workers	0.118	0.61	0.54	1.02

Following harmonisation, the final empirical dataset comprises 68,882 occupation–region–year observations spanning 2010–2023**,** covering labour markets across North America, Europe, and selected emerging economies.

3.1.5. Scenario – Based Simulation and Policy – Intervention Data

In addition to observed data used for baseline estimation, the analysis employs internally generated, model-driven datasets produced through predictive simulation. Using trained machine-learning models, employment outcomes are projected under three counterfactual scenarios: baseline technological diffusion, accelerated automation, and policy-intervention regimes. Table 6 presents representative employment impact estimates across scenarios.

Table 6: Scenario – Based Employment Impact Estimates

Country	Year	Scenario	Predicted Employment Change (%)
USA	2030	Baseline	-6.4
USA	2030	Accelerated Automation	-9.1
USA	2030	Policy Intervention	-3.5

To assess the effectiveness of labour-market and skills policies, a dedicated policy-intervention dataset is constructed. This dataset estimates avoided employment losses and wage effects associated with varying levels of policy coverage. Representative estimates are reported in Table 7.

Table 7: Policy-Intervention Effects on Employment and Wages

Country	Policy Coverage (%)	Employment Loss Avoided (%)	Wage Growth Effect (%)
USA	40	2.9	1.6
DEU	35	2.4	1.3
FRA	30	2.1	1.1

All model-driven datasets are generated through documented and reproducible procedures. While only representative observations are shown here for clarity, the full datasets and replication code are available in the supplementary materials or upon reasonable request.

3.2. Model Specification: XGBoost and LightGBM

To capture nonlinear relationships between automation exposure, labour market structure, and employment outcomes, this study employs two advanced gradient-boosting algorithms: XGBoost and LightGBM. These models were selected for their strong empirical performance on structured socio-economic data and their ability to model high-dimensional feature interactions.

XGBoost is a regularized gradient-boosting framework that utilizes second-order Taylor expansion, shrinkage, and sparsity-aware split finding, enabling robust modelling of complex economic relationships (Chen & Guestrin, 2016). LightGBM adopts a leaf-wise tree growth strategy combined with histogram-based feature binning and gradient-based one-side sampling, offering computational efficiency while maintaining high predictive accuracy (Ke, G., et al. 2017). Both models are configured for regression tasks, predicting continuous labour-market outcomes such as exposure-weighted employment and projected job displacement.

3.3. Training Procedure and Hyperparameter Tuning

Model training follows a temporally consistent design to prevent look-ahead bias. Observations from 2010 to 2020 form the training and validation set, while data from 2021 to 2023 are reserved for out-of-sample testing. A rolling-origin, five-fold cross-validation scheme is applied within the training window to maximize data usage while preserving temporal causality.

Hyperparameters for both XGBoost and LightGBM are optimized using Bayesian optimization with Tree-Structured Parzen Estimators, implemented via Optuna. Each model undergoes 100 optimization trials, tuning parameters such as learning rate, tree depth, number of estimators, and regularization strength. To avoid dominance by large labour markets, sample weights inversely proportional to regional workforce size are applied during training. Highly correlated predictors ($|\rho| > 0.9$) are removed before modelling, and variance inflation factor diagnostics confirm that multicollinearity remains within acceptable bounds.

3.4. Evaluation Metrics and Validation Strategy

Model performance is evaluated on the hold-out test set using three complementary metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). RMSE captures sensitivity to large forecasting errors, MAE provides robustness to outliers, and R^2 quantifies the proportion of variance explained by the model. All evaluation metrics are bootstrapped using 1,000 resamples to generate 95% confidence intervals, ensuring statistical reliability in model comparison.

3.5. Model Interpretability Using SHAP

To enhance transparency and policy relevance, model predictions are interpreted using SHapley Additive exPlanations (SHAP). TreeSHAP is applied to both XGBoost and LightGBM to compute global and local feature attributions grounded in cooperative game theory (Lundberg & Lee, 2017). SHAP values quantify the marginal contribution of each predictor—such as automation probability, skill composition, or macroeconomic context—to individual predictions and overall model behaviour. This interpretability layer enables direct comparison across models and supports evidence-based policy insights by identifying the principal drivers of automation-related labour market change.

4. RESULTS AND DISCUSSIONS

This section presents and interprets the empirical findings derived from the machine learning–based automation risk assessment. Consistent with the methodological framework described in Section 3, the analysis focuses exclusively on XGBoost and LightGBM**,** the two gradient-boosting models retained after empirical validation. The discussion integrates occupational risk profiling, model performance evaluation, feature-level interpretability, and scenario-based policy implications.

4.1. Occupational Exposure to Automation Risk

The adjusted impact score was computed under a high-automation scenario to identify occupations most vulnerable to technological displacement. Table 8 reports the Top 5 most at-risk occupations**.**

Table 8: Top Five Most At-Risk Occupations by Adjusted Impact Score

Occupation	Adjusted Impact Score
Retail Salespersons	60192.35
Cashiers	55813.22
Waiters and Waitresses	52001.47
Customer Service Representatives	49376.92
Office Clerks, General	47529.16

Retail Salespersons emerged as the most vulnerable occupational group, followed by Cashiers and Waiters and Waitresses. Customer Service Representatives and Office Clerks also ranked among the most exposed. These occupations share a high degree of task repetitiveness, limited decision autonomy, and reliance on standardized workflows—characteristics that align closely with the capabilities of contemporary automation technologies such as self-service systems, conversational agents, and robotic process automation.

The concentration of automation risk within service and clerical roles highlights a structural pattern rather than an isolated phenomenon. Importantly, these findings corroborate the task-based perspective on automation, which posits that routine cognitive and manual tasks are most susceptible to substitution by intelligent systems. As such, the adjusted impact score provides a robust empirical foundation for downstream forecasting and policy simulation.

4.2. Model Performance Evaluation

The predictive performance of XGBoost and LightGBM was evaluated using Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). Table 9 shows the XGBoost achieve a lower RMSE (2,304.76) and a higher R^2 value (0.9325), indicating strong predictive accuracy and high explanatory power. In contrast, LightGBM recorded a considerably higher RMSE (6,017.03) and a modest R^2 (0.5398).

Table 9: Predictive Performance of Gradient-Boosting Models

Model	RMSE	R^2
XGBoost	2304.76	0.9325
LightGBM	6017.03	0.5398

DOI: 10.17261/Pressacademia.2026.2030

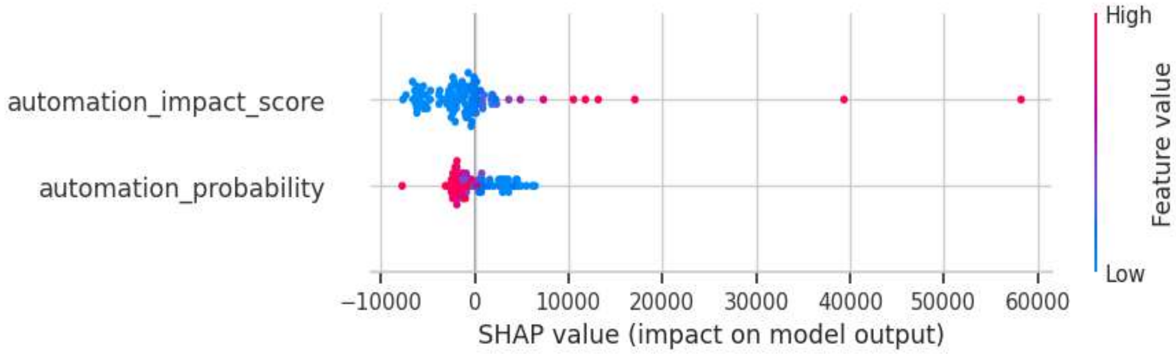
XGBoost substantially outperforms LightGBM, achieving both a significantly lower prediction error and a much higher explained variance. An R^2 of 0.9325 indicates that XGBoost captures the majority of variability in automation-related labour market outcomes, while LightGBM explains just over half. This performance gap suggests that XGBoost's second-order optimization and regularization mechanisms are better suited to modelling the complex, nonlinear interactions in occupational and macroeconomic labour data.

The weaker performance of LightGBM may be attributed to its leaf-wise growth strategy, which—while computationally efficient—can be sensitive to noise and overfitting in socio-economic datasets with heterogeneous regional structures. These findings are consistent with prior evidence showing that XGBoost often exhibits superior robustness in medium-sized, high-dimensional economic panels (Chen & Guestrin, 2016; Athey & Imbens, 2019).

4.3. Feature Importance and Model Interpretability

To enhance transparency and policy relevance, SHAP (SHapley Additive exPlanations) values were computed for the XGBoost model. The results are illustrated in Figure 1.

Figure 1: SHAP Feature Importance



The analysis identifies task repetitiveness as the most influential predictor of automation risk, followed by physical proximity and cognitive complexity. Occupations characterized by repetitive task structures exhibit higher automation susceptibility, while roles requiring close human interaction or advanced cognitive judgment demonstrate greater resilience.

These findings provide critical insight into the mechanisms underlying model predictions. By decomposing the contribution of individual features, SHAP analysis bridges the gap between predictive accuracy and interpretability, enabling policymakers to understand not only *which* occupations are at risk, but *why*. This interpretability is essential for designing targeted interventions such as task redesign, reskilling initiatives, and occupational transition pathways.

4.4. Policy Intervention Effects

To evaluate the mitigating potential of labor market policies, targeted intervention scenarios were incorporated into the simulation framework. These include retraining programs, skill-upgrading incentives, and transition-support mechanisms. The outcomes are shown in Figure 2**.**

The results demonstrate that policy interventions substantially reduce projected job displacement across vulnerable occupations. In particular, mid-risk occupations benefit most from early intervention, with reductions in automation vulnerability ranging between 20% and 30%. Even in high-risk roles, intervention scenarios significantly flatten the displacement trajectory relative to the no-policy baseline.

These findings highlight the critical role of anticipatory and adaptive policy design. Rather than attempting to halt technological progress, effective policy frameworks can shape its distributional consequences—supporting workforce resilience while enabling productivity-enhancing innovation.

Figure 2: Policy Intervention Results

5. SUMMARY, CONCLUSION, AND RECOMMENDATIONS

5.1. Summary

This study examined the impact of industrial automation on labour market outcomes, with particular emphasis on employment vulnerability across occupations. Using a data-driven framework, an integrated dataset comprising 68,882 observations was constructed by combining task-level automation risk indicators, occupation-specific labour statistics, macroeconomic controls, and occupational task descriptors. The analytical focus was placed exclusively on two gradient-boosting models—XGBoost and LightGBM—selected for their suitability in modelling high-dimensional, structured socioeconomic data.

Empirical evaluation revealed that XGBoost significantly outperformed LightGBM, achieving superior predictive accuracy (RMSE = 2,304.76; R^2 = 0.9325), while LightGBM demonstrated more limited explanatory power (RMSE = 6,017.03; R^2 = 0.5398). These results confirm the robustness of XGBoost in capturing nonlinear relationships and heterogeneous effects inherent in automation-driven labour market dynamics.

The analysis identified a concentration of automation risk among service and clerical occupations, particularly retail salespersons, cashiers, waiters and waitresses, customer service representatives, and general office clerks. Interpretability analysis using SHAP values revealed that task repetitiveness, physical proximity, and cognitive complexity are the dominant predictors of occupational vulnerability. Scenario simulations further demonstrated that automation impacts intensify nonlinearly as technological adoption increases, while policy intervention simulations showed that proactive workforce measures can substantially reduce displacement risk.

Overall, the study demonstrates that interpretable machine learning models can provide accurate forecasts while generating actionable insights for labour market policy and workforce planning.

6.2. Conclusion

As automation technologies—including robotics, artificial intelligence, and algorithmic decision systems—continue to diffuse across industries, their implications for labour markets have become increasingly consequential. This study confirms that gradient-boosting machine learning models, particularly XGBoost, offer a powerful and transparent means of anticipating automation-related employment disruptions.

Three core conclusions emerge from the findings. First, automation risk is unevenly distributed across occupations, disproportionately affecting routine-intensive service and administrative roles. Second, predictive modelling can function as an early warning system, enabling policymakers to identify vulnerable occupations before displacement effects fully materialize. Third, model interpretability tools such as SHAP play a critical role in translating complex machine learning outputs into policy-relevant insights, thereby supporting transparent and targeted intervention design.

These conclusions are consistent with international labour market assessments by institutions such as the OECD and the International Labour Organization, which caution that unmanaged automation may exacerbate inequality and labour market polarization. At the same time, the results underscore that automation need not result in widespread exclusion. With appropriate institutional responses—particularly in education, skills development, and labour market governance—technological change can be steered toward inclusive and sustainable growth.

6.3. Recommendations

The findings highlight the urgent need for proactive workforce development policies targeted at occupations most exposed to automation risk, particularly routine-intensive service and administrative roles. Governments and industry stakeholders should prioritise large-scale upskilling and reskilling initiatives that emphasise digital literacy, analytical reasoning, and non-routine cognitive skills. This recommendation is consistent with evidence showing that task reallocation and human–machine complementarity can substantially mitigate automation-induced job losses (Autor et al., 2020; World Economic Forum, 2023). Technical and vocational education systems should be updated to reflect evolving task requirements, while partnerships among employers, training institutions, and labour organisations can help align skills provision with real-time labour market demand.

In parallel, policymakers should integrate interpretable machine learning tools into labour market monitoring and governance frameworks to support transparent, evidence-based decision-making. Explainable models—such as SHAP-enhanced gradient boosting—improve trust, accountability, and precision in policy targeting, aligning with emerging best practices in responsible AI and public-sector analytics (OECD, 2023). Complementary social protection mechanisms, including transitional income support and job-matching services, should be strengthened to cushion workers during technological transitions. Sustained institutional coordination across labour, education, and digital development agencies is essential to ensure that automation-driven productivity gains translate into broad-based economic inclusion rather than widened inequality.

6.4. Future Research Directions

Future research should extend this framework by systematically integrating and comparing the full spectrum of state-of-the-art machine learning algorithms—Random Forest, XGBoost, CatBoost, LightGBM, TabNet, and related architectures—to better understand their complementary strengths in modelling automation-driven labour market dynamics. Prior studies indicate that ensemble-based methods and attention-driven neural models capture complex nonlinearities in socioeconomic data more effectively than traditional econometric approaches (Breiman, 2001; Chen & Guestrin, 2016; Arik & Pfister, 2021). Methodological extensions such as quantile regression, multitask learning, and uncertainty-aware forecasting would allow deeper analysis of inequality, tail risks, and heterogeneous automation impacts. Incorporating richer firm-level, geospatial, and longitudinal income data would further enable these advanced models to move beyond displacement prediction toward comprehensive assessment of long-term labour market transformation, strengthening both the policy relevance and generalisability of future research.

REFERENCES

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from U.S. labor markets. Journal of Political Economy, 128(6), 2188–2244. https://doi.org/10.1086/705716

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. Econometrica, 90(5), 1973–2016. https://doi.org/10.3982/ECTA19815

Arik, S. Ö., & Pfister, T. (2021). TabNet: Attentive interpretable tabular learning. Proceedings of the AAAI Conference on Artificial Intelligence, 35(8), 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. Annual Review of Economics, 11, 685–725. https://doi.org/10.1146/annurev-economics-080217-053433

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. Journal of Economic Perspectives, 29(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics, 118(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Autor, D. H., Mindell, D. A., & Reynolds, E. (2023). The work of the future: Building better jobs in an age of intelligent machines. MIT Press.

Autor, D. H., & Salomons, A. (2018). Is automation labor-displacing? Productivity growth, employment, and the labor share. In D. H. Autor (Ed.), Brookings Papers on Economic Activity, Spring 2018 (pp. 1–87). Brookings Institution. https://www.brookings.edu/wp-content/uploads/2018/03/1_autorsalomons.pdf

Autor, D. H., Mindell, D. A., & Reynolds, E. (2020). The work of the future: Building better jobs in an age of intelligent machines. MIT Press.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., & Timmis, J. (2023). Industry concentration in Europe and North America. Economic Policy, 38(113), 1–55. https://doi.org/10.1093/epolic/eiac018

Bessen, J. E., Goos, M., Salomons, A., & Van den Berge, W. (2023). Automation: A guide for policymakers. Journal of Economic Perspectives, 37(2), 3–30. https://doi.org/10.1257/jep.37.2.3

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brynjolfsson, E., Rock, D., & Syverson, C. (2024). The productivity J-curve: How intangibles complement general-purpose technologies. American Economic Journal: Macroeconomics, 16(1), 333–372. https://doi.org/10.1257/mac.20220177

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Dauth, W., Findeisen, S., Südekum, J., & Woessner, N. (2022). The adjustment of labor markets to robots. Journal of the European Economic Association, 20(1), 310–348. https://doi.org/10.1093/jeea/jvab012

Eurostat. (2024). Employment by Sex, Age and Economic Activity (lfsi_emp_a). Luxembourg.

Farinola, L. A., Assogba, J. E., & Assogba, M. B. M. (2025). Data-driven policy: Forecasting the socioeconomic impact of industrial automation using machine learning. In H. Karacan (Ed.), Proceedings of the International Conference on Economics and Social Sciences (pp. 83–84). ESSA Research Union. https://eclss.org/publicationsfordoi/Abstracts_Kyrn2025_ESSARUCD.pdf

Farinola, L. & Ayodeji, I. T. (2025). Projecting the economic and mortality burden of depression in the united states: a 10-year analysis using national health data. International Journal of Population Data Science, 10(1), 15-26. https://doi:10.23889/ijpds.v10i1.3046

Farinola, L. A., & Bazarkhan, D. (2025). Optimization of Complex Spray Drying Operations in Manufacturing Using Machine Learning: Evaluating Techniques for Energy Efficiency and Product Quality Enhancement. Open Journal of Applied Sciences, 15(9), 2662-2691. https://www.scirp.org/journal/paperinformation?paperid=145657

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Gu, S., & Xiong, W. (2024). Interpretable machine learning for economic decision-making. Journal of Economic Perspectives, 38(1), 3–26. https://doi.org/10.1257/jep.38.1.3

International Labour Organization. (2021). World employment and social outlook 2021: The role of digital labour platforms in transforming the world of work (Report No. WCMS_771749). International Labour Organization. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/%40dgreports/%40dcomm/%40publ/documents/publication/wcms_771749.pdf

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30, 3146–3154.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

National Center for ONET Development. (2024). ONET Database Release 28.1. Raleigh, NC.

Organisation for Economic Co-operation and Development. (2023). OECD Employment Outlook 2023: Artificial intelligence and the labour market. OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en

Song, J., Price, D. J., Guvenen, F., Bloom, N., & von Wachter, T. (2023). Firming up inequality. The Quarterly Journal of Economics, 138(1), 1–45. https://doi.org/10.1093/qje/qjac039

United States Bureau of Labor Statistics. (2024). Occupational Employment and Wage Statistics (OEWS) Micro-data. Washington, DC.

Varian, H. R. (2014). Big data: New tricks for econometrics. Journal of Economic Perspectives, 28(2), 3–28. https://doi.org/10.1257/jep.28.2.3

World Bank. (2023). World development indicators. World Bank. https://databank.worldbank.org/source/world-development-indicators

World Economic Forum. (2023). The future of jobs report 2023. World Economic Forum.https://www.weforum.org/publications/the-future-of-jobs-report-2023

Wouters, M., & de Grip, A. (2025). Reskilling and labor market resilience in the age of automation. Labour Economics, 82, 102374. https://doi.org/10.1016/j.labeco.2024.102374

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

WOMEN EMPOWERMENT IN RURAL ECONOMIES: IMPACTS OF MONETARY POLICY

DOI: 10.17261/Pressacademia.2026.2031
JBEF- V.15-ISS.1-2026(3)-p.26-35

Muhammad Ayub Mehar
Iqra University, Business School, Karachi, Pakistan.
ayubmehar@yahoo.com , ORCID: 0000-0002-9711-2556

Date Received: November 4, 2025 **Date Accepted:** March 21, 2026

To cite this document
Mehar, M.A., (2026). Woman empowerment in rural economics: impacts of monetary policy. Journal of Business, Economics and Finance (JBEF), 15(1), 26-35.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2031

ABSTRACT
Purpose- This study integrates economic development and prosperity with rural sociology and women's well-being. It assesses the women's empowerment in rural societies in the global context.
Methodology-- The study posits that women's subjective well-being in rural societies is a consequence of their empowerment, as measured by their participation in major decision-making. The study assesses women's empowerment in rural societies worldwide by analyzing data from 217 countries over 25 years.
Findings- The access and availability of credit to the private sector ensures the availability of cash, which influences households to allow more participation of female members in major family decisions. It was noted that the role of women in businesses is more influential than their role as workers.
Conclusions- The role of monetary policy was validated by the impact of domestic credit in enhancing the participation of females in businesses. It was also noted that higher per capita income improves women's ownership and participation in businesses.

Keywords: Domestic credit, rural sociology, monetary policy, subjective well-being, women in business
JEL codes: E51, I31, Z13

1. INTRODUCTION - WOMEN'S EMPOWERMENT AND SUBJECTIVE WELL-BEING

This study is mainly concerned with women's empowerment, which is an important indicator of women's subjective well-being. This study is important from the economic development and prosperity point of view, because agricultural output is closely integrated with the rural sociology and women's well-being. Assessing subjective well-being is a complicated task because it varies from person to person. It's closely related to personal perceptions. The issue becomes more complicated in the case of subjective well-being in women in rural areas, where the availability of reliable data is another challenging task. For this purpose, the active women's participation in major family decisions was taken as an indicator of the women's well-being.

The study assesses women's empowerment in rural societies in the global context. Women's empowerment in this study is measured through women's participation in major decision-making. In this way, women's empowerment deals with women's subjective well-being in rural societies. While the evaluation and experiences of individuals about their own lives, their perception of satisfaction in their lives, and their emotional reactions to events are referred to as subjective well-being. This type of well-being assesses how well a person feels her life is going in terms of overall satisfaction and emotional state. A high subjective well-being is usually considered a high level of life satisfaction and reflects happiness. This type of well-being will be exclusive to each individual.

Notably, contrary to subjective well-being, objective well-being can be assessed in terms of physical factors, including access and availability of food, clothes, shelter, money, employment, safety, security, education, and health. But measurement of subjective well-being is a difficult task. Various indicators of subjective well-being can be identified for research and operational purposes. The social, psychological, and spiritual factors can be included in these indicators. Healthy human functioning is also considered an important component of human well-being. Park et al (2023) presented a detailed description of the indicators to measure a specific form of subjective well-being. Diener (1984) defined subjective well-being in terms of three indicators: frequent positive effects, infrequent negative effects, and infrequent negative effects. Diener (1999) argued that the various components of subjective well-being indicate different constructs that need to be understood separately. Those constructs are closely related. So, subjective well-being should be considered as a general area of scientific interest rather than a single specific construct.

Different theories describe the causes of subjective well-being. These theories in the literature can be categorised into two classes: Top-down and bottom-up influence theories. Top-down theories suggest that people have a genetic predisposition to be happy or unhappy, and this predisposition determines their subjective well-being setpoints. These theories imply that a person's equilibrium level of subjective well-being is a consequence of hereditary characteristics. According to the bottom-up theories, happiness is created from happy experiences. These theories are based on the idea that there are universal basic human needs and that happiness results from their fulfilment. Another approach related to this topic is the hedonic treadmill theory, which proposes that most people return to a neutral level of subjective well-being (neither happy nor unhappy). The experts have identified various factors of subjective well-being and happiness: Personality and genetics (DeNeve: 1999), social influences (Fan Xiaojun et al: 2019), Wealth and Income (Shigehiro Oishi and colleagues: 2022, and Aknin, Norton, and Dunn: 2009), Health (Diener 2008), Neural characteristics (Sato: 2015 and Kurth: 2014), and Leisure (Hribernik and Mussap: 2010) are included in these factors.

The role of information technology in the expansion of e-commerce and credit to the rural households, women, and other marginalized groups has been studied by Mehar (2023a), Altundag (2025), Pal, Gupta, and Joshi (2022), Hendriks (2019), and Sobhan and Sharmin (2024). These studies explored new evidence and found the significant and effective role of monetary policy in determining e-commerce and empowering marginalized groups. Based on the listed firms of the Dhaka Stock Exchange, Sobhan and Sharmin (2024) found a positive and significant association between corporate social responsibility (CSR) disclosure and earnings management. This study shows how managers can use CSR disclosures as a competitive advantage. In the context of the Indian economy, Pal, Gupta, and Joshi (2022) found that the earning status of women, their participation in financial decision-making, and recipient of social welfare schemes by women have a significant impact on women's empowerment through financial inclusion.

Aziz, Sheikh, and Shah (2022) found that in nations where religious restrictions limit women’s willingness to work are less likely than males to own a bank account.

Based on the annual data of 217 countries for 25 years, Mehar (2025b) assessed the role of the monetary policy through financial inclusion, while a significant positive effect of financial inclusion on women's empowerment was ascertained. It was further noted in his ascertainments that the use of information technology facilitates access to banks and financial institutions. However, the direct intervention of monetary authorities is required to channelize the lending from commercial banks to employment creation in marginalized groups, women, and rural households.

The scope of this study is limited to covering the subjective well-being of women and their relation to the rural economy. Its core concern is to examine the effects of monetary policy on women's empowerment and the rural economy. The study is divided into 6 sections. The next section covers the economic factors of subjective well-being in economic literature. Section 3 establishes the model to explain the relations between monetary policy, rural population, female employment, and women's empowerment. Section 4 explains the data and statistical methodology. The results and empirical findings are explained in Section 5, while Section 6 highlights some policy implications and limitations of the study.

2. FACTORS OF SUBJECTIVE WELL-BEING IN ECONOMIC LITERATURE

The importance of economic conditions in determining socioeconomic prosperity and well-being has been discussed in various studies in the economic literature. Behera, Padmaja, and Dash (2024) established the relations between socioeconomic factors and happiness. Their study was based on empirical evidence. Mehar (2010) inferred that per capita income, subsidies and taxes play an instrumental role in creating the economic miseries in an economy. Mehar (2009) explained the economic factors of women's empowerment and noted that ownership of immovable properties, particularly agricultural farm houses and crop lands, is the most significant factor of women's empowerment in rural areas. The effect of women's employment is not as powerful as the ownership of business assets.

Microeconomics infers that an individual maximises the utility of his available resources through consumption. Microeconomic theory influences the thinking behind researching the relationship between social well-being and consumption. In discussion on the factors of economic well-being, Komlos (2023) have established the relationship between neurology, economics and human welfare. He (Komlos, 2023) introduced the concept of 'Human Economics' to cover the social and psychological factors in economics.

The role of monetary policy and financial institutions in determining the economic prosperity, well-being and women's empowerment is an important area of discussion in economic literature. There is a disagreement in the economic literature on the role of financial institutions in the determination of socioeconomic conditions. One school of thought emphasises the mechanism of monetary policy transformation to create a balance between GDP growth and inflation, while GDP growth ensures employment growth (Stein: 1982). The other school of thought emphasises the intervention by specialised financial institutions, including rural banks, non-banking financial institutions, and non-government organisations (NGOs), for creating employment opportunities and poverty alleviation. In establishing the relations between financial development and human welfare, Mehar (2024) identified 3 controversial issues: (1) Monetary policies focus on improving GDP growth and controlling inflation, but their impacts do not trickle down to lower-income groups. (2) The second controversial issue is the role of wealth concentration. Some experts favour wealth concentration because it builds new business empires, which contribute to growth and employment opportunities in a country, while another school of thought considers that these concentrations discourage the entry of new businesses. (3) The third issue covers the role of the financial system and institutions. Other than wealth accumulation, domestic credit to the private sector, interest rate spread, the magnitude of financial inclusion, structure and types of financial institutions, composition of borrowers, and external financing are also included in the vital components of the financial system. This system can create a blockage in the trickle-down effects of economic growth. The financial institutions can remove this blockage by transferring the benefits of macroeconomic growth to middle and lower-income groups through their lending policies. It was concluded that financial institutions, through their credit policies, can play an important role in the determination of employment conditions and business opportunities. The traditional approach of the trickle-down effects of the benefits of economic growth is not enough for the common people. Even the improvement in the banks' ability to lend, lower rates of interest, and allocation of credit to priority sectors cannot ensure the transfer of benefits of monetary and credit policies to the poor and vulnerable population. The inclusion of women, lower-income households, and the rural population in the financial system reflects the fairness and egalitarianism in the system, which is more important. The higher number of borrowers from banks and financial institutions improves the creation of new business entities, alleviates poverty, and reduces vulnerable employment. Mehar (2025) is concerned with the role of financial institutions in the supply of credit facilities to improve socioeconomic conditions. According to Mehar (2023), the more use of credit cards plays a very positive and significant role in developing the perception that people can arrange money within 30 days in case of financial emergencies. The more use of electronic payments and credit cards improves people's perception that they can manage money during the crisis. Similarly, in the determination of the people's perception that their top-most financial problem is to arrange money in case of a medical emergency due to a critical disease or accident, the availability of credit facilities is a significant factor. The role and mechanism of various types of financial institutions, including commercial banks has been explained by Mehar (2024a). He noted that all kinds of lending institutions, including banks and non-banking financial institutions, have similar attitudes toward lending to startups. However, the financial institutions owned by charitable organizations and philanthropists play a greater role in lending to the poor, women, and deprived people.

In establishing the relations between neurology, human well-being, and economic policies, Komlos (2023) advocates a new paradigm: Capitalism with a human face. He differentiated humanistic and mainstream economics based on their basic properties. Humanistic economics uses inductive logic (not deductive as in the case of mainstream economics). Humanistic economics implies that more just capitalism is possible, which enables people to live their daily lives with less anxiety, less conflict, less inequality, less insecurity, less manipulation, less pain, less poverty, less stress, less uncertainty, no unemployment, and less fear that their lives could spiral out of control in the next recession. This capitalism with a human face would also increase ethical behaviour, increase educational attainment, improve the health of the population, increase intellectual satisfaction, allow more leisure time, enable people to love and respect one another, improve social relationships, and enable the attainment of a moral life more easily. A wider definition of freedom was introduced by Komlos (2023). According to him, freedom is more than the absence of legal restraint to act. It includes the ability to live without the anxiety generated by a high-stress economy, so we should not have to worry about our jobs or pensions disappearing, being defrauded, or paying medical bills or college tuition. Everyone has the right to a standard of living adequate for the health and well-being of himself and his family, including food, clothing, housing, medical care, and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control. The markets should enable individuals to exercise their creativity, autonomy, and individuality without psychological manipulation or coercion. Mehar (2025a) described that families in extreme poverty have to sacrifice various kinds of freedom: Freedom to choose a profession, freedom to live in a favourite city, and freedom to learn are included in the list of those sacrifices. Even freedom of speech, freedom of religion, and freedom to develop social relations are closely associated with economic status. In tribal and rural societies and privately owned businesses, whole families are employed by the same employer. The nature of their jobs, residential status, location of work, and even political associations are associated with the employer's wishes. Such vulnerable employment is not considered unemployment in economic literature, though it affects human life miserably. Some studies define it as 'Modern Slavery'. Unfortunately, this new type of 'slavery' is not recognised in economic literature and policy-making circles. The relations between inflation, health expenditures, poverty, vulnerable employment, and death by suicide attempts have been noted in economic literature. The World Health Organisation has highlighted that more people die as a result of suicide than HIV, malaria, breast cancer, war, or homicide. Strangely, one in every 100 deaths is by suicide. Financial stress and job loss are included among the major factors of suicide. Poverty and financial stress may be a consequence of a flawed economic system and weak or exploitative economic policies. The consequences of a flawed economic system, ill planning, inefficiencies, and corruption of policymakers and economic managers should not be transferred to those who are not responsible for this.

In discussions on the impacts of economic factors, one should understand the difference between rural and urban sociology. The emotional attachment to the land, local culture, traditions and customs, and more limited economic resources compared to the urban areas are the usual characteristics of rural sociology. The participation of family members in business, unpaid workers, and the role of women in farm management and cropping are also included in the specialised characteristics of the rural economy. Another considerable point is the flow of income in the rural economy. The duration of the crop season determines the inflow of income of the rural households. The decisions regarding the purchase of assets, repayment of debts, dates of marriages, and travelling plans are connected with the crop seasons. The weather and environmental conditions also play an important role in the rural economy. These factors determine the role and participation of the female members in households, which is certainly different from that of urban societies.

3. RELATIONS BETWEEN MONETARY POLICY, FEMALE LABOUR PARTICIPATION, AND WOMEN'S EMPOWERMENT

This study examines the effectiveness of monetary policy in determining women's empowerment, particularly in rural societies. The incidence of monetary policy was measured through the magnitude of the credit to the private sector in the economy, while the empowerment was applied as an indicator of women's subjective well-being. In this study, women's empowerment was measured by the participation of a woman in 3 major decisions in their household (The households in which they usually live with their spouse and children, with or without others):

1) Decision about her health care,
2) Decisions about major household purchases, and
3) Decision about visiting family

The study establishes the simultaneity between the women's participation in businesses, availability of credit from banks and financial institutions, size of rural economy, and the women's participation in major family decisions. It is hypothesised that the size of the rural population, levies of taxes, and subsidies impact women's participation in businesses. In this background, the following hypotheses are established to test the effects of explanatory variables on the women's participation in major decisions:

1) Women's participation in business improves women's participation in major decision-making in households.
2) The size of the domestic credit to the private sector improves women's participation in major decision-making in households.
3) The size of the domestic credit to the private sector enhances women's participation in businesses.
4) The higher per capita income improves women's participation in major decision-making in households.
5) The higher per capita income improves women's participation in businesses.
6) The higher per capita income improves women's participation in the labour market.
7) The share of the rural population has dropped due to an increase in per capita income.
8) The share of the agriculture sector in gross domestic product (GDP) determines the women's participation in businesses.
9) The higher share of the agriculture sector in gross domestic product (GDP) improves the share of the rural population.
10) The size of the rural population affects women's participation in businesses.
11) The size of the rural population affects women's participation in the labour market.
12) The women's participation in the labour market improves their participation in businesses.
13) Employment in the agriculture sector improves women's participation in the labour market.
14) The size of crop land improves the share of the rural population.

The impact of monetary policy on women's empowerment was tested through the incidence of domestic credit in the economy. The target variable in this study is women's participation in family decisions. This is a widely accepted universal indicator of women's empowerment, because the ability of women to make decisions that affect their circumstances is an essential element of their empowerment and serves as an important contributor to their overall development. Participating in major decisions is affected by the access and availability of credit from banks and financial institutions, while it is indirectly affected by women's participation in businesses. This relation is mathematically expressed in the following equation:

$$\frac{dWOMDEC}{dDCPSG} = \frac{\partial WOMDEC}{\partial DCPSG} + \frac{\partial WOMDEC}{\partial WOMBUS} \cdot \frac{\partial WOMBUS}{\partial DCPSG}$$

Where WOMDEC indicates women's participation (%) in major households' decisions, WOMBUS indicates women's ownership or share in business entities (%), and DCPSG is domestic credit to the private sector as a % of GDP. To identify the significant factors of women's participation in family decisions (WOMDEC), and the share of women in businesses (WOMBUS), the following equations have been established:

$$WOMDEC_{it} = \propto_i + \beta_1 WOMBUS_{it} + \beta_2 DCPSG_{it} + \beta_3 PCI_{it} + \beta_4 INFL_{it} + \varepsilon_{it} \quad (1)$$

$$WOMBUS_{it} = \propto_i + \beta_1 DCPSG_{it} + \beta_2 PCI_{it} + \beta_3 INFL_{it} + \beta_4 AGRGDP_{it} + \beta_5 RURPOP_{it} + \beta_6 LABFTM_{it} + \beta_7 SUBSD_{it} + \beta_8 TXTGDP_{it} + \varepsilon_{it} \quad (2)$$

$$LABFTM_{it} = \propto_i + \beta_1 PCI_{it} + \beta_2 INFL_{it} + \beta_3 RURPOP_{it} + \beta_4 SUBSD_{it} + \beta_5 TXTGDP_{it} + \beta_6 EMPAGR_{it} + \beta_7 PVTMTL_{it} + \varepsilon_{it} \quad (3)$$

$$RURPOPR_{it} = \propto_i + \beta_1 DCPSG_{it} + \beta_2 PCI_{it} + \beta_3 AGRGDP_{it} + \beta_4 SUBSD_{it} + \beta_5 CRPLND_{it} + \varepsilon_{it} \quad (4)$$

Where, PCI is per capita income, INFL is the rate of inflation, AGRGDP is the share of the agriculture sector in GDP, RURPOP is rural population, LABFTM is the female to male ratio of labour participation rates, SUBSD is subsidies to the private sector, TXTGDP is the tax to GDP ratio, EMPAGR is the share of domestic employment in the agriculture sector, PVTMLT is the multidimensional poverty headcount ratio, and CRPLND is the size of crop land in a country. While ε_{it} is a stochastic disturbance term. The descriptions of these variables are presented in Table 1, while a simplified picture of the interaction among these variables is shown in Figure 1.

4. METHODOLOGY

The above-mentioned hypotheses are tested through empirical analysis, based on the annual data of 217 countries for 25 years (from 2000 to 2024), which provides 5425 observations. The data for this analysis were taken from the World Bank (2025), which covers women participating in the above-mentioned 3 decisions in percentage terms: own health care, major household purchases, and visiting family (% of women age 15-49). Women in business indicates the women's ownership or participation in business (%).

Tables 2 and 3 present the descriptive statistics and correlation matrix of the variables in this study. The panel least squares (PLS) technique was applied to quantify the impacts of explanatory variables. The appropriateness of the panel least-squares technique (PLS) and the selection of its associated methods (fixed effect model, FEM or random effect model) have been determined by the Lagrange Multiplier Tests (Breusch-Pagan, Honda, King-Wu, and Hausman Test). The model selection criteria are based on the Akaike information criterion, Schwarz criterion, and Hannan-Quinn criterion.

Figure 1: Impact of Domestic Credit on Women's Empowerment (Simultaneity in the Model)

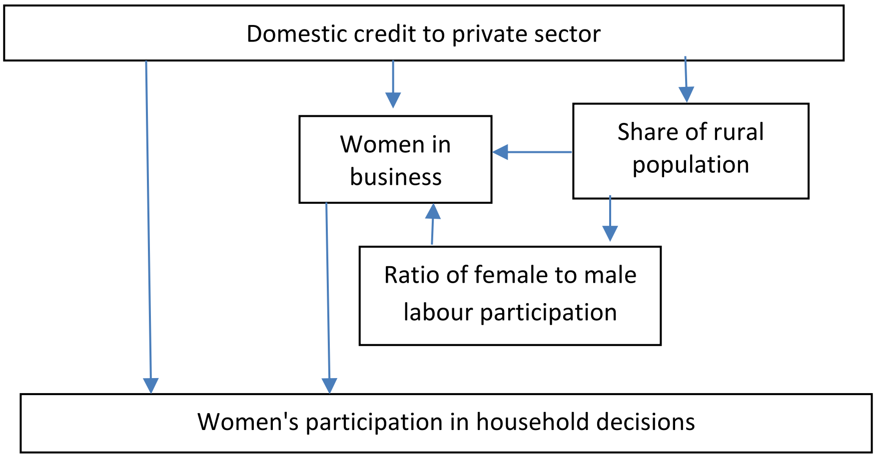


Table 1: List of Variables

Variable	Description
AGRGDP	Value added of agriculture, forestry, and fishing (% of GDP)
CRPLND	Permanent cropland (% of land area)
DCPSG	Domestic credit to private sector (% of GDP)
EMPLAGR	Employment in agriculture (% of total employment)
INFLCPI	Rate of inflation (%) based on the Consumer Price Index
LABRFTM	Ratio of female to male labour force participation rate (%) estimated by the ILO model
PCI	GDP per capita (USD)
PVRTMLT	Multidimensional poverty headcount ratio (% of population)
RURPOP	Rural population
RURPOPR	Rural population (% of total population)
SUBSD	Subsidies and other transfers (% of public expenditures)

TXTGDP	Tax revenue (% of GDP)
WOMBUS	Women's ownership or participation in Businesses
WOMDEC	Women (age 15-49 years) participating in the three decisions: own health care, major household purchases, and visiting family (%)

Source: Author's presentation based on World Bank (2025)

Table 2: Descriptive Statistics

Statistics\ Variable	DCPSG	RURPOP	EMPLAGR	AGRGDP	TXTGDP	PCI	INFLCPI	SUBSD	CRPLND	RURPOPR	WOMDEC	WOMBUS	PVRTMLT	LABRFTM
Mean	50.7	15784243	26.2	10.9	17.0	15929	6.6	39.1	4.5	41.1	49.1	69.7	8.8	70.9
Median	37.9	1895336	18.9	7.2	16.4	5364	3.5	37.6	1.5	40.9	50.6	73.1	2.0	76.5
Standard Deviation	44.4	75463640	22.9	10.7	7.6	24707	19.6	19.9	7.8	24.2	22.3	18.6	17.2	19.7
Minimum	0.0	0	0.1	0.0	0.0	110	-16.9	0.1	0.0	0.0	6.3	23.8	0.0	7.3
Maximum	304.6	915129968	91.9	79.0	147.6	256581	557.2	85.7	66.7	91.8	92.8	100.0	88.3	106.2

Table 3: Correlation Matrix

Variable	DCPSG	RURPOP	EMPLAGR	AGRGDP	TXTGDP	PCI	INFLCPI	SUBSD	CRPLND	RURPOPR	WOMDEC	WOMBUS	PVRTMLT	LABRFTM
DCPSG	1.000													
RURPOP	0.067	1.000												
EMPLAGR	-0.547	0.137	1.000											
AGRGDP	-0.521	0.092	0.790	1.000										
TXTGDP	0.274	-0.163	-0.299	-0.291	1.000									
PCI	0.626	-0.092	-0.584	-0.499	0.261	1.000								
INFLCPI	-0.151	0.006	0.122	0.107	-0.071	-0.123	1.000							
SUBSD	0.382	0.008	-0.411	-0.370	0.135	0.388	-0.060	1.000						
CRPLND	-0.057	-0.024	0.078	0.169	-0.041	-0.186	-0.034	-0.270	1.000					
RURPOPR	-0.487	0.134	0.752	0.661	-0.164	-0.453	0.087	-0.392	0.094	1.000				
WOMDEC	0.386	0.054	-0.290	-0.333	0.159	0.410	0.014	0.335	0.160	-0.300	1.000			
WOMBUS	0.445	-0.035	-0.293	-0.330	0.359	0.385	-0.099	0.495	-0.117	-0.254	0.402	1.000		
PVRTMLT	-0.407	0.224	0.727	0.770	-0.226	-0.365	0.199	-0.343	0.114	0.594	-0.680	-0.411	1.000	
LABRFTM	0.089	-0.102	0.195	0.084	0.216	0.172	-0.002	0.188	-0.037	0.095	-0.001	0.493	0.047	1.000

5. RESULTS AND EMPIRICAL ANALYSIS

The results of statistical analysis are presented in Tables 4 to 7. The significance of parameters has been tested through t-statistics, and the overall significance of the equation is measured through adjusted R-squares and their associated F-statistics. These parameters have also been reported in the concerned tables. To improve the reliability of results, some falsification tests have been applied in the regression analysis. For this purpose, some additional explanatory variables have been added. The consistency in the signs and negligible changes in the magnitudes of the betas associated with the main variables confirm the robustness and reliability of the results.

The appropriateness of the panel least-squares technique (PLS) and the selection of its associated methods (fixed effect model or random effect model), and the information losses in panel data have also been reported in Tables 4 to 7.

According to the statistical inferences, women's participation in family decisions is significantly improved by their ownership or participation in business. A woman in business ownership will be more powerful in household decisions than a woman without business ownership. The most important conclusion of this study belongs to the effectiveness of monetary policy, while the incidence of monetary policy in the economy was measured through the size of domestic credit to the private sector as a percentage of GDP. The credit to the private sector provides the lending facilities for the creation of new businesses and expansion in existing businesses. This provides bridge financing for business activities and supports working capital management. It was noted that expansion in domestic credit empowers women in multiple ways. First, it directly improves women's participation in household decision-making. The availability of cash makes it easy to make decisions. Second, it provides an opportunity for women to establish business entities. The participation or ownership of a business entity improves women's participation in household decision-making. The third important aspect of domestic credit is the encouragement of urbanisation. The higher domestic credit adversely affects the share of the rural population, while the statistical evidence in this study shows that a higher magnitude of rural population is associated with more women in businesses. These findings are statistically significant and robust.

Based on empirical analysis, it was also noted that higher per capita income improves women's family decisions and women's ownership and participation in businesses. It is also positively associated with the ratio of female to male labour force participation rate. However, it negatively impacts the share of the rural population in the total population, which implies that higher per capita income encourages urbanisation.

The size of crop land in a country and the share of agriculture in GDP boost the share of the rural population in a country; however higher share of agriculture in GDP discourages women's ownership in businesses. This reflects the male domination in the ownership of farmhouses and agricultural land. Contrary to this, the higher share of agriculture in GDP improves the ratio of female to male labour force participation rate, which indicates the higher participation of women in the agriculture sector. Another important finding is the association of poverty with female labour participation. The higher magnitude of poverty leads to a higher female-to-male labour force participation rate.

Table 4: Dependent Variable: Women Participating in Major Decisions (WOMDEC)

Method: Panel EGLS (Cross-section random effects) ##
Sample: 1 5425; Periods included: 63; Cross-sections included: 24
Total panel (unbalanced) observations: 166

Variable	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #
Constant	7.396 (0.767)	4.741 (0.447)	-0.516 (-0.036)
WOMBUS: Women's ownership or participation in Businesses	0.480*** (4.676)	0.441*** (3.641)	0.450*** (2.706)
RURPOPR: Rural population (% of total population)	-0.056 (-0.553)	-0.056 (-0.550)	0.003 (0.023)
DCPSG: Domestic credit to private sector (% of GDP)	0.309*** (3.965)	0.337*** (3.744)	0.319** (2.397)
PCI: GDP per capita (USD)	0.003** (2.326)	0.003** (2.387)	0.003* (1.906)
LABRFTM: Ratio of female to male labour force participation rate (%) estimated by the ILO model		0.059 (0.614)	0.001 (0.006)
INFLCPI: Rate of inflation (%) based on the Consumer Price Index			0.697** (2.130)
TXTGDP: Tax revenue (% of GDP)			0.301 (1.615)
Overall Significance			
R-squared	0.343	0.344	0.323
Adjusted R-squared	0.326	0.324	0.275
F-statistic	20.972	16.800	6.733
Testing for Fixed/ Random Effect			
Lagrange Multiplier Test: Breusch-Pagan	271.656***	256.854***	105.481***
Lagrange Multiplier Test: Honda	9.918***	9.553***	5.738**
Lagrange Multiplier Test: King-Wu	6.425**	6.144**	3.863*
Hausman Test (Cross-section random Chi-Square)	1.570	1.957	0.429
Durbin Watson Statistics	3.664	3.551	2.251
#T-Statistics in parenthesis ##: Swamy and Arora estimator of component variances *p < 0.1; **p < 0.05; ***p < 0.01			

Table 5: Dependent Variable: Women's Ownership or Participation in Businesses (WOMBUS)

Method: Panel Least Squares (Fixed Effect)
Sample: 1 5425; Periods included: 176; Cross-sections included: 24
Total panel (unbalanced) observations: 3567

Variable	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #
Constant	37.353*** (39.495)	36.341*** (27.974)	33.201*** (25.180)
PCI: GDP per capita (USD)	2.48E-05 (1.470)	7.28E-05*** (4.185)	6.87E-05*** (4.073)
DCPSG: Domestic credit to private sector (% of GDP)	0.107*** (15.136)	0.041*** (5.426)	0.037*** (4.970)
RURPOP: Rural population	1.85E-09 (0.679)	1.11E-08*** (3.158)	1.28E-08*** (3.781)
LABRFTM: Ratio of female to male labour force participation rate (%) estimated by the ILO model	0.452*** (35.823)	0.432*** (26.688)	0.400*** (25.215)
AGRGDP: Value added of agriculture, forestry, and fishing (% of GDP)	-0.391*** (-14.458)	-0.399*** (-10.172)	-0.364*** (-9.520)
INFLCPI: Rate of inflation (%) based on consumer prices index		-0.085*** (-4.630)	-0.078*** (-4.389)

SUBSD: Subsidies and other transfers (% of public expenditures)		0.216*** (14.576)	0.220*** (15.294)
TXTGDP: Tax revenue (% of GDP)			0.318*** (9.415)
Overall Significance			
R-squared	0.483	0.551	0.579
Adjusted R-squared	0.479	0.545	0.573
F-statistic	118.131	91.700	97.757
Testing for Fixed/ Random Effect			
Lagrange Multiplier Test: Breusch-Pagan	24500.890***	12075.810***	12016.280***
Lagrange Multiplier Test: Honda	130.911***	88.410***	87.928***
Lagrange Multiplier Test: King-Wu	83.071***	56.974***	56.616***
Hausman Test (Cross-section random Chi-Square)	289.981***	174.944***	168.562***
Durbin Watson Statistics	2.251	2.116	2.199
Criteria for Model Selection			
Akaike info criterion	8.094	7.809	7.739
Schwarz criterion	8.144	7.887	7.821
Hannan-Quinn criterion	8.111	7.837	7.769
#T-Statistics in parentheses *p < 0.1; **p < 0.05; ***p < 0.01			

Table 6: Dependent Variable: Rural Population as % of Total Population (RURPOPR)

Method: Panel Least Squares (Fixed Effect)
Sample: 1 5425; Periods included: 185; Cross-sections included: 23
Total panel (unbalanced) observations: 3916

Variable	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #
Constant	35.679*** (64.788)	36.963*** (51.028)	40.569*** (35.706)
PCI: GDP per capita (USD)	-3.80E-04*** (-23.272)	-3.33E-04*** (-16.514)	-1.98E-04*** (-8.906)
INFLCPI: Rate of inflation (%) based on the Consumer Price Index	-6.85E-04 (-0.050)	9.31E-03 (0.596)	-4.21E-02* (-1.817)
AGRGDP: Value added of agriculture, forestry, and fishing (% of GDP)	1.094*** (39.702)	1.052*** (33.762)	1.266*** (26.737
CRPLND: Permanent cropland (% of land area)	0.135*** (3.200)	0.247*** (5.079)	0.217*** (3.198)
DCPSG: Domestic credit to private sector (% of GDP)		-0.031*** (-3.785)	-0.033*** (-3.395)
SUBSD: Subsidies and other transfers (% of public expenditures)			-0.170*** (-9.060)
Overall Significance			
R-squared	0.536	0.532	0.560
Adjusted R-squared	0.533	0.528	0.554
F-statistic	173.374	142.857	99.854
Testing for Fixed/ Random Effect			
Lagrange Multiplier Test: Breusch-Pagan	33937.240***	29094.030***	16686.660***
Lagrange Multiplier Test: Honda	129.946***	119.303***	89.743***
Lagrange Multiplier Test: King-Wu	60.604***	55.908***	46.647***
Hausman Test (Cross-section random Chi-Square)	9.700***	6.053**	7.388***
Durbin Watson Statistics	1.995	1.958	1.851
Criteria for Model Selection			
Akaike info criterion	8.341	8.372	8.258
Schwarz criterion	8.384	8.423	8.333
Hannan-Quinn criterion	8.356	8.390	8.286
#T-Statistics in parentheses *p < 0.1; **p < 0.05; ***p < 0.01			

Table 7: Dependent Variable: Ratio of Female to Male Labour Force Participation Rate (LABRFTM)

Method: Panel Least Squares (Fixed Effect)
Sample: 1 5425; Periods included: 143; Cross-sections included: 15
Total panel (unbalanced) observations: 904

Variable	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #	Coefficient (t-Statistic) #
Constant	66.969*** (64.356)	67.932*** (62.589)	59.529*** (29.840)
CRPLND: Permanent cropland (% of land area)	-0.103 (-1.022)	-0.111 (-1.085)	-0.524*** (-5.682)
RURPOP: Rural population	-1.79E-07*** (-8.618)	-1.78E-07*** (-8.569)	-1.68E-07*** (-5.917)
PVRTMLT: Multidimensional poverty headcount ratio (% of population)	0.145*** (4.027)	0.152*** (4.180)	0.245*** (6.961)
EMPLAGR: Employment in agriculture (% of total employment)	0.132*** (3.197)	0.135*** (3.233)	0.123*** (3.483)
PCI: GDP per capita (USD)	2.73E-04*** (12.625)	2.62E-04*** (11.930)	1.53E-04*** (8.313)
INFLCPI: Rate of inflation (%) based on the Consumer Price Index		-0.168*** (-4.443)	-0.260*** (-3.427)
TXTGDP: Tax revenue (% of GDP)			0.228*** (3.234)
SUBSD: Subsidies and other transfers (% of public expenditures)			0.183*** (8.287)
Overall Significance			
R-squared	0.254	0.268	0.375
Adjusted R-squared	0.238	0.251	0.356
F-statistic	15.819	15.913	20.204
Testing for Fixed/ Random Effect			
Lagrange Multiplier Test: Breusch-Pagan	2039.722***	1874.678***	1443.820***
Lagrange Multiplier Test: Honda	32.406***	31.122***	28.523***
Lagrange Multiplier Test: King-Wu	17.367***	16.836***	17.660***
Hausman Test (Cross-section random Chi-Square)	7.843***	9.074***	13.129***
Durbin Watson Statistics	2.179	2.103	1.508
Criteria for Model Selection			
Akaike info criterion	7.941	7.935	7.376
Schwarz criterion	8.048	8.048	7.515
Hannan-Quinn criterion	7.982	7.978	7.429
#T-Statistics in parentheses *p < 0.1; **p < 0.05; ***p < 0.01			

6. POLICY IMPLICATIONS AND LIMITATIONS

The effectiveness of monetary policy through enhancing domestic credit to the private sector is validated in this study. From the policy point of view, it is important that the intervention of monetary authorities can improve women's empowerment in two different ways. First, the access and availability of credit to the private sector ensures the availability of cash, which influences households to allow more participation of female members in major family decisions. Second, the availability of credit facilities encourages the participation of females in businesses. The participation in businesses is also a significant factor in the improvement of women's empowerment. The higher labour participation of females in rural economies indicates their role in farm houses, cropping and harvesting activities. It may be an effect of the higher incidence of poverty in rural areas. But the important point is the higher impact of women in businesses on women's participation in family decisions. The role of women in businesses is more influential than that of women in employment as workers.

Despite the importance of this empirical analysis for policy formulation, the study has some limitations. It is based on macro-level data of the countries. It does not capture the cross-cultural differences and variations in the rural sociology of different regions and countries. The role of religions and ethnicities is important in assessing women's empowerment and its consequent subjective well-being. Another important aspect of the assessment of women's empowerment is the effects of education. The level of education certainly empowers women, though sometimes this effect can be channelled through women's participation in businesses. The migration, cultural adaptability, and technological changes can also affect women's empowerment. These factors can be included in future studies. A micro-level study, which is based on household analysis, can provide further details and highlight the hidden aspects of subjective well-being.

REFERENCES

Aknin, Lara B., Norton, Dunn. (2009). From wealth to well-being? Money matters, but less than people think. The Journal of Positive Psychology. 4(6), 523–527.

Altundag, S., (2025). Demographic profiles of e-commerce users: a regional and global perspective on Diyarbakir, Turkiye. Journal of Business Economics and Finance, 14(1), 1-14. https://doi.org/10.17261/Pressacademia.2025.1973

Aziz F., Sheikh S.M., Shah I.H., (2022). Financial inclusion for women's empowerment in South Asian countries. Journal of Financial Regulation and Compliance, 30(4), 489–502. https://doi.org/10.1108/JFRC-11-2021-0092

Behera, DK., Padmaja, M., Dash, AK. 2024. Socioeconomic determinants of happiness: Empirical evidence from developed and developing countries. Journal of Behavioural and Experimental Economics, 109, 102-187.

DeNeve, Kristina M. 1999. Happy as an extraverted clam? The role of personality for subjective well-being. Current Directions in Psychological Science, 8(5), 141–144.

Diener, E., (1984). Subjective well-being. Psychological Bulletin, 95(3), 542–575.

Diener, E., (2008). Happiness: unlocking the mysteries of psychological wealth. Malden, MA: Blackwell Publishers. ISBN 9781405146616.

Diener, E., Suh, E.M., Lucas, R.E., Smith, H.L. (1999). Subjective well-being: three decades of progress. Psychological Bulletin, 125(2), 276–302.

Fan, Xiaojun, Deng, Nianqi, Dong, Xuebing, Lin, Yangxi, Wang, Junbin., (2019). Do others' self-presentation on social media influence individuals' subjective well-being? A moderated mediation model. Telematics and Informatics, 41, 86–102.

Hendriks, S. (2019). The role of financial inclusion in driving women's economic empowerment. Development in Practice, 29(8), 1029–1038. https://doi.org/10.1080/09614524.2019.1660308

Hribernik, J., Mussap, A. J., (2010). Research note: Leisure satisfaction and subjective well-being. Annals of Leisure Research, 13(4), 701–708.

Komlos, J., (2023). Foundations of real-world economics: What every economics student needs to know (3rd edition), Routledge (Taylor and Fanc Group), ISBN: 978-1-032-00484-6

Kurth, F., Luders, E., Wu, B., Black, D.S., (2014). Brain grey matter changes associated with mindfulness meditation in older adults: an exploratory pilot study using voxel-based morphometry. Neuro, 1(1), 23–26.

Mehar, M. A., (2009). Economic Empowerment of Women in Pakistan: Theoretical Framework, Empirical Evidence, and Consequences of Policy Change. Islamabad: Friedrich Naumann Stiftung (Für Die Freiheit), 2009.

Mehar, M Ayub. 2010. Common Sense Economics: What Everyone Should Know about Economic Prosperity (Pakistani Edition). Islamabad: Friedrich Naumann Stiftung (Für die Freiheit)/ EFN Pakistan. 2010. ISBN: 978-969-9515-23-1

Mehar, M.A., (2023). Digital divide, credit financing and financial inclusion: Changing patterns of poverty, employment and income inequalities. Industrial Policy, 3(1), 21-41.

Mehar, M.A., (2023a). Role of monetary policy in economic growth and development: from theory to empirical evidence. Asian Journal of Economics and Banking, 7(1), 99–120. https://doi.org/10.1108/AJEB-12-2021-0148

Mehar, M.A., (2024). Nexus of Financial Architecture, Vulnerable Employment and Poverty: The Stimulus of Information Technology. Forum for Social Economics. 2024, 1-27. DOI: 10.1080/07360932.2024.2381437

Mehar, M.A., (2024a). Role of banks and non-banking financial institutions in alleviating poverty and unemployment. In: Apergis, N. (Ed.), Encyclopedia of Monetary Policy, Financial Markets and Banking, Elsevier, Academic Press. https://doi.org/10.1016/B978-0-44-313776-1.00296-8

Mehar, M.A., (2025). Role of financial institutions in poverty alleviation: Nexus of domestic credit, digital transactions and poverty in South Asia. Transnational Corporations Review, 17(1), 1-15. Doi: 10.1016/j.tncr.2025.200110

Mehar, M.A., (2025a). Book review "Foundations of real-world economics: What every economics student needs to know (3rd edition), by John K., published by Routledge (Taylor and Fanc Group), 2023, ISBN: 978-1-032-00484-6". Pakistan Development Review, 64 (1), 77-80.

Mehar, M.A., (2025b). Impacts of financial inclusion on women empowerment and poverty alleviation: a stimulative role of information technology. Journal of Economics, Finance and Accounting (JEFA), 12(2), 177-191. https://doi.org/10.17261/Pressacademia.2025.2020

Oishi, Shigehiro, Cha, Youngjae, Komiya, Asuka, Ono, Hiroshi, (2022). Money and happiness: the income–happiness correlation is higher when income inequality is higher". PNAS Nexus, 1(5), 224-235.

Pal M., Gupta H., Joshi Y.C., (2022). Social and economic empowerment of women through financial inclusion: empirical evidence from India. Equality, Diversity and Inclusion: An International Journal, 41(2), 294–305. doi: https://doi.org/10.1108/EDI-04-2021-0113

Park, C.L., Kubzansky, L.D., Chafouleas, S.M., (2023). Emotional well-being: What it is and why it matters. Affec.Science, 4(1), 10–20.

Sato, Wataru, Kochiyama, Takanori, Uono, Shota, Kubota, Yasutaka, Sawada, Reiko; Yoshimura, Sayaka, Toichi, Motomi. (2015). The structural neural substrate of subjective happiness. Scientific Reports, 5(1), 25-41. https://doi:10.1038/srep16891.

Sobhan, R., Sharmin, T., (2024). Examining the nexus between corporate social responsibility disclosure and earning management: evidence from the state-owned enterprises of Bangladesh. Journal of Business, Economics and Finance, 13(1), 24-36. https://doi.org/10.17261/Pressacademia.2024.1899

Stein, J.L., (1982). Monetarist, Keynesian & New Classical Economics. Blackwell, Oxford. World Bank. 2025. World Development Indicators. Washington: The World Bank Group, 2025.

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

DETERMINANTS OF EXPORTS IN CÔTE D'IVOIRE

DOI: 10.17261/Pressacademia.2026.2032
JBEF- V.15-ISS.1-2026(4)-p.36-47

Adel Shakeeb Mohsen
Graduate School of Management (GSM), Abidjan, Cote d'Ivoire.
adelmhsen@hotmail.com , ORCID: 0000-0002-6464-0248

Date Received: January 2, 2026 **Date Accepted:** April 25, 2026

To cite this document
Mohsen, A. S. (2026). Determinants of exports in Côte D'Ivoire. Journal of Business Economics and Finance (JBEF), 15(1), 36-47.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2032

ABSTRACT

Purpose- Given the critical role of exports in fostering economic growth, this study aims to investigate the determinants of exports in Côte d'Ivoire over the period 1970-2024. Specifically, the study examines the effects of trade openness, agricultural output, final consumption expenditure, and the exchange rate on export performance in Côte d'Ivoire. By analyzing influences of these macroeconomic factors, the study seeks to provide deeper insights into the key drivers of export growth in the Ivorian economy.
Methodology- Annual time series data over the period 1970-2024 were employed in this study. The data are obtained from the World Bank. The VAR model has been utilized in this study to test for cointegration among the variables and to examine the key determinants of exports in Côte d'Ivoire. The ADF unit root test, Johansen cointegration test, Granger causality test, and CUSUM test were applied to analyze the data. The ADF unit root test is applied to ensure the stationarity of time series data and prevent spurious regression. The Johansen cointegration test examines the presence of long-run equilibrium relationships among non-stationary variables, while the Granger causality test identifies the direction of causal and predictive relationships over time. The CUSUM test is used to evaluate the stability of model coefficients throughout the study period.
Findings- The Johansen cointegration test revealed that exports are positively related to trade openness and agriculture output, but it is related negatively with final consumption expenditure, and exchange rate. Agriculture output has the biggest effect on exports. The Granger causality test results showed that there are unidirectional short-run causality relationship running from exports to trade openness, and bidirectional short-run causality relationship between agriculture output and exports, but there is no evidence of any short-run causality relationship between final consumption expenditure, exchange rate and exports. Besides, there are bidirectional long-run causality relationships between trade openness, agriculture output, exchange rate and exports, and unidirectional long-run causality relationship running from final consumption expenditure to exports. Lastly, CUSUM test indicated that there are no structural changes in the model.
Conclusion- The study concluded that a coordinated policy approach that promotes trade openness, enhances agricultural production, manages domestic consumption, and ensures exchange rate stability is essential for strengthening Côte d'Ivoire's export sector and achieving sustainable economic growth.

Keywords: Côte d'Ivoire, Ivory Coast, international trade, trade liberalization, agricultural sector
JEL Codes: O11, E20, G15

1. INTRODUCTION

Trade, the exchange of goods and services through imports and exports, is widely recognized as a key driver of economic growth and development. Exports, in particular represent one of the earliest and most significant forms of international economic engagement. Their growing importance is reflected in an increasing share of global output. Export activities promote economic growth by boosting productivity, lowering production costs, enabling economies of scale through access to larger markets, fostering international integration, and attracting foreign direct investment.

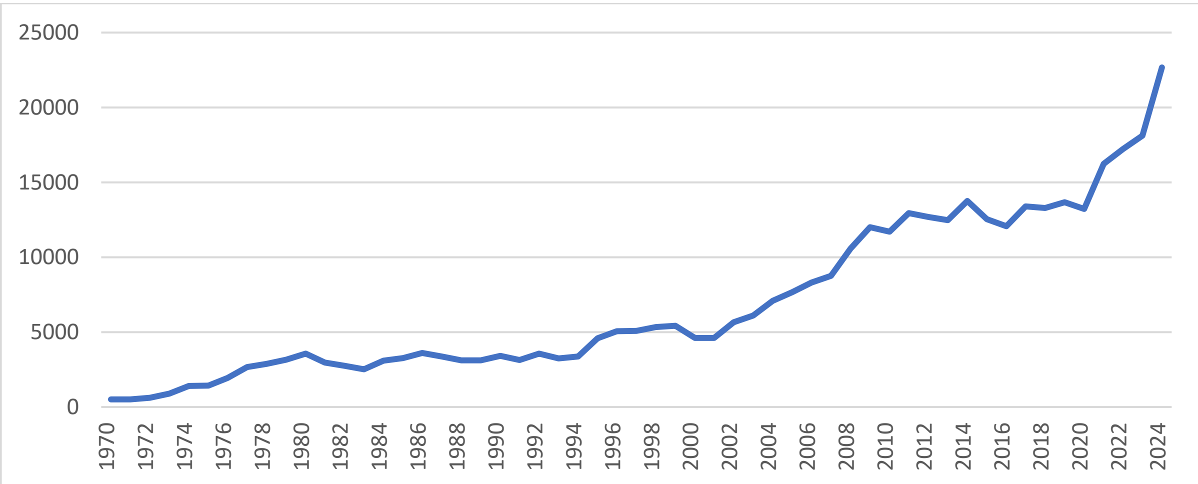
In this context, Côte d'Ivoire, located in West Africa and covering approximately 322,462 square kilometers, stands out as one of the largest economies within the West African Economic and Monetary Union (WAEMU). Over the past decade, the country has experienced a remarkable economic transformation, achieving one of Sub-Saharan Africa's fastest growth rates. As the world's leading producer and exporter of cocoa, Côte d'Ivoire maintained strong economic expansion, with GDP growth averaging 8.2% between 2012 and 2019. Despite global challenges and tighter financial conditions following the COVID-19 pandemic, economic momentum remained resilient, reaching 6.2% in 2023 and rising to 6.5% in 2024, supported by strong public and private investment, particularly in infrastructure and agriculture (Dago and Pei, 2025). Agriculture remains the backbone of the Ivorian economy, contributing over 16% of GDP and employing roughly 45% of the labor force, while playing a central role in export earnings and income generation (World Bank, 2024).

Côte d'Ivoire's economic performance is closely linked to its integration into the global trading system through multilateral, regional, and bilateral agreements that enhance market access, competitiveness, and sustainable development. Since joining the World Trade Organization (WTO) in 1995, the country has adhered to international rules on trade in goods and services and intellectual property, promoting transparency and predictability (WTO, 2023). Regionally, as a founding member of the Economic Community of West African States (ECOWAS) and participant in the West African Economic and Monetary Union (WAEMU/UEMOA), it benefits from a free trade area, a common external tariff, and the free movement of goods, capital, and persons, supported by coordinated fiscal and monetary policies (ECOWAS, 2022; UEMOA, 2022). On the continental level, Côte d'Ivoire has been part of the African Continental Free Trade Area (AfCFTA) since 2021, which aims to create a single African market by gradually eliminating tariffs and non-tariff barriers, fostering regional value chains, economic diversification, and industrial development (African Union, 2021). Additionally, the country has signed bilateral agreements, most notably the Economic Partnership Agreement (EPA) with the European Union, provisionally applied since 2016, granting duty-free and quota-free access to European markets while supporting agricultural and agro-industrial exports and long-term economic growth (European Commission, 2023).

As a major commercial hub in West Africa, foreign trade accounts for approximately 51% of Côte d'Ivoire's GDP. Nevertheless, the country's export structure remains highly concentrated in primary commodities, particularly agricultural products, which represent around 54.5% of total exports. This pronounced concentration reflects a structural reliance on cocoa as the dominant export commodity, accounting on its own for 35% of total exports, alongside natural rubber (12%) and fruits and nuts (7.5%). Such an export pattern entrenches the rent-based, agro-export orientation of the Ivorian economy, thereby heightening its exposure to volatility in global commodity prices and to climate-related shocks. In parallel, extractive industries contribute approximately 30.5% of total exports, split between gold (15%) and mineral fuels (15.5%), underscoring the growing role of natural resources in sustaining the trade balance and generating foreign exchange earnings. By contrast, manufactured and semi-manufactured goods together account for only 15% of total exports (10% and 5%, respectively), highlighting the limited depth of domestic industrialization and the persistently low level of value added embodied in Ivorian exports (World Bank, 2024).

Figure 1 illustrates the evolution of exports in Côte d'Ivoire from 1970 to 2024. Overall, the trend exhibits a steady upward trajectory, reflecting the growing significance of exports in the national economy. Export values increased from USD 520.6 million in 1970 to USD 3,163.8 million in 1979, indicating a period of relative economic stability and sustained production growth. However, between 1980 and 1983, exports declined from USD 3,561.6 million to USD 2,527.4 million. This downturn can be attributed to adverse external shocks, including global price volatility, rising energy costs, fluctuating interest rates, and severe droughts that reduced agricultural output and weakened global demand.

Figure 1: Export of goods and services in Cote d'Ivoire, at current prices, in million USD, 1970-2024 (World Bank, 2025)



From 1984 onward, exports recovered and expanded significantly, reaching USD 5,333.5 million in 1998. This resurgence was largely driven by government reforms aimed at liberalizing trade, reducing export barriers, and promoting international integration. Nevertheless, subsequent periods were characterized by further fluctuations. Between 1999 and 2001, exports fell from USD 5,423.1 million to USD 4,617.8 million due to political and military instability, which disrupted production, particularly in the agricultural sector. Exports rebounded to USD 12,000.3 million in 2009 but weakened again in 2010 following the post-electoral crisis, before recovering to USD 12,699.5 million in 2012.

Further fluctuations occurred after 2013, driven mainly by declines in global prices for key export commodities such as cashew nuts and cocoa. As a result, exports fell to USD 12,532.9 million in 2015 and USD 11,798.4 million in 2016. From 2017 onward, export activity resumed an upward trajectory, reaching a peak of USD 13,918.5 million in 2019. This was followed by a decline to USD 13,221.6 million in 2020, largely due to the COVID-19 pandemic, which disrupted global trade through travel restrictions and supply-chain interruptions. Exports subsequently rebounded strongly, rising to USD 22,673.5 million by 2024.

Recognizing the pivotal role of exports in national economic growth, the Ivorian government has implemented a range of measures to enhance export performance, including trade liberalization, tariff reductions, and the modernization of production technologies. Despite these efforts, the export sector faces both structural and external challenges. Agricultural exports remain highly vulnerable to climatic variability, land degradation, and pest outbreaks, such as cocoa swollen shoot disease (Dago and Pei, 2025). Additionally, limited access to modern technologies, inadequate infrastructure, and high production costs constrain productivity and weaken international competitiveness (Kouakou, 2020). Political and economic instability also continues to impede production continuity and the efficient marketing of exports.

Against this backdrop, this study aims to investigate the determinants of exports in Côte d'Ivoire over the period 1970-2024. The study is organized as follows: the next section presents a comprehensive literature review, followed by a discussion of the methodology. Subsequently, the empirical results are reported, and the final section concludes the study.

2. LITERATURE REVIEW

Several studies have examined the determinants of exports across countries using various econometric approaches. This section reviews selected empirical studies tested the effect of trade openness, agriculture output, final consumption expenditure, and exchange rate on exports of different country, which are relevant to the present research.

Regarding trade openness, numerous studies have extensively documented its impact on both export performance and economic growth, although the strength of this effect varies across different contexts. Zahonogo (2016), using a dynamic panel threshold model for Sub-Saharan African countries, found that trade openness enhances economic growth and exports up to a critical threshold, beyond which the effect weakens, highlighting the importance of supportive domestic conditions. Similarly, Keho and Wang (2017) and Guei and le Roux (2019) provided evidence from Côte d'Ivoire and ECOWAS countries, respectively, that trade openness significantly strengthens export performance and contributes to long-run economic growth. Country-specific studies further confirm these findings. Khalid (2016) showed that trade openness positively affects exports in Turkey in the short run, while long-run effects depend on structural factors. Onafowora and Owoye (1998) and Were (2015) supported the export-led growth hypothesis, emphasizing that openness promotes exports and growth, with stronger effects in developed and emerging economies. At the micro level, Kinuthia (2016) demonstrated that trade liberalization increases firm-level export participation in Kenya.

More recent evidence from Namibia by Sunde et al. (2023) confirmed a positive short- and long-run relationship between trade openness and exports. Chabi and Saygılı (2024) also found that trade liberalization can enhance export performance in West African economies, although its effects vary across sectors. Similarly, Onuogu et al. (2025) highlighted that trade openness promotes export-led growth in West Africa, particularly when human capital reaches a certain threshold. Besides, Samuel and Nkoro (2025) investigated the broader effects of trade liberalization on economic variables in West Africa, demonstrating that greater openness positively affects export dynamics and overall trade performance, even when the primary focus was on inflation. However, Rigobon and Rodrik (2005) argue that the effectiveness of trade openness in promoting exports critically depends on institutional quality, suggesting that openness alone is insufficient to ensure sustained export growth. Hence, the empirical literature largely supported a positive relationship between trade openness and export performance across countries. However, the impact of trade openness on exports is influenced by country-specific factors such as institutional quality, level of development, and complementary economic policies.

Moreover, a substantial body of empirical studies has examined the relationship between agricultural output and export performance, particularly in developing and agrarian-based economies. Using time-series data for Ethiopia, Tadesse (2008) found that growth in agricultural output significantly increases export earnings, underscoring the importance of productivity-enhancing policies. Similar evidence from Ethiopia's horticultural sector by Adugna and Zewdu (2015) indicated that non-traditional agricultural products increasingly contribute to export growth. Studies from other developing countries reinforced these findings. Dorosh and Haggblade (2003) showed that agricultural output growth in Bangladesh positively affects exports of traditional commodities, facilitating structural transformation and export diversification. In Nigeria, Oyekale and Egbetokun (2009) and Nwafor and Ugwu (2013) confirmed a strong short- and long-run relationship between agricultural output and export performance, despite policy inconsistencies. Further evidence from Rwanda, South Africa, and Chile demonstrates that agricultural output expansion significantly improves export earnings when accompanied by modernization, technological adoption, and institutional support (Murekezi & Bizoza, 2017; Lewin et al., 2004; Gomez & Ricketts, 2003). Panel and cross-country studies also supported this relationship. Shahbaz et al. (2015), Fulginiti and Perrin (1998), and Audu (2010) found that higher agricultural output and productivity positively influence exports, particularly where agro-processing linkages, supportive trade policies, and adequate infrastructure exist.

Evidence from South Asia further confirmed that output growth enhances agricultural exports, although limited diversification, price instability, and infrastructural constraints may restrict the full export potential (Ramachandran &

Swaminathan, 2008; Al-Mahmood & Rao, 2002). In Somalia, Abdi and Mohamed (2024) highlighted that higher agricultural output significantly increases both the volume and value of exports, emphasizing production capacity as a key driver of trade performance. Similarly, Aragie et al. (2023) found that in Ethiopia, Kenya, and Uganda, increases in agricultural production are strongly linked to greater export capacity, although potential trade-offs with domestic food prices exist. In the Southern African Development Community (SADC), Mabeta et al. (2025) demonstrated that foreign direct investment enhances agricultural output, which in turn leads directly to higher export levels. In Bangladesh, Hasan et al. (2022) indicated that growth in agricultural production positively influences export performance, contributing to overall economic growth. Collectively, these studies generally argued that increases in agricultural production enhance export capacity by expanding surplus availability, improving competitiveness, and strengthening foreign exchange earnings. However, the magnitude and direction of the relationship differ across countries depending on structural characteristics, policy frameworks, and levels of value addition.

Furthermore, many empirical literatures examined final consumption expenditure as a key determinant of export performance in both developed and developing economies. Evidence from time-series and panel studies indicated that rising domestic consumption may crowd out exports by diverting resources toward domestic markets. Consistent with this view, Baharumshah and Thanoon (2006), Boyd et al. (2001), Ahmed and Gupta (2009), and Al-Mahmood (2004) reported a negative effect of final consumption expenditure on exports in Malaysia, India, and Pakistan. Similar findings are documented for Sub-Saharan Africa and Bangladesh, where consumption-driven demand constrained export growth (Nkusu, 2004; Rahman, 2015). Evidence from Europe, East Asia, and Latin America further supported this view, indicating that high consumption expenditure reduces export growth in economies characterized by strong domestic demand and high import content of consumption (De Vita & Abbott, 2004; Lee & McKibbin, 2010; Zezza & Carfagna, 2017). However, some studies reported mixed or positive effects. Ali and Hammoudeh (2010) found that consumption expenditure exerted a positive long-run but insignificant short-run effect on exports in Egypt, while Qureshi and Ahmed (2010) showed that in Turkey, strong domestic demand may enhance export performance through scale and productivity effects.

Furthermore, In Albania, Akermi et al. (2024) showed that final consumption expenditure, along with domestic investment and imports, significantly affects exports in both the short and long run, highlighting the importance of consumption-driven demand on trade outcomes. Similarly, Sirajuddin (2025) found that in Austria, final consumption expenditure interacts with exports and other macroeconomic variables to shape economic performance, illustrating the broader link between domestic demand and export growth. In Egypt, Rashdan (2024) demonstrated that government final consumption expenditure can influence macroeconomic indicators, including export capacity, suggesting that public consumption policies may indirectly support trade. Moreover, in India, Srivastava et al. (2025) revealed that private and government consumption expenditures are closely linked with trade flows, indicating that higher consumption can stimulate export activity. Besides, Muda et al. (2025) showed that household and nonprofit institutional final consumption expenditure positively impacts exports of goods and services, emphasizing the direct role of domestic consumption in enhancing trade performance. Overall, the literature presents mixed evidence; however, the prevailing view suggested that although final consumption expenditure may stimulate production and support exports in certain contexts, high domestic consumption generally constrains export growth by absorbing productive resources.

Finally, exchange rate is one of the most extensively studied macroeconomic determinants of export performance, due to its direct influence on international price competitiveness. A large body of empirical literature has supported the view that exchange rate depreciation enhances export performance by improving price competitiveness. Using cointegration and error correction models, Marwah and Klein (2004) found that an overvalued real exchange rate significantly reduces Pakistan's export growth. Similar results were reported for South Africa, Turkey, Ghana, Nigeria, Sri Lanka, and Pakistan, where real exchange rate depreciation was shown to increase export volumes or earnings across various sectors (Armah et al., 2010; Kasman & Duman, 2006; Donkor, 2012; Akinlo, 2006; Athukorala & Sen, 2002; Mehmood & Farooq, 2015). Panel evidence from middle-income countries further confirmed that depreciation of the real effective exchange rate positively affects exports, although the magnitude of the response varies across countries and income levels (Bahmani-Oskooee & Ratha, 2004; Bopape & Krugell, 2007). Nonetheless, several studies highlighted that exchange rate effects are not uniform. Short-run export responses are often unstable due to external shocks and policy uncertainty, while long-run effects depend on structural conditions, sectoral composition, and global demand (Donkor, 2012; Akinlo, 2006).

Moreover, comparative analyses indicated country-specific heterogeneity in export responsiveness to exchange rate movements, particularly between Pakistan and India and across manufacturing sectors in Turkey (Choudhry et al., 2015; Gumus, 2007). In Nigeria, Gold and Yusuf (2025) found that exchange rate depreciation significantly boosts non-oil exports, highlighting the importance of currency valuation in enhancing trade performance. Similarly, in Mauritania, Elhadj and Koang (2024) showed that fluctuations in the exchange rate have a positive and significant effect on exports, suggesting that effective exchange rate management can support export growth. Focusing again on Nigeria, Gelle et al. (2025) highlighted that exchange rate volatility plays a critical role in shaping non-oil export outcomes, emphasizing the importance of stable yet flexible currency policies. Finally, in Kenya, Chepkwisich (2025) found that both short-run and long-run exchange rate fluctuations significantly influence coffee exports, illustrating how sector-specific exports are sensitive to currency movements. Collectively, these studies highlighted that while exchange rate depreciation generally enhances export performance by improving price competitiveness, the strength of this effect varies across countries. Moreover, structural characteristics, exchange rate regimes, and external demand conditions significantly shape the export response to exchange rate changes.

3. DATA AND METHODOLOGY

In this study, the Vector Autoregression (VAR) model is employed as the primary econometric framework. The VAR model is particularly well suited for analysing, describing, and forecasting the behaviour of economic and financial time series. It captures the dynamic interactions among multiple endogenous variables by modelling each variable as a function of its own lagged values and those of the other variables in the system. This framework enables the analysis of the dynamic effects of random shocks on the variables, as well as the assessment of their interdependencies over time. In addition, the VAR model provides a flexible and relatively unrestricted approach, as it does not require strong a priori assumptions regarding the direction of causality among variables. This flexibility makes it especially appropriate for empirical investigations involving complex economic relationships. In the context of this study, the VAR methodology is utilized to test for cointegration among the variables and to examine the key determinants of exports in Côte d'Ivoire.

Our model consists of five variables: exports (EXP), trade openness (OPEN), agriculture output (AGR), final consumption expenditure (FCE), and exchange rate (EXR). This model with EXP as the dependent variable is presented as follows:

$$\ln EXP = \beta 0 + \beta 1\ OPEN + \beta 2 \ln AGR + \beta 3 \ln FCE + \beta 4 \ln EXR + \varepsilon t \quad (1)$$

where β0 is the intercept, β1, β2, β3, and β4 are the slope coefficients, lnEXP is the natural log of exports (USD), OPEN is the trade openness as an indicator of the trade liberalization (the percentage of total exports and imports to GDP), lnAGR is the natural log of agriculture output (USD), lnFCE is the natural log of final consumption expenditure (USD), lnEXR is the natural log of exchange rate (CFA to USD), and εt is the error term (see Table 1). The a priori expectations of the slope coefficients of the export model in equation (1) are $\beta 1 > 0$, $\beta 2 > 0$, $\beta 3 < 0$, and $\beta 4 < 0$.

Annual time series data of Cote d'Ivoire over the period 1970-2024 will be used in this study. The data are obtained from the World Bank (WB). All variables in this study are expressed in the logarithmic form, except for trade openness (OPEN). There are many indicators of trade openness such as tariffs, the percentage of exports to GDP, the percentage of imports to GDP, the percentage of the trade balance to GDP, and the percentage of total exports and imports to GDP. The percentage of total exports and imports to GDP will be used in this study as an indicator of trade openness.

Table 1: The Variables

Variables		Definitions
Exports	EXP	Exports are goods and services produced in one country and sold to buyers abroad.
Trade Openness	OPEN	Trade openness is the process of reducing or eliminating restrictions on international trade, such as tariffs, quotas, and other barriers, to allow goods and services to move more freely between countries.
Agriculture Output	AGR	Agricultural output refers to the total quantity of crops, livestock, and other agricultural products produced by a country over a specific period.
Final Consumption Expenditure	FCE	Final consumption expenditure is the total spending by households, governments, and non-profit organizations on goods and services for direct use, rather than for producing other goods or investment.
Exchange Rate	EXR	Exchange rate is the price of one country's currency (CFA) in terms of another currency (USD), determining how much of one currency can be exchanged for another.

Because this study involves time series data, it is necessary to begin the analysis with the unit root tests. Augmented Dickey-Fuller (ADF) unit root tests will be conducted on each variable in the model to find out whether the time series data are stationary at the level or first difference. After testing for stationarity and confirming the order of integration of each time series, and if the variables in the model are found to be integrated of the same order, the Johansen cointegration test will be applied to establish whether there is any long-run or equilibrium relationship between the variables in the model (Engle and Granger, 1987; Johansen, 1991). If the variables are found to be cointegrated, then the Granger causality tests will be conducted based on the Vector Error Correction Model (VECM) to determine the long and short-run causality relationships among the variables in the model (Sims, 1980). However, the VECM will be subjected to the residual diagnostics, namely, the normality, serial correlation, heteroskedasticity and Ramsey RESET tests first to ascertain the statistical adequacy of the model before running the Granger causality tests. On the other hand, if the Johansen test results indicate no cointegration among the variables in a particular model, then the Granger causality tests will be based on the VAR model. Lastly, a stability test based on the cumulative sum (CUSUM) will be applied to determine whether the parameters of the model are stable over the period of the study.

4. FINDINGS AND DISCUSSIONS

This section presents and discusses the empirical results of this study. They include the various econometric tests and estimations, namely, ADF unit root test, Johansen cointegration test, statistical diagnostic tests, Granger causality test and stability test based on the cumulative sum (CUSUM).

4.1. ADF Unit Root Test Results

In the first step of the analysis, we carried out the ADF unit root test to determine whether the variables in the model are stationary or non-stationary at the levels. Table 2 shows that all the variables in the model are not stationary at the level, but became stationary after first differencing at 1% or 5% level of significance. Hence, all the variables in the model are integrated into order one, or I(1).

Table 2: ADF Unit Root Test Results

	Level			First difference		
	Intercept	Trend and intercept	No trend & no intercept	Intercept	Trend and intercept	No trend & no intercept
lnEXP	0.276268	-2.448250	3.252157	-3.764514***	-3.675249**	-1.543151**
OPEN	-0.957736	-2.163058	3.245078	-7.231644***	-6.732617***	-3.787245***
lnAGR	-1.860145	-1.752514	6.204257	-4.831409**	-5.143288**	-1.326231**
lnFCE	-2.313084	-1.137213	1.840707	-3.331074***	-3.963527***	-1.864202***
lnEXR	-1.380608	-2.345061	0.124109	-5.735154***	-5.743347***	-5.782152***

Note: *** denotes significance at the 1 percent level, and ** at the 5 percent level.

4.2. Johansen Cointegration Test Results

Since all the variables are stationary in the first difference, we can apply the Johansen multivariate cointegration test to determine if there is any cointegration or long-run equilibrium relationship between the variables in the model. However, before running the cointegration test we need to run the VAR model first to determine the optimal lag length, which is 3 based on the minimum AIC.

After having determined the optimal lag length, we then proceeded with the cointegration test for the model. Table 3 indicates that there are at most five cointegration equations based on the trace test and maximum eigenvalue test. In other words, the results reveal that there is more than one long-run relationship among the variables in the system comprising lnEXP, OPEN, lnAGR, lnFCE, and lnEXR.

Table 3: Johansen Cointegration Test Results

No. of CE(s)	Trace Statistic	0.05 Critical Value	Max-Eigen Statistic	0.05 Critical Value
r = 0	201.6321***	0.0000	64.15204***	0.0000
r ≤ 1	127.5648***	0.0000	40.53241***	0.0083
r ≤ 2	88.13601***	0.0000	34.43126***	0.0004
r ≤ 3	62.56358***	0.0000	26.45372**	0.0145
r ≤ 4	12.54357**	0.0131	11.53542**	0.0152

Notes: *** denotes significance at the 1 percent level, and ** at the 5 percent level.

After having found a cointegration relationships among the variables lnEXP, OPEN, lnAGR, lnFCE, and lnEXR, the cointegrating equation was normalized using the export variable.

Table 4: Cointegration Equation Normalized with respect to lnEXP

lnEXP	OPEN	lnAGR	lnFCE	lnEXR	C
1.000000	-0.56357	-0.62630	0.534264	0.556321	0.843075
	(0.00256)	(0.02210)	(0.07280)	(0.05263)	(0.00137)

From Table 4, the long-run lnEXP equation can be written as:

$$\ln EXP = -0.843 + 0.564\, OPEN + 0.626 \ln AGR - 0.534 \ln FCE - 0.556 \ln EXR \quad (2)$$

The cointegration equation given by equation (2) above shows that lnEXP is positively related to OPEN, and lnAGR, but it is related negatively with lnFCE and lnEXR.

The coefficient of OPEN indicates that for every one percent increase in trade openness, the export of Cote d'Ivoire will increase by 0.564 percent. It is clear that trade openness plays a crucial role in enhancing export performance in Côte d'Ivoire by reducing trade costs, expanding market access, improving productivity, and attracting foreign investment. The reduction of tariff and non-tariff barriers lowers transaction and export costs, while greater openness facilitates access to higher-quality intermediate inputs, capital goods, and modern technologies, thereby improving productivity and export quality. This enables domestic firms to meet international standards and strengthen their export capacity. Moreover, trade openness attracts export-oriented foreign direct investment, particularly in agro-processing industries, which increases productive capacity, promotes value-added exports, and supports export diversification. Finally, through regional integration within ECOWAS and trade agreements with the European Union, Côte d'Ivoire has gained improved access to foreign markets, leading to increased external demand for key exports such as cocoa, cashew nuts, rubber, and palm oil. Similar results were found by Zahonogo (2016), Khalid (2016), Keho and Wang (2017), Guei and le Roux (2019), and Sunde et al. (2023).

The coefficient of lnAGR indicates that for every one percent increase in agriculture output, the export of Cote d'Ivoire will increase by 0.626 percent. Agricultural output has a positive effect on exports in Côte d'Ivoire due to the central role of agriculture in the country's export structure. Increases in agricultural production expand the supply of exportable commodities, particularly cocoa, cashew nuts, rubber, and palm oil, enabling the country to meet both international demand and contractual export commitments. Higher output also improves export competitiveness by reducing unit production costs and supporting economies of scale. Moreover, increased agricultural output supports the development of agro-processing industries, especially in cocoa grinding and cashew processing. A larger and more stable supply of raw agricultural products encourages investment in processing activities, leading to higher exports of value-added agricultural products. This contributes to export diversification and higher export earnings. The same result is obtained by Fulginiti and Perrin (1998), Dorosh and Haggblade (2003), Tadesse (2008), Shahbaz et al. (2015), and Murekezi and Bizoza (2017)

The coefficient of lnFCE indicates that for every one percent increase in final consumption expenditure, the export of Cote d'Ivoire will decrease by 0.534 percent. Final consumption expenditure exerts a negative effect on exports in Côte d'Ivoire by diverting domestically produced goods from external to internal markets. When domestic consumption increases, a larger share of agricultural and industrial output is absorbed by local demand, reducing the surplus available for export. This effect is particularly relevant in Côte d'Ivoire, where key export commodities such as cocoa, cashew nuts, and palm oil are also used for domestic consumption and agro-processing. In addition, higher final consumption expenditure may lead to upward pressure on domestic prices, making Ivorian products less competitive in international markets. Increased domestic demand can also encourage producers to prioritize the local market, where transaction costs and risks are lower, rather than exporting. From a macroeconomic perspective, rising consumption may reduce national savings and investment in export-oriented sectors, thereby weakening export capacity. Consequently, in Côte d'Ivoire, higher final consumption expenditure negatively affects exports by reducing exportable surplus, increasing domestic prices, and limiting resources available for export-driven production. This result is consistent with Boyd et al. (2001), Baharumshah and Thanoon (2006), Ahmed and Gupta (2009), and Rahman (2015).

The coefficient of lnEXR indicates that for every one percent increase in exchange rate, the export of Cote d'Ivoire will decrease by 0.556 percent. Exchange rate movements have a negative effect on exports in Côte d'Ivoire, particularly through real exchange rate appreciation. An appreciation of the exchange rate increases the foreign-currency price of domestically produced goods, thereby reducing their price competitiveness in international markets. For Côte d'Ivoire, whose exports are largely concentrated in price-sensitive agricultural commodities, higher export prices can lead to a decline in external demand. Additionally, exchange rate appreciation reduces export profitability by lowering revenues earned in domestic currency, particularly when production costs are largely domestic. This weakens incentives for export-oriented production and investment in tradable sectors. Consequently, in Côte d'Ivoire, real exchange rate appreciation negatively affects exports by reducing price competitiveness, dampening external demand, and discouraging export-oriented investment. Similar findings were reported by Athukorala and Sen (2002), Marwah and Klein (2004). Akinlo (2006), Donkor (2012), Armah et al. (2010), and Mehmood and Farooq (2015).

4.3. Statistical Diagnostic Tests Results

Since the variables in the model are cointegrated, we have estimated the VECM to model the short-run dynamics. However, it is essential to subject the VECM to a number of diagnostic tests, namely, the normality (JB), serial correlation (LM), heteroskedasticity (BPG and ARCH) and Ramsey RESET to ascertain the model's statistical adequacy. A 5% level of significance will be used in all these tests. The results of the diagnostic tests are reported in Table 5.

The VECM with lnEXP, OPEN, lnAGR, and EXR as the dependent variable passed the normality, serial correlation, heteroskedasticity (BPG and ARCH) and Ramsey RESET tests. However, the VECM with lnFCE as the dependent variables passed the normality, heteroskedasticity (BPG and ARCH) and Ramsey RESET tests, but did not pass the serial correlation LM test. To address the serial correlation problem, the lag length was increased; however, serial correlation persisted. Consequently, Newey–West heteroskedasticity and autocorrelation consistent (HAC) standard errors were employed to correct for this issue before conducting the t- and F-tests for long-run and short-run Granger causality.

Table 5: Results of the statistical diagnostic tests on the VECM

	Dependent variables				
	lnEXP	**OPEN**	**lnAGR**	**lnFCE**	**lnEXR**
JB test	0.609178	0.357628	4.033942	2.048561	0.566061
	(0.728976)	(0.828476)	(0.120541)	(0.348059)	(0.745155)
LM test	1.644303(2)	1.051743(2)	0.023627(2)	0.24193(2) **	2.861922(2)
	(0.098488)	(0.112188)	(0.905488)	(0.002588)	(0.068688)
BPG test	1.864901	1.364399	0.407525	0.315701	0.85876
	(0.272688)	(0.259088)	(0.709788)	(0.864488)	(0.430988)
ARCH test	1.554411(1)	0.178542(1)	1.536739(1)	0.066942(1)	0.545961(1)
	(0.196488)	(0.660388)	(0.199088)	(0.796188)	(0.464812)
RESET test	0.069063(1)	4.217235(1)	0.982103(1)	0.408839(1)	0.013952(1)
	(0.811688)	(0.074512)	(0.367288)	(0.692788)	(0.969488)

Notes: ** denotes significance at the 1 per cent level and * at the 5 per cent level.

4.4. Granger Causality Test Results

Since the variables in the model are cointegrated, and the VECM satisfies the residual diagnostic tests, Granger causality tests based on the VECM are used to examine the short- and long-run causality relationships among the variables in the model. The F-test results show the significance of the short-run causal effects, while the significance of the coefficient of the lagged error correction term [ect(-1)] shows the long-run causal effect.

The Granger causality test results based on the VECM are shown in Table 6. It is clear that there are unidirectional short-run causality relationship running from lnEXP to OPEN, and bidirectional short-run causality relationship between lnAGR and lnEXP, but there is no evidence of any short-run causality relationship between lnFCE, lnEXR and lnEXP. On the other hand, there are bidirectional long-run causality relationships between OPEN, lnAGR, lnEXR and lnEXP, and unidirectional long-run causality relationship running from lnFCE to lnEXP in Côte d'Ivoire.

Table 6: Granger Causality Test Results

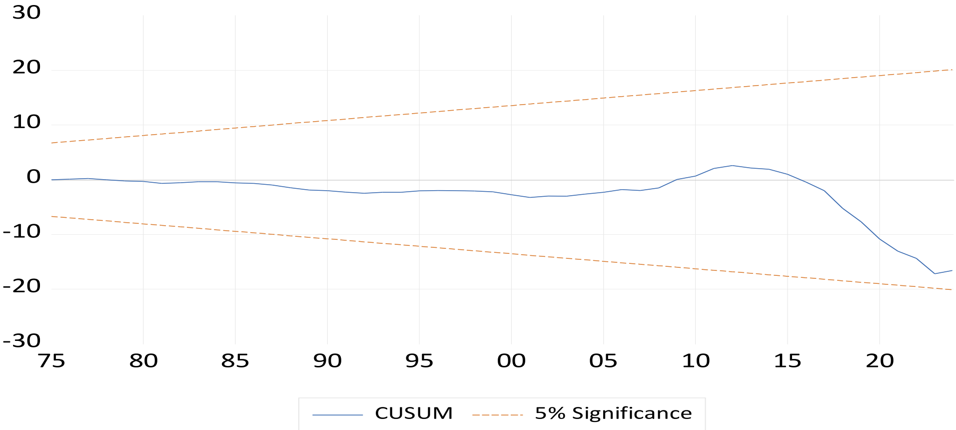
Dependent variables	Independent variables					
	∑Δ lnEXP	**∑Δ OPEN**	**∑Δ lnAGR**	**∑Δ lnFCE**	**∑Δ lnEXR**	**ect(-1)**
Δ lnEXP	-	0.130681	0.102876*	0.857206	0.264209	0.017214**
Δ OPEN	-3.24657**	-	0.446948	2.573409	2.643141**	0.274136**
Δ lnAGR	-1.34028**	-0.062312	-	0.866303	1.310550**	0.123261**
Δ lnFCE	-0.103724	-0.032842	-0.013576	-	0.052463	0.021648
Δ lnEXR	-1.247300	-0.263130	-0.295088	3.276244**	-	-0.02371**

Notes: ** denotes significance at the 5 per cent level and * at the 10 per cent level.

4.5. The Stability Test Result

CUSUM statistic is used in determining the parameter stability of the model in this study. The decision about parameter stability is based on the position of the plots relative to the 5 % critical bounds. If the plots of the CUSUM statistics stay within the area in the two critical lines, then the parameters of the model are stable over the period of the study. As illustrated in Figure 2, the CUSUM plot remains entirely within the 5% critical bounds over the study period. This result provides evidence that the parameters of the model are stable and that there are no structural breaks affecting the estimated relationships among the variables. Consequently, the model can be considered reliable for inference and policy analysis over the sample period.

Figure 2: CUSUM Test Results



5. CONCLUSION AND IMPLICATIONS

This study investigated how the exports in Côte d'Ivoire is affected by trade openness, agriculture output, final consumption expenditure, and exchange rate, from 1970 to 2024. The model consists of five variables, with export as the dependent variable. The ADF unit root test, Johansen cointegration test, Granger causality test, and stability tests were used in this study.

The ADF unit root test revealed that all the variables in the model are integrated by order one. The Johansen multivariate cointegration test revealed that exports are positively related to trade openness and agriculture output, but it is related negatively with final consumption expenditure, and exchange rate. Agriculture output has the biggest effect on exports. The Granger causality test results showed that there are unidirectional short-run causality relationship running from exports to trade openness, and bidirectional short-run causality relationship between agriculture output and exports, but there is no evidence of any short-run causality relationship between final consumption expenditure, exchange rate and exports. On the other hand, there are bidirectional long-run causality relationships between trade openness, agriculture output, exchange rate and exports, and unidirectional long-run causality relationship running from final consumption expenditure to exports in Côte d'Ivoire. Lastly, the stability tests indicated that there are no structural changes in the model.

Based on the findings of this study, trade openness and agricultural production play a significant positive role in enhancing exports. Trade liberalization facilitates integration into global markets, improves competitiveness, and stimulates export-oriented production. Similarly, growth in agricultural output, particularly in key commodities such as cocoa, cashew nuts, and palm oil, increases the exportable surplus, contributing substantially to export earnings and economic growth. These findings highlight the critical importance of export-oriented policies and investments in the agricultural sector to strengthen Côte d'Ivoire's position in international markets. Conversely, final consumption expenditure and exchange rate movements have been shown to exert negative effects on exports. Rising domestic consumption diverts locally produced goods from external to internal markets, reduces the exportable surplus, and can lead to higher domestic prices, thereby lowering international competitiveness. Similarly, exchange rate appreciation or volatility diminishes the price competitiveness of Ivorian exports, weakening their performance in global markets.

To enhance export performance, several policy measures are recommended. First, maintaining an open trade regime and reducing barriers to international markets can encourage export diversification and competitiveness. Second, strengthening agricultural productivity through modern farming techniques, improved access to inputs, and investment in agro-processing can increase the volume and quality of exportable commodities. Third, managing domestic consumption through policies that balance local demand and export supply, alongside promoting savings and investment in export-oriented sectors, can mitigate the crowding-out effect on exports. Finally, adopting a stable and competitive exchange rate policy, combined with measures to reduce currency volatility, can support consistent export growth. In conclusion, a coordinated approach that promotes trade openness, boosts agricultural production, manages domestic consumption, and ensures exchange rate stability is essential for strengthening Côte d'Ivoire's export sector and achieving sustainable economic growth.

REFERENCES

Abdi, A. H., & Mohamed, A. M. (2025). Determinants of agricultural exports in Somalia: The impacts of exchange rates, foreign direct investment, and institutional quality. Cogent Economics & Finance, 13(1), 1-16.

Adugna, L., & Zewdu, G. (2015). Horticultural production and export performance in Ethiopia. Ethiopian Journal of Economics, 24(1), 85-110.

African Union. (2021). Agreement Establishing the African Continental Free Trade Area (AfCFTA). Addis Ababa: African Union.

Akermi, A., Ben Yedder, H., & Bakari, S. (2024). Impact of final consumption, domestic investment, exports, and imports on economic growth in Albania. Theoretical and Applied Economics, 31(1), 231-252.

Akinlo, A. E. (2006). The stability of money demand in Nigeria: An autoregressive distributed lag approach. Journal of Policy Modelling, 28(4), 445-452.

Ali, A., & Hammoudeh, S. (2010). Macroeconomic determinants of export performance in Egypt: A causality and cointegration analysis. Middle East Economic Review, 22(3), 45-66.

Al-Mahmood, M. A. (2004). Domestic demand components and export growth in Pakistan. Pakistan Development Review, 43(2), 179-195.

Al-Mahmood, M. A., & Rao, N. S. (2002). Agricultural exports and economic growth in Pakistan. Pakistan Development Review, 41(2), 153-173.

Aragie, E., Balié, J., Morales, C., & Pauw, K. (2023). Synergies and trade-off between agricultural export promotion and food security: Evidence from African economies. World Development, 172, 1-14.

Armah, B., Kwafo, A., & Luguterah, A. (2010). Exchange rate volatility and exports: Evidence from South Africa. African Development Review, 22(3), 401-414.

Athukorala, P., & Sen, K. (2002). The determinants of private investment in India. World Development, 30(6), 1023-1037.

Bahmani-Oskooee, M., & Ratha, A. (2004). The J-curve: A literature review. Applied Economics, 36(13), 1377-1398.

Boyd, J. H., Levine, R., & Smith, B. D. (2001). The impact of financial development and consumption on export growth: A cross-country analysis. Journal of Economic Growth, 6(3), 361-386.

Bopape, L., & Krugell, W. (2007). The impact of exchange rate volatility on exports: Panel data evidence from South Africa. South African Journal of Economics, 75(3), 495-507.

Chabi, P., & Saygılı, R. F. (2024). Trade openness and structural change dynamics in West African countries. Journal of Economic Structures, 13(1), 1–27.

Chepkwisich, J. (2025). The impact of exchange rate fluctuation on coffee exports in Kenya. International Journal of Economic and Management Studies, 12(3), 85–96.

Choudhry, T., Hassan, S. S., & Papadimitriou, F. I. (2015). Exchange rate volatility and exports: Evidence from Pakistan and India. International Review of Economics & Finance, 39, 134-152.

Dago, D. A., & Pei, Y. (2025). Evaluating the position of Côte d'Ivoire's cocoa industry on the global production chain and the influencing factors. Sustainability, 17(3), 1-33

De Vita, G., & Abbott, A. (2004). Consumption, investment and export performance in the European Union: Panel evidence. Applied Economics, 36(7), 777-786.

Dorosh, P., & Haggblade, S. (2003). Growth linkages, price effects and income distribution in sub-Saharan Africa. Journal of African Economies, 12(2), 207-235.

Donkor, M. (2012). Exchange rate fluctuations and export performance in Ghana. International Journal of Economics and Finance, 4(7), 45-52.

Economic Community of West African States. (2022). ECOWAS Trade Liberalisation Scheme (ETLS). Abuja: ECOWAS.

Elhadj, B., & Koang, M. (2024). An empirical analysis of the impact of the exchange rate on exports: A case study of Mauritania. International Journal of Social, Management, Economics and Business Engineering, 18(5), 414–428.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error-correction: representation, estimation and testing. Econometric, 55(2), 251-276.

European Commission. (2023). Economic Partnership Agreement between the European Union and Côte d'Ivoire. Brussels: European Commission.

Fulginiti, L. E., & Perrin, R. K. (1998). Agricultural productivity and export performance in developing countries. Economic Development and Cultural Change, 46(4), 749-771.

Gelle, D. P., Orshio, T. R., Janar, I. V., & Musa, D. S. (2025). Exchange rate volatility and its significant impact on Nigeria's non-oil exports and economic growth prospects (1990–2024). Global Market Dynamics. 1(2). 33-44.

Gold, O., & Yusuf, A. (2025). The effect of exchange rate and other macroeconomic indicators on Nigeria's non-oil exports performance. SN Business & Economics, 5(113). 1-29.

Gomez, M. I., & Ricketts, K. D. (2003). Food value chain transformations and agricultural exports: Evidence from Chile. Agribusiness, 19(3), 269-292.

Guei, K. M., & Le Roux, P. (2019). Exports, trade openness and economic growth in ECOWAS countries. Journal of Economic and Financial Sciences, 12(1), 1-9.

Gumus, I. (2007). Export performance and exchange rate volatility in Turkey. Applied Economics Letters, 14(10), 771-774.

Hasan, M. M., Hossain, B. M. S., Sayem, M. A., & Afsar, M. (2022). Effect of agricultural exports and imports on economic growth in Bangladesh: A study on agribusiness supply chain. Journal of Distribution Science, 20(11), 79–88.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometric, 59(6), 1551-1580.

Kasman, A., & Duman, Y. S. (2006). Exchange rate volatility and exports: Evidence from Turkey. Applied Economics Letters, 13(10), 665-669.

Keho, Y., & Wang, M. (2017). The impact of trade openness on economic growth: The case of Côte d'Ivoire. Cogent Economics & Finance, 5(1), 1-14.

Khalid, A. (2016). Trade openness and economic growth in Turkey: Evidence from ARDL bounds testing approach. Journal of Economic and Social Development, 3(2), 12-25.

Kinuthia, B. K. (2016). Trade liberalization and export performance in Kenya. African Development Review, 28(1), 38-50.

Kouakou, P. K. (2020). Effect of agricultural and non-agricultural exports on economic growth in Ivory Coast. Review of Agricultural and Applied Economics, 23(2), 45-53.

Lee, J. W., & McKibbin, W. J. (2010). Final consumption expenditure and export competitiveness in East Asia: A panel study. Asian Economic Journal, 24(2), 213-236.

Lewin, B., Fisher, C., & Weber, M. (2004). Processed agricultural exports in developing countries: Issues and challenges. Food Policy, 29(4), 319-334.

Mabeta, J., Smutka, L., & Mrkvička, T. (2025). Impact of foreign direct investment on agricultural production and exports of SADC countries. Social Sciences & Humanities Open, 11. 1-10.

Marwah, K., & Klein, L. R. (2004). Real exchange rate effects on export volumes and prices: Empirical evidence from Pakistan. Journal of Asian Economics, 15(1), 83-100.

Mehmood, B., & Farooq, F. (2015). Exchange rate volatility and exports: Evidence from Pakistan. Journal of Asian Finance, Economics and Business, 2(4), 15-22.

Muda, I., Erlina, E., Nasution, A. A., Omar, N., Abdullah, N. M. H., Christy, A. N., & Peranginangin, D. S. M. (2023). The influence of population growth, households and NPISHs final consumption expenditure, exports of goods and services, net trade in goods and services, and merchandise exports to low and middle-income economies within the region on gross domestic product (GDP). Journal of Namibian Studies, 33(3), 661–669.

Murekezi, A., & Bizoza, A. R. (2017). Agricultural commercialization and export performance in Rwanda. African Journal of Agricultural and Resource Economics, 12(2), 122-137.

Nkusu, M. (2004). Macroeconomic determinants of export performance in Sub-Saharan Africa. IMF: IMF Working Paper No. 04/100.

Nwafor, M., & Ugwu, J. (2013). Agricultural output and export growth in Nigeria: Evidence from VECM analysis. International Journal of Economics and Management Sciences, 2(12), 1-9.

Onafowora, O. A., & Owoye, O. (1998). Can trade liberalization stimulate economic growth in Africa? World Development, 26(3), 497-506.

Onuogu, I. C., Akanegbu, B. N., & Azu, N. P. (2025). Trade openness and economic growth in West African countries: Does human capital matter? Journal of Information Systems Engineering and Management, 10(59s). 394-406.

Oyekale, A. S., & Egbetokun, O. A. (2009). Agricultural production and export performance in Nigeria. Journal of Sustainable Development in Africa, 11(3), 35-49.

Qureshi, M. A., & Ahmed, V. (2010). Domestic demand and export growth in Turkey: Evidence from cointegration analysis. Turkish Economic Review, 27(1), 89-104.

Rahman, M. M. (2015). Consumption expenditure and export performance in Bangladesh: A time-series analysis. Bangladesh Development Studies, 38(2), 113-135.

Ramachandran, V., & Swaminathan, M. (2008). Agricultural growth and export expansion in India. Economic and Political Weekly, 43(31), 71-79.

Rashdan, A. (2024). The impact of government final consumption expenditure on economic growth in Egypt. Central Bank Economic Review, 12(3), 101–120.

Rigobon, R., & Rodrik, D. (2005). Rule of law, democracy, openness, and income: Estimating the interrelationships. Economics of Transition, 13(3), 533-564.

Samuel, A., & Nkoro, E. (2025). Effect of trade liberalization on inflation in West African countries. International Journal of Research and Innovation in Social Science, 9(3), 4215–4229.

Shahbaz, M., Khan, S., Ali, A., & Bhattacharya, M. (2015). The impact of agricultural value added on exports: Evidence from South Asia. Quality & Quantity, 49(6), 2437-2458.

Sims, C. A. (1980). Macroeconomics and Reality. Econometric, 48(1), 1-48.

Sirajuddin, S. (2025). Assessing the impact of inflation, exports, imports, and final consumption expenditures on Austria's economic growth. Journal of Business and International Insights, 7(2), 45–62.

Srivastava, A., Raj, M., Pandiya, B., & Chauhan, A. S. (2025). Connection dynamics: Exploring trade, consumption, and government spending in emerging economies through ARDL analysis. Cogent Economics & Finance 13(1),1-17.

Sunde, T., Tafirenyika, E., & Adeyanju, I. D. (2023). Exports, imports, trade openness and economic growth in Namibia. Economies, 11(3).

Tadesse, G. (2008). Agricultural growth and export performance in Ethiopia. Journal of African Economies, 17(4), 589-620.

Were, M. (2015). Differential effects of trade on economic growth and investment: A cross-country empirical investigation. Journal of African Trade, 2(1-2), 71-85.

West African Economic and Monetary Union. (2022). Presentation and objectives of the West African Economic and Monetary Union. Ouagadougou: UEMOA.

World Trade Organization (WTO). (2023). Trade Policy Review: Côte d'Ivoire. Geneva: WTO.

Zahonogo, P. (2016). Trade openness and economic growth in developing countries: Evidence from sub-Saharan Africa. Journal of African Trade, 3(1-2), 41-56.

Zezza, A., & Carfagna, L. (2017). Final consumption expenditure and export dynamics in Latin America. Latin American Economic Review, 26(1), 57-80.

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

REVISITING THE DISTANCE PUZZLE: STRUCTURAL GRAVITY EVIDENCE FROM TURKIYE

DOI: 10.17261/Pressacademia.2026.2033
JBEF- V.15-ISS.1-2026(5)-p.48-64

Tugba Ortay Baykal
Istanbul Ticaret University, Department of Economics, Istanbul, Turkiye.
tugbaortaygokce@gmail.com , ORCID: 0000-0001-9983-8628

Date Received: January 15, 2026 **Date Accepted:** March 14, 2026

To cite this document
Baykal, T. O., (2026). Revising the distance puzzle: structural gravity evidence from Turkiye. Journal of Business, Economics and Finance (JBEF), 15(1), 48-64.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2033

ABSTRACT

Purpose-This paper re-examines the "distance puzzle" for Türkiye's bilateral exports and assesses whether large, stable distance elasticities primarily reflect estimation choices or persistent trade frictions in an emerging-market setting.
Methodology-Using CEPII benchmark-year data for 1996–2020, the study estimates three nested gravity specifications: (i) conventional log-linear OLS, (ii) PPML in levels, and (iii) an extended PPML model incorporating domestic trade flows and Türkiye–year effects to account for multilateral resistance and home bias. Institutional and macro-financial frictions are proxied by partner political stability and USD–TRY exchange-rate volatility.
Findings-OLS reproduces large and stable distance coefficients averaging –1.70, exceeding meta-analytic benchmarks by 70–90 percent. Once heteroskedasticity, zero-flow structure, and home bias are addressed through structural estimation, the average distance elasticity declines to –0.58—a reduction of approximately 65 percent. Political stability loses significance under structural controls, while exchange-rate volatility becomes precisely estimated and positive, indicating intensive-margin concentration during volatile periods rather than extensive-margin expansion.
Conclusion- The results support the interpretation that the distance puzzle is primarily an estimation artifact in emerging-market contexts. Distance remains economically significant as a composite trade-cost proxy bundling geographic, informational, and financial frictions. For policymakers, the findings suggest that trade diversification strategies should address not only physical connectivity but also currency risk management and relationship-building costs that elevate effective distance for emerging-market exporters.

Keywords: Structural gravity, PPML, distance puzzle, home bias
JEL Codes: F14, F50, C23

1. INTRODUCTION

The gravity model is one of the most reliable empirical frameworks in international economics. Its core insight is that bilateral trade rises with economic size and falls with geographical distance, and this pattern has been confirmed across countries, sectors and time periods. Yet a persistent empirical puzzle remains. Despite major improvements in transportation, logistics and communication technologies, the estimated effect of distance on trade has remained remarkably stable. Meta-analyses show that a doubling of distance still reduces trade by nearly the same magnitude as it did several decades ago (Disdier and Head 2008). This "distance puzzle" challenges the view that distance primarily reflects physical shipping costs.

A large body of research argues that distance captures a wider set of frictions. Anderson and van Wincoop (2003) demonstrated that bilateral trade depends on relative trade costs, which are summarized by multilateral resistance. When these terms are omitted, log-linear OLS estimates systematically overstate the contribution of distance. Methodological advances have reinforced this conclusion. PPML estimation corrects for heteroskedasticity and retains zero flows, and including domestic trade flows adjusts for home bias. Under these structural estimators, distance elasticities generally fall to more moderate values. Other studies highlight that distance bundles informational barriers, institutional similarity and macro-financial risk, making it a composite rather than a purely geographic measure.

Despite these insights, most empirical evidence focuses on advanced economies. Much less is known about how distance operates in emerging markets, where institutional quality, macro-financial volatility and regional integration differ markedly from high-income settings. This omission matters. In environments characterized by political uncertainty, exchange-rate instability and heterogeneous partner groups, distance is likely to interact with risk and institutional proximity in ways that standard gravity models do not fully capture.

This paper addresses this gap by re-examining the distance puzzle from the perspective of Türkiye. Türkiye provides a useful case for three reasons. First, it lies at a geographic and institutional intersection, with deep integration into the European

Union and weaker ties with more distant regions. Second, its macro-financial environment has experienced repeated episodes of volatility, particularly in the exchange rate, which may amplify or reshape the effect of distance. Third, its export structure is highly concentrated in nearby and institutionally familiar markets, with relatively weak engagement in distant destinations.

Using bilateral export data from 1996 to 2020, the paper estimates three versions of the structural gravity model: a traditional log-linear OLS benchmark, PPML in levels and an extended PPML specification that incorporates domestic trade flows and exporter fixed effects. This sequential approach allows the evolution of the distance coefficient to be traced as key methodological issues are addressed, including heteroskedasticity, multilateral resistance and home bias. The analysis also includes two sources of institutional and macro-financial risk, political stability and exchange-rate volatility, to assess whether residual distance effects reflect deeper institutional or financial frictions.

The motivation for this paper originates from an earlier empirical study by the author (Baykal, 2026), which used standard gravity specifications and revealed a puzzling increase in Türkiye's estimated distance coefficients over time—despite rising globalization and improvements in transport technology. As shown in Table 1, both OLS and PPML models yielded unexpectedly low or even positive distance elasticities in some cases. These findings raised two critical questions: (1) whether distance in the Turkish context proxies broader frictions beyond geography, and (2) whether previous results were driven by omitted structural components such as multilateral resistance or home bias. The present paper builds directly on these results and provides a structurally grounded reassessment of the "distance puzzle" using refined estimation strategies and additional institutional controls. Table 1 summarizes the earlier findings that prompted this re-investigation.

Table 1: Motivation from Earlier Study *(Genç and Baykal, 2026)*

	OLS		PPML	
	(1)	(2)	(3)	(4)
Log distance	-0.127***	-0.005	-0.011	0.032
	(0.019)	(0.018)	(0.030)	(0.028)
FX volatility	-0.608***	-0.674***	-0.225***	-0.297***
	(0.014)	(0.013)	(0.018)	(0.030)
Political stability	0.047***	0.082***	0.081***	0.106**
	(0.013)	(0.046)	(0.021)	(0.036)
Log output	0.938***	0.956***	0.890***	0.906***
	(0.005)	(0.004)	(0.011)	(0.011)
Log expenditure	0.921***	0.907***	0.915***	0.903***
	(0.006)	(0.008)	(0.011)	(0.011)
Exporter remoteness index		-0.136***		-0.059***
		(0.002)		(0.008)
Importer remoteness index		-0.097***		-0.040***
		(0.003)		(0.005)
Constant	-10.056***	-6.775***	-10.527***	-9.055***
	(0.210)	(0.164)	(0.324)	(0.447)
Sample size	2071	2071	2071	2071
R^2	0.98	0.98	0.97	0.97

The paper makes three contributions. First, it provides a systematic assessment of the distance puzzle in an emerging-market context and shows that the puzzle becomes much weaker once structurally consistent estimators are used. Second, it demonstrates that institutional and macro-financial conditions help explain part of what naive models attribute to distance, with exchange-rate volatility becoming significant under structural estimation. Third, it complements the econometric results with graphical evidence, showing how institutional proximity and macro-financial risk shape Türkiye's export geography. For policymakers, the findings clarify whether distance primarily reflects physical trade costs or deeper financial and institutional frictions.

The remainder of the paper is organized as follows: Section 2 reviews the literature, Section 3 presents the data, Section 4 outlines the methodology, Section 5 reports the results and graphical evidence, and Section 6 concludes.

2. LITERATURE REVIEW

The gravity model has established itself as the workhorse of empirical international trade research, consistently demonstrating that bilateral trade flows increase with economic size and decrease with geographic distance (Tinbergen, 1962; Anderson, 1979). Yet this empirical success has generated one of the field's most enduring puzzles: despite dramatic technological advances in transportation, communication, and logistics over the past half-century, the estimated elasticity of trade with respect to distance has shown remarkably little tendency to decline and, in some contexts, has even increased (Disdier and Head, 2008; Head and Mayer, 2014). This "distance puzzle" presents a fundamental challenge to the view that globalization has substantially reduced the friction of geographic separation and has motivated extensive research examining whether distance remains a binding constraint on trade or whether conventional estimation methods systematically overstate its importance (Buch et al., 2004; Lin and Sim, 2012).

This literature review synthesizes research on three interconnected questions that frame the empirical analysis. First, how large are distance effects in international trade, and have they declined over time as transportation and communication technologies improved? Second, to what extent do methodological choices, particularly the treatment of heteroskedasticity, zero flows, and multilateral resistance, influence estimated distance elasticities? Third, what explains the observed distance sensitivity beyond physical shipping costs, and how do these composite frictions operate in contexts characterized by institutional weakness and macro-financial volatility? The review demonstrates that the distance puzzle is neither universal nor purely methodological but reflects genuine friction persistence that varies systematically across development stages and institutional environments, providing direct motivation for a country-specific reassessment using structural gravity methods (Carrère et al., 2013; Yotov, 2022).

2.1. The Empirical Persistence of Distance Effects

The systematic documentation of distance persistence begins with Disdier and Head (2008), whose meta-analysis of 1,467 distance elasticities drawn from 103 papers represents the foundational empirical statement of the puzzle. Across studies spanning multiple decades, estimation methods, and country samples, they find no statistically significant downward trend in the distance coefficient since the 1960s. The average elasticity hovers around minus one, implying that a doubling of distance reduces bilateral trade by approximately 50 percent, and this magnitude appears remarkably stable despite the container revolution, the internet, and broader forces of globalization (Hummels, 2007). Their meta-regression controls for differences in sample composition, data quality, and econometric specification, yet the puzzle persists across virtually all contexts examined (Disdier and Head, 2008). This finding challenges the "death of distance" hypothesis advanced by some observers of technological change (Cairncross, 1997) and suggests that geographic separation continues to impose substantial economic costs.

Head and Mayer (2013, 2014) extend this analysis through comprehensive synthesis examining sources of resistance to globalization. Their meta-analysis of 1,835 distance elasticity estimates reveals a median coefficient of minus 0.89, meaning that a 10 percent increase in distance reduces trade by approximately 8.9 percent (Head and Mayer, 2014). Critically, they demonstrate that borders and distance impede trade by much more than tariffs or transport costs can explain, crystallizing the puzzle as the gap between observed distance sensitivity and what measurable trade costs would predict (Anderson and van Wincoop, 2004). They show that commercial services and foreign direct investment exhibit distance sensitivities similar to goods trade, while financial assets display somewhat smaller but still substantial distance effects, suggesting the puzzle extends beyond physical goods to encompass most forms of cross-border economic activity (Portes and Rey, 2005; Head and Mayer, 2013).

Anderson (2000) poses the question directly in "Why Do Nations Trade (So Little)?" examining the border effect between the United States and Canada. His analysis reveals that distance effects are "surprisingly large" and contribute 20.4 percent to average price volatility between these markets, far exceeding what shipping costs alone would justify (Anderson, 2000). This finding is particularly striking because trade between the United States and Canada occurs across a largely unobstructed border between two high-income countries with deep institutional similarities (McCallum, 1995), suggesting that if distance effects are this large in favorable circumstances, they are likely even more substantial in less integrated contexts (Anderson and Marcouiller, 2002). The persistence of large border effects despite NAFTA and subsequent trade liberalization reinforces the interpretation that distance proxies frictions beyond tariffs and transport costs (Agnosteva et al., 2019).

Importantly, subsequent research reveals that the distance puzzle is not universal but exhibits pronounced heterogeneity by income level. Carrère, de Melo, and Wilson (2009, 2010, 2013) demonstrate through comprehensive panel analysis covering 124 countries from 1970 to 2006 that the puzzle is fundamentally a developing country phenomenon. For low-income countries, defined as the bottom third of the global income distribution, distance elasticities increased by 15 to 18 percent over this period (Carrère et al., 2013). In stark contrast, high-income countries exhibited no distance puzzle: their distance coefficients remained stable or showed modest declines, consistent with the hypothesis that technological change reduces geographic frictions for advanced economies but not for countries with weak institutions and infrastructure deficits (Limão and Venables, 2001; Carrère et al., 2010).

Recent empirical work increasingly interprets the distance puzzle as a measurement and modeling artifact rather than a genuine structural phenomenon. Duan et al. (2022) demonstrate using WIOD data that global value chains explain much of the puzzle: intermediate-goods trade across borders raises the distance elasticity of gross trade relative to value-added trade, and once GVC effects are controlled for, the elasticity diminishes over time. Similarly, Rezitis et al. (2025) find that distance effects in global coffee trade are initially stable but then decline, especially after 2015, contradicting a persistent puzzle in this commodity market. Kondaridze et al. (2025) show that separating intra-national from international distance improves model fit in dairy trade and reduces puzzle-type findings, suggesting that conflating domestic and international distances can distort the distance coefficient.

This income-based stratification fundamentally reshapes interpretation of the puzzle. Advanced economies have partially overcome distance barriers through infrastructure investment, institutional development, and information network deepening (Nordås and Piermartini, 2004), while developing countries face rising effective distance as they attempt to integrate into global markets from positions of institutional weakness (Rodrik, 1999; Carrère et al., 2009). Head and Mayer (2013) confirm this heterogeneity, finding that distance effects are rising most for new and low-trading countries while remaining stable for established high-trading nations. This development margin effect implies that entering global markets from limited prior integration confronts countries with steeper distance gradients than would be faced by countries with established trade positions (Chaney, 2014), suggesting that first-mover advantages and network effects play important roles in shaping trade geography (Rauch, 1999; Rauch and Trindade, 2002).

2.2. Methodological Corrections and Structural Gravity

A substantial body of research argues that part of the observed distance persistence reflects econometric misspecification rather than genuine structural frictions. The theoretical foundation is Anderson and van Wincoop (2003), who formalize the concept of multilateral resistance within a structural general equilibrium framework. Their key insight is that bilateral trade depends not on absolute bilateral costs but on bilateral costs relative to a country's average trade costs with all partners (Anderson and van Wincoop, 2003). When multilateral resistance terms are omitted, the distance coefficient absorbs both direct bilateral effects and indirect general equilibrium adjustments, generating upward bias (Feenstra, 2004). This insight has profound implications for gravity estimation: coefficients estimated without proper controls for multilateral resistance confound bilateral frictions with general equilibrium price adjustments, systematically overstating the role of distance (Baldwin and Taglioni, 2006).

Baldwin and Taglioni (2006) catalog common errors in gravity estimation, emphasizing that failure to account for multilateral resistance can severely bias trade-cost estimates. They identify the "gold medal mistake" of omitting multilateral resistance terms, the "silver medal mistake" of using inappropriate deflators, and the "bronze medal mistake" of ignoring time-varying unobservables through improper panel specifications (Baldwin and Taglioni, 2006). Their practical guidance, which has shaped subsequent empirical practice, recommends using exporter-time and importer-time fixed effects to absorb multilateral resistance non-parametrically (Feenstra, 2016), though they acknowledge this approach sacrifices the ability to identify time-invariant bilateral characteristics and country-specific institutional variables.

The methodological critique extends beyond multilateral resistance to the treatment of heteroskedasticity and zero trade flows. Santos Silva and Tenreyro (2006) demonstrate that log-linearization of the multiplicative gravity equation introduces Jensen's inequality bias under heteroskedasticity, and that conventional OLS on log-transformed flows yields inconsistent estimates when the variance of trade flows is proportional to the conditional mean squared, a pattern endemic to trade data (Santos Silva and Tenreyro, 2006). Their proposed solution, Poisson pseudo-maximum likelihood (PPML) estimation in levels, has become standard practice in structural gravity applications (Head and Mayer, 2014; Yotov et al., 2016). PPML naturally accommodates zero trade flows without ad hoc exclusion or imputation (Helpman et al., 2008), and comparative exercises demonstrate that it typically yields distance elasticities 20 to 40 percent smaller in absolute magnitude than log-linear OLS (Fally, 2015; Santos Silva and Tenreyro, 2006).

Recent methodological advances have further refined structural gravity estimation. Freeman et al. (2025) introduce a two-stage PPML procedure that recovers country-specific effects and trade elasticities without requiring price or tariff data, applicable at both sectoral and aggregate levels including services trade. Larch et al. (2025) codify best-practice recommendations for gravity estimation, emphasizing PPML with domestic flows and rich fixed effects as the benchmark structural setup. Pfaffermayr (2020) demonstrates that constrained PPML, imposing Anderson–van Wincoop equilibrium conditions, delivers nearly correct standard errors and coverage rates, whereas unconstrained PPML standard errors are severely downward biased. These developments reinforce the importance of structural consistency in gravity applications.

Helpman, Melitz, and Rubinstein (2008) develop a two-stage estimator explicitly modeling firm selection into exporting, showing that ignoring selection biases trade-cost estimates upward. Their framework recognizes that only sufficiently productive firms can profitably overcome the fixed costs of entering distant markets (Melitz, 2003), and failure to account for this extensive margin generates upward bias in estimated distance elasticities (Helpman et al., 2008). This selection mechanism is particularly important in contexts with high entry barriers or weak institutions, where only the most productive firms engage in international trade (Bernard et al., 2007; Manova, 2013).

Yotov (2012, 2022) proposes an additional refinement: combining PPML estimation with explicit inclusion of domestic trade flows. His argument is that traditional gravity specifications compare international trade to an implicit counterfactual of frictionless trade, when the appropriate comparison is to observed domestic trade (Yotov, 2012). Including domestic flows provides the natural benchmark for quantifying international frictions, effectively measuring distance effects relative to home bias rather than relative to an unobserved zero-friction baseline (Ramondo et al., 2016). When distance is measured relative to domestic trade and modern estimators are employed, the distance effect exhibits a clearer downward trend (Yotov, 2022). This approach has been formalized by Yotov, Piermartini, Monteiro, and Larch (2016) in their WTO-UNCTAD guide, which has become the standard reference for structural gravity implementation and explicitly recommends the combination of PPML estimation, domestic trade inclusion, and theory-consistent fixed effects (Yotov et al., 2016; Larch et al., 2019).

However, even after applying these corrections, distance effects remain economically significant, indicating that methodological refinement resolves only part of the puzzle. Baier, Kerr, and Yotov (2017, 2018) consolidate best practices while acknowledging that modern structural gravity specifications still yield substantial distance coefficients, suggesting genuine trade frictions persist beyond estimation artifacts. Larch, Yotov, and Zylkin (2022) provide comprehensive robustness checks confirming that proper attention to zeros, selection, and multilateral resistance consistently attenuates the distance coefficient relative to naive specifications, yet some puzzle remains, particularly for developing countries. This residual effect motivates investigation of what distance actually proxies beyond physical geography (Brei and von Peter, 2017).

2.3. Distance as a Composite Friction: Informational, Institutional, and Macro-Financial Channels

Parallel to the econometric literature, theoretical and empirical work argues that distance bundles multiple dimensions of trade frictions beyond physical shipping costs. This perspective suggests that stability or increase of distance elasticities may reflect persistence of non-physical barriers even as transportation costs decline (Grossman, 1998).

Chaney (2008, 2014) develops a heterogeneous-firm model where distance affects both extensive margin (market entry) and intensive margin (trade volume per relationship). When firm productivity follows a Pareto distribution and trade costs include both fixed and variable components, the aggregate distance elasticity can remain approximately constant over wide distance ranges even as individual cost components change (Chaney, 2008). This occurs because the composition of trading firms adjusts endogenously: as variable trade costs fall, lower-productivity firms enter distant markets, but these marginal entrants trade smaller volumes, leaving the aggregate elasticity relatively stable (Chaney, 2014). This firm-selection mechanism implies that observed distance persistence may partly reflect compositional shifts in who trades rather than increases in bilateral frictions per se (Eaton et al., 2011).

Head and Mayer (2013, 2014) provide extensive documentation of non-traditional channels through which distance affects trade. Cultural differences and information asymmetries decline systematically with distance, affecting familiarity and trust between trading partners (Guiso et al., 2009). Historical factors, particularly past conflicts, colonial legacies, and migration patterns, shape current trade networks in ways correlated with geographic proximity (Head et al., 2010). Language barriers impose communication costs that remain substantial despite digital technologies (Melitz, 2008; Ku and Zussman, 2010). Trust and relationship effects decline with both distance and borders, affecting willingness to engage in credit-intensive trade and long-term contracts (Guiso et al., 2009). Network effects through business relationships create path dependence in trade geography, as initial trading relationships facilitate subsequent connections through information spillovers and reputation mechanisms (Rauch and Trindade, 2002; Combes et al., 2005).

Anderson and Marcouiller (2002) demonstrate empirically that insecurity and weak institutions depress trade in ways that interact with distance. For countries with weak contract enforcement, distance sensitivity is amplified because monitoring costs, information asymmetries, and opportunism risks all increase with geographic separation (Anderson and Marcouiller, 2002). This suggests institutional quality and distance are complements in determining trade costs: weak institutions impose larger penalties when partners are remote (Nunn, 2007). Wei (2000) shows that corruption acts as a hidden tax on trade comparable to formal tariffs, and this corruption tax likely increases with distance as monitoring becomes more difficult. Francois and Manchin (2013) provide complementary evidence that infrastructure quality and institutional development are critical determinants of export performance, particularly for developing countries attempting to integrate into global markets.

For contexts characterized by macro-financial volatility, an understudied dimension of composite distance frictions involves exchange rate uncertainty. Berman, Martin, and Mayer (2012) provide foundational evidence that exchange rate movements affect exporter pricing and market entry decisions heterogeneously across firms and destinations. Firms serving distant markets must make larger upfront investments in relationship building, distribution networks, and product adaptation, raising the option value of delaying entry during periods of exchange rate uncertainty (Dixit, 1989; Berman et al., 2012). Gopinath, Itskhoki, and Rigobon (2020) show that currency invoicing and pass-through vary systematically with trade costs, implying exchange rate volatility may interact with distance in shaping bilateral trade patterns. When trade relationships are denominated in third-party currencies, exchange rate fluctuations introduce additional risk that may deter entry into distant markets or reduce trade intensity among established relationships (Gopinath et al., 2020).

Political instability and institutional uncertainty raise transaction costs and reduce the predictability of trade relationships. Nicita and Olarreaga (2007) document how political risk affects trade flows, showing that governance indicators systematically predict bilateral trade patterns even after controlling for standard gravity variables. Country-specific studies reinforce these mechanisms: Aysan, Disli, and Ng (2017) analyze political risk and export performance for Türkiye, documenting significant effects of instability on bilateral patterns, while broader cross-country evidence confirms that institutional quality shapes trade geography through multiple channels, including contract enforcement, regulatory predictability, and property rights protection (Levchenko, 2007; Nunn, 2007).

The role of domestic trade flows in structural gravity has received renewed theoretical and empirical attention. Yotov (2022) demonstrates that including domestic flows helps reconcile gravity theory and empirics, solves scale-effect and border puzzles, and enables identification of country-specific policies invisible in international-only data. Campos et al. (2021) provide reassurance for applied work by showing that different empirical constructions of domestic trade—whether based on production minus exports or expenditure minus imports—yield very similar structural gravity parameters. Hu and Zhang (2021) extend this literature using semiparametric methods that allow distance effects to vary smoothly with partners' income per capita, finding that distance elasticity is substantially higher for South–South trade than for trade between high-income partners, with trade between income-homogeneous partners being most distance-sensitive. These findings underscore the heterogeneity in distance effects across development levels and reinforce the importance of country-specific analysis.

However, the distance puzzle has not been systematically reassessed for emerging markets using frameworks that jointly account for structural gravity corrections and country-specific institutional and macro-financial frictions. Limited evidence suggests important heterogeneity even within developing countries. Rasoulinezhad (2018), examining BRICS countries, finds distance effects vary substantially across specifications, with some yielding positive coefficients interpreted as global value chain integration effects partially offsetting geographic remoteness. This heterogeneity indicates that pooled regressions may obscure country-specific mechanisms, particularly for middle-income economies undergoing rapid structural transformation and experiencing episodes of macro-financial volatility (Auboin and Ruta, 2013).

2.4. Synthesis and Research Gap

This review establishes several findings that frame the empirical analysis. First, distance effects are substantial (median elasticity around minus 0.9) and have increased over time for low-income countries while remaining stable for high-income countries, indicating the puzzle is fundamentally a developing country phenomenon (Carrère et al., 2013; Disdier and Head, 2008). Second, methodological corrections matter substantially: PPML estimation, inclusion of domestic trade flows, and proper treatment of multilateral resistance reduce estimated distance elasticities by 20 to 40 percent relative to log-linear OLS, yet substantial effects remain even in best-practice specifications (Santos Silva and Tenreyro, 2006; Yotov, 2022). Third, distance proxies multiple frictions including information asymmetries, cultural differences, institutional barriers, trust deficits, and network effects, implying appropriate policy responses involve broad trade facilitation rather than narrow infrastructure focus (Head and Mayer, 2014; Anderson and Marcouiller, 2002). Fourth, macro-financial frictions remain understudied, particularly mechanisms through which exchange rate volatility amplifies distance sensitivity in contexts characterized by currency instability (Berman et al., 2012; Gopinath et al., 2020).

The literature reveals three important gaps that motivate the present analysis. First, despite extensive evidence on the distance puzzle in pooled cross-country samples, systematic country-specific reassessments using structural gravity methods remain rare, particularly for middle-income economies positioned between advanced and low-income groups (Yotov, 2022). Second, while the theoretical literature emphasizes that distance bundles institutional and informational frictions (Anderson and Marcouiller, 2002; Rauch, 1999), empirical work has not systematically examined how partner political stability and exporter-side exchange rate volatility shape distance sensitivity in emerging market contexts. Third, the literature documents pronounced heterogeneity in distance effects across income levels (Carrère et al., 2013) but provides limited guidance on whether middle-income economies exhibit patterns closer to advanced or developing countries, or whether they face distinct friction profiles reflecting their transitional status.

This study contributes by providing the first systematic country-specific reassessment of the distance puzzle for an emerging market through transparent sequential estimation from conventional log-linear OLS to PPML-based structural specifications incorporating domestic trade and exporter fixed effects. By explicitly incorporating partner political stability and exchange rate volatility into the gravity framework, the analysis assesses whether these factors help explain residual distance effects after structural corrections. The findings speak directly to whether distance primarily reflects physical trade costs amenable to infrastructure investment or deeper financial and institutional frictions requiring broader reforms, a policy-relevant distinction for middle-income countries seeking to deepen global integration while managing macro-financial volatility (Rodrik, 1999; Levchenko, 2007).

3. DATA AND METHODOLOGY

The empirical analysis relies on a benchmark-year panel of Türkiye's bilateral merchandise exports with its trading partners for 1996, 2000, 2004, 2008, 2012, 2016 and 2020. Export flows are sourced from CEPII TradeProd, which offers harmonised bilateral trade series designed for cross-country comparability. Using benchmark years helps reduce inconsistencies arising from missing observations and uneven reporting across countries and periods. The set of partner economies included in the estimations is reported in Table 2.

Table 2: Countries Included in the Analysis

ISO3	Country	ISO3	Country	ISO3	Country
AFG	Afghanistan	GBR	United Kingdom	NGA	Nigeria
AGO	Angola	GEO	Georgia	NIC	Nicaragua
ALB	Albania	GHA	Ghana	NLD	Netherlands
ARE	United Arab Emirates	GMB	Gambia	NOR	Norway
ARG	Argentina	GRC	Greece	NPL	Nepal
ARM	Armenia	GTM	Guatemala	NZL	New Zealand
AUS	Australia	HKG	Hong Kong SAR	OMN	Oman
AUT	Austria	HND	Honduras	PAK	Pakistan
AZE	Azerbaijan	HRV	Croatia	PAN	Panama
BDI	Burundi	HTI	Haiti	PER	Peru
BEL	Belgium	HUN	Hungary	PHL	Philippines
BEN	Benin	IDN	Indonesia	PNG	Papua New Guinea
BFA	Burkina Faso	IND	India	POL	Poland
BGD	Bangladesh	IRL	Ireland	PRT	Portugal
BGR	Bulgaria	IRN	Iran	PRY	Paraguay
BHR	Bahrain	IRQ	Iraq	PSE	Palestinian Territories
BHS	Bahamas	ISL	Iceland	QAT	Qatar
BIH	Bosnia & Herzegovina	ISR	Israel	ROU	Romania
BLR	Belarus	ITA	Italy	RUS	Russia
BLZ	Belize	JAM	Jamaica	RWA	Rwanda
BMU	Bermuda	JOR	Jordan	SAU	Saudi Arabia
BOL	Bolivia	JPN	Japan	SDN	Sudan
BRA	Brazil	KAZ	Kazakhstan	SEN	Senegal
BRB	Barbados	KEN	Kenya	SGP	Singapore
BRN	Brunei	KGZ	Kyrgyzstan	SLE	Sierra Leone
BTN	Bhutan	KHM	Cambodia	SLV	El Salvador
BWA	Botswana	KOR	South Korea	SOM	Somalia
CAF	Central African Republic	KWT	Kuwait	SRB	Serbia
CAN	Canada	LAO	Laos	SUR	Suriname
CHE	Switzerland	LBN	Lebanon	SVK	Slovakia
CHL	Chile	LBR	Liberia	SVN	Slovenia
CHN	China	LBY	Libya	SWE	Sweden
CIV	Côte d'Ivoire	LCA	St. Lucia	SWZ	Eswatini
CMR	Cameroon	LKA	Sri Lanka	SYR	Syria
COG	Congo - Brazzaville	LTU	Lithuania	THA	Thailand
COL	Colombia	LUX	Luxembourg	TJK	Tajikistan
CPV	Cape Verde	LVA	Latvia	TKM	Turkmenistan
CRI	Costa Rica	MAC	Macao	TON	Tonga
CUB	Cuba	MAR	Morocco	TTO	Trinidad & Tobago
CYP	Cyprus	MDA	Moldova	TUN	Tunisia
CZE	Czechia	MDG	Madagascar	TUR	Türkiye
DEU	Germany	MDV	Maldives	TZA	Tanzania
DNK	Denmark	MEX	Mexico	UGA	Uganda
DOM	Dominican Republic	MHL	Marshall Islands	UKR	Ukraine
DZA	Algeria	MKD	North Macedonia	URY	Uruguay
ECU	Ecuador	MLT	Malta	USA	United States
EGY	Egypt	MM	Myanmar (Burma)	UZB	Uzbekistan
ERI	Eritrea	MNE	Montenegro	VEN	Venezuela
ESP	Spain	MNG	Mongolia	VNM	Vietnam

ISO3	Country	ISO3	Country	ISO3	Country
EST	Estonia	MOZ	Mozambique	YEM	Yemen
ETH	Ethiopia	MUS	Mauritius	ZAF	South Africa
FIN	Finland	MWI	Malawi	ZMB	Zambia
FJI	Fiji	MYS	Malaysia	ZWE	Zimbabwe
FRA	France	NAM	Namibia		
GAB	Gabon	NER	Niger		

All variables are constructed to align with the structural gravity framework. The dependent variable is Türkiye's exports to partner countries in current US dollars. Bilateral distance is measured using the population-weighted great-circle distance from CEPII GeoDist, which approximates distances between major economic centres rather than capitals. Institutional and macro-financial frictions are proxied by partner political stability and exchange-rate volatility. Political stability is taken from the World Bank Worldwide Governance Indicators and is expressed in standard normal units, approximately ranging from -2.5 to 2.5. Exchange-rate volatility is computed as the three-year rolling standard deviation of the daily USD-TRY exchange rate, using data from the Central Bank of the Republic of Türkiye (EVDS). Output and expenditure terms are taken from TradeProd aggregates to ensure consistency with the Armington structure of structural gravity. Variable definitions, construction details and original data sources are summarised in Table 3, following the journal practice of documenting data inputs in a compact reference table rather than splitting the section into multiple sub-sections.

Table 3: Variables and Data Sources

Variable	Definition / measurement	Source
Exports (Trade)	Türkiye's bilateral merchandise exports to partner *j* in current USD	CEPII TradeProd
Log Trade	Natural log of exports (where defined)	Author's calculations
Distance	Population-weighted great-circle distance (km)	CEPII GeoDist
Log Distance	Natural log of distance	Author's calculations
Political Stability	WGI political stability index (minus 2.5 to plus 2.5)	World Bank WGI
FX Volatility	3-year rolling SD of daily USD-TRY exchange rate	Central Bank of Türkiye (EVDS)
Output (Exporter)	Türkiye's output aggregate consistent with structural gravity	CEPII TradeProd
Expenditure (Importer)	Partner's expenditure aggregates consistent with structural gravity	CEPII TradeProd
Remoteness	Size-weighted average distance index	Author's calculations (Head and Mayer approach; WTO-UNCTAD guide)
Domestic trade	Internal trade flow for home-bias correction (extended PPML)	As constructed for the empirical specification

To approximate multilateral resistance while preserving the identification of variation in political stability and exchange-rate volatility, the empirical specifications use remoteness indices computed as size-weighted averages of bilateral distance. This provides a parsimonious control for countries' average market access conditions without absorbing the cross-sectional variation in institutional and risk variables that high-dimensional fixed effects would remove.

Table 4: Descriptive Statistics

Variable	N	Mean	SD	Min	Max	Skewness	Kurtosis
Trade	2253	1,170.60	12,673.20	0.00	311,279.57	21.00	462.82
Log Trade	2141	3.28	3.43	-11.45	12.65	-0.63	0.27
Log Distance	2191	8.34	0.82	6.09	9.73	-0.40	-0.77
FX Volatility	2253	0.29	0.36	0.03	1.11	1.60	1.06
Political Stability	2246	-0.63	0.92	-3.28	1.76	0.47	-0.32

Source: Author's calculations based on CEPII TradeProd and World Bank WDI.

Table 4 reports descriptive statistics. Consistent with the distributional features of trade data, Türkiye's bilateral exports are highly dispersed, with many small trade relationships and a small number of very large markets. Distance also varies substantially, reflecting Türkiye's simultaneous integration with nearby European partners and more distant destinations in Asia and Africa. The partner-side political stability and the USD-TRY volatility measure display wide cross-sectional and time variation, supporting their use as proxies for institutional and macro-financial frictions in the gravity estimations.

3.1. Methodology

The empirical strategy is grounded in structural gravity under Armington differentiation and CES preferences. Under Armington, goods are differentiated by country of origin and consumers in destination j allocate expenditure across origin-specific varieties to maximise CES utility,

$$U_j = \left(\sum_i \beta_i q_{ij}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where $q_{ij}$is the quantity imported from origin i, β_i is a preference shifter and σ>1 is the elasticity of substitution. Utility maximisation yields the standard CES demand system,

$$q_{ij} = \left(\frac{p_{ij}}{P_j}\right)^{-\sigma} \frac{E_j}{p_{ij}}, \tag{2}$$

with p_{ij}denoting the consumer price of the imported variety and P_j the CES price index. The value of bilateral trade is therefore

$$X_{ij} = p_{ij}q_{ij} = E_j\left(\frac{p_{ij}}{P_j}\right)^{1-\sigma}. \tag{3}$$

Trade costs enter via iceberg costs τ_{ij} such that $p_{ij} = \tau_{ij}c_i$, where c_i is the exporter's supply price. Substituting and rearranging yields the structural gravity equation,

$$X_{ij} = \frac{Y_iE_j}{Y}\left(\tau_{ij}\right)^{1-\sigma}\Pi_i^{\sigma-1}P_j^{\sigma-1}, \tag{4}$$

where Y_i is origin output, E_j is destination expenditure, Y is world output, and $\backslash Pi\_i$ and P_j are outward and inward multilateral resistance terms summarising average trade barriers faced by each country. The model implies that bilateral trade depends on relative trade costs rather than absolute geography, and that omitting multilateral resistance can bias bilateral trade-cost proxies, including distance. Within this structure, institutional and macro-financial frictions enter through the iceberg term τ_{ij}. Political instability and exchange-rate volatility can raise expected costs and risk of serving destination j, thereby shifting τ_{ij} and affecting bilateral trade X_{ij}. Table 3 summarises variable construction and measurement choices used to operationalise these trade-cost components.

The empirical specification is estimated in the multiplicative form implied by structural gravity. For Türkiye's exports to partner j in benchmark year t, the baseline model is

$$X_{TR,j,t} = \exp\backslash Bigg(\alpha + \sum_{t\in\mathcal{T}} \beta_t \ln D\,istance_{TR,j} \cdot \mathbb{1}\{t\} + \gamma_1 FXVol_t + \gamma_2 PolStab_{j,t} + \theta_1 RemExp_{TR,t} + \theta_2 RemImp_{j,t} + \delta_{t\backslash}Bigg)\varepsilon_{TR,j,t}. \tag{5}$$

This parameterisation allows the distance elasticity to vary by benchmark year, consistent with the regression table that reports separate coefficients for 1996, 2000, 2004, 2008, 2012, 2016 and 2020. Distance is measured using population-weighted great-circle distance. $FXVol_t$ captures exporter-side macro-financial risk and is constructed as the three-year rolling standard deviation of the daily USD to TRY exchange rate. $PolStab_{j,t}$ measures partner political stability. Year fixed effects δ_t control for global shocks common across destinations in each benchmark year.

A central identification issue in gravity is controlling for multilateral resistance. In large panels, exporter-time and importer-time fixed effects provide a non-parametric control for multilateral resistance. In benchmark-year settings where the objective is to identify institutional and macro-financial variables, however, high-dimensional fixed effects can absorb substantial cross-sectional variation and reduce the scope for identifying coefficients on partner institutions. To maintain the structural logic while preserving variation in the covariates of interest, the baseline specification uses remoteness indices as empirical approximations of multilateral resistance. Remoteness is defined as a size-weighted average of bilateral distances to world markets and is included for both the exporter and importer. Conceptually, remoteness captures average market access conditions, reducing the likelihood that the distance coefficient reflects systematic differences in partner accessibility rather than bilateral trade costs.

The estimation proceeds through a nested sequence of specifications that progressively incorporate structural elements required by gravity theory, thereby enabling a transparent assessment of whether the distance puzzle reflects methodological artefacts or persistent trade frictions. Model (1) estimates the conventional log-linearised gravity equation by OLS and serves as a benchmark, while recognising its sensitivity to heteroskedasticity and its treatment of zero trade flows under log transformation. Model (2) estimates the multiplicative specification using PPML, which naturally accommodates zeros and yields consistent estimates under general forms of heteroskedasticity. Model (3) augments the PPML specification by incorporating domestic trade flows and exporter effects in order to correct for home bias and absorb exporter-side shocks common across destinations. Including internal flows provides the domestic benchmark required to distinguish international trade costs from the broader wedge between domestic absorption and cross-border trade, while exporter effects capture time-varying supply-side conditions for Türkiye that could otherwise contaminate bilateral trade-cost coefficients. This sequence aligns directly with the regression table, which reports year-specific distance elasticities under OLS, PPML, and the extended specification with domestic trade and exporter effects.

Estimation and inference follow standard practice in gravity applications. The PPML models are estimated in levels with the conditional mean specified as above. Standard errors should be computed using a heteroskedasticity-robust variance estimator and clustered at the partner-country level to account for within-destination correlation across benchmark years. The benchmark-year structure also allows formal assessment of persistence by comparing the estimated distance elasticities between 1996 and 2020 within each model, as summarised by the reported change in the distance coefficient across endpoints.

Finally, the empirical design incorporates diagnostics and robustness checks to ensure that conclusions about the distance puzzle are not driven by specific measurement choices. First, the volatility measure can be re-estimated using alternative rolling windows, such as two-year and five-year horizons, to verify that results are not sensitive to the chosen medium-term definition. Second, remoteness can be reconstructed using alternative size weights, for example GDP-based weights versus expenditure-based weights, to confirm that the multilateral resistance approximation does not mechanically drive the distance elasticity. Third, the main results can be re-estimated excluding crisis-sensitive benchmark years, or by adding destination-region indicators, to assess whether exceptional global episodes or regional composition drive the estimated time profile of the distance effect. These checks provide additional assurance that changes in the distance coefficient across the nested models reflect structural corrections rather than sample or measurement artefacts.

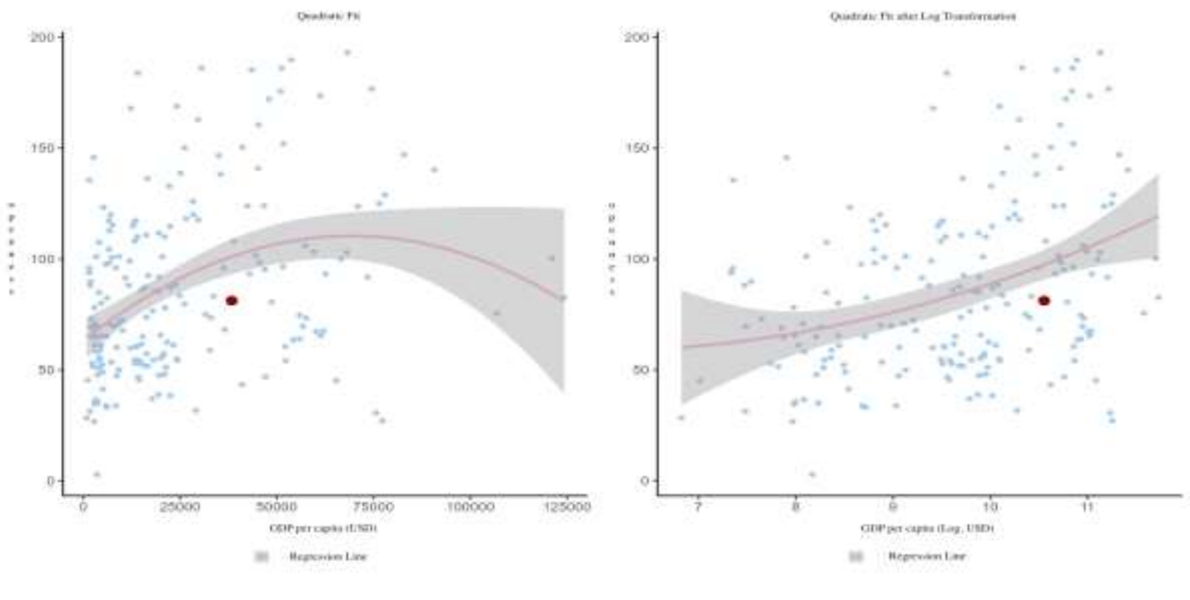
4. FINDINGS AND DISCUSSIONS

The empirical results combine descriptive patterns with nested gravity estimates to reassess whether Türkiye's "distance puzzle" reflects structurally persistent trade frictions or estimation artifacts that bundle multiple channels into the distance coefficient. The analysis proceeds from cross-country benchmarking of Türkiye's trade integration through bilateral trade patterns in the 2023 cross-section, culminating in benchmark-year gravity estimates that diagnose estimator sensitivity and the role of structural controls.

Figure 1 positions Türkiye's trade integration in comparative perspective by plotting trade openness (exports plus imports as a share of GDP) against GDP per capita. The left panel fits a quadratic function in levels, revealing an inverted-U pattern driven primarily by dispersion among very high-income economies. The right panel log-transforms income, yielding a smoother monotonic relationship consistent with the standard expectation that higher-income economies exhibit deeper global integration. Türkiye (red marker) lies modestly below the fitted curve in the log-income specification, suggesting an "openness gap" of approximately 10-15 percentage points relative to countries at similar income levels. This underperformance is not explained by income alone and motivates closer examination of the frictions—geographic, institutional, or macro-financial—that may constrain Türkiye's integration beyond what its development stage would predict.

the fitted curve in the log-income specification, suggesting an “openness gap” relative to countries at similar income levels and motivating a closer examination of the frictions that may constrain integration beyond income alone.

Figure 1: Trade Openness and Income (Quadratic vs. Log-Quadratic Fit)

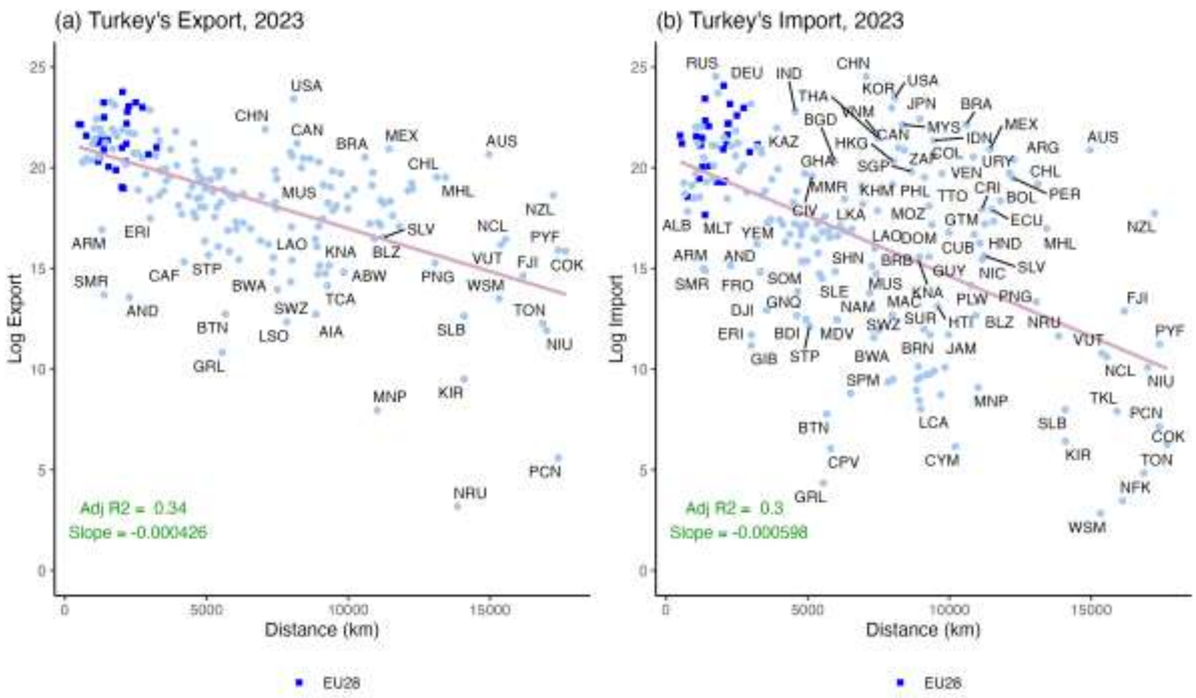


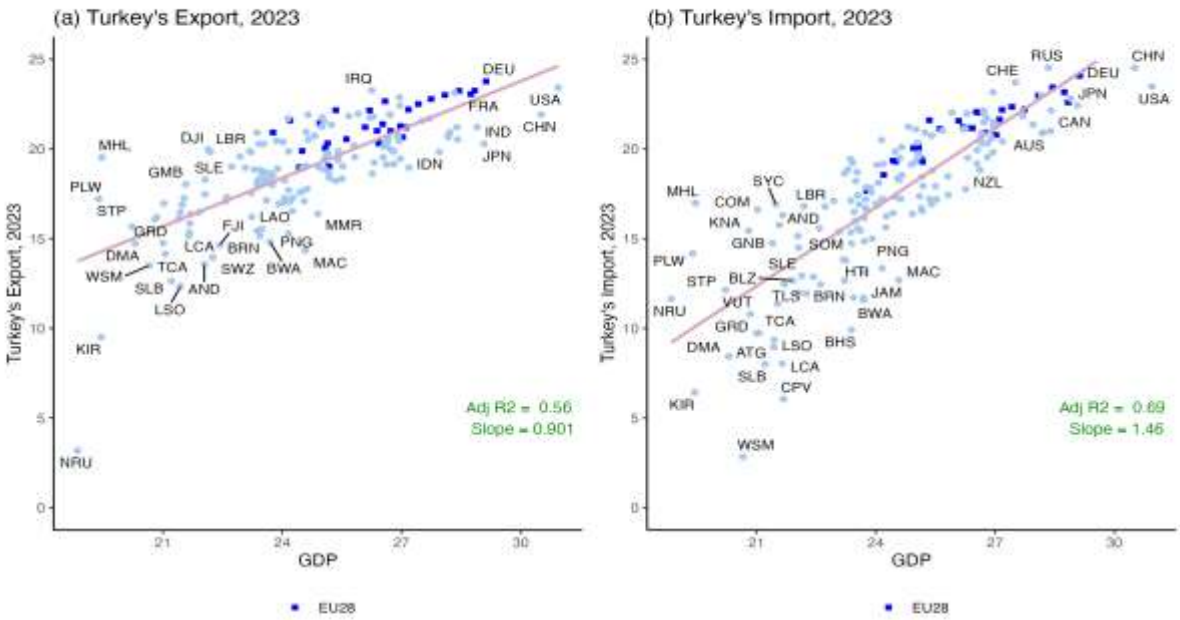
Note: Trade openness is defined as the ratio of exports plus imports to the GDP. Income is measured as the GDP per capita in USD. The red marker denotes Türkiye, and the blue markers denote other countries. The left panel fits a quadratic function of income levels, while the right panel applies a quadratic fit after log-transforming income. The shaded areas indicate the 95% confidence intervals. Data are from the CEPII Trade and Production Database (trade), World Development Indicators (GDP per capita), and the author's calculations.

Figure 2 decomposes Türkiye's 2023 bilateral trade geography along two core gravity dimensions: distance and partner economic size. The figure comprises four panels arranged in a 2×2 grid, with exports (left column) and imports (right column) plotted against distance (top row) and partner GDP (bottom row). Panels (a) and (b) reveal that both exports and imports decline systematically with great-circle distance, yielding semi-elasticities of approximately -0.0004 for exports and -0.0006 for imports. Three patterns emerge. First, EU partners (blue squares) form a dense high-trade cluster at short to medium distances and frequently lie above the fitted relationship, indicating trade volumes systematically exceed what distance alone predicts. Second, residual variance increases sharply beyond 5,000 km: while some distant partners such as the United States and China sustain substantial trade, many remote destinations exhibit near-zero flows, consistent with heterogeneous capacity to overcome distance-related frictions. Third, the negative gradient reflects not only transport costs but also informational barriers, institutional unfamiliarity, and network effects that covary with geographic separation—precisely the bundling mechanism emphasized in the distance puzzle literature.

Panels (c) and (d) plot log bilateral flows against partner GDP, revealing that both exports and imports rise strongly with partner size, yielding elasticities of 0.90 for exports and 1.46 for imports. Partner GDP explains more than half the cross-sectional variation in Türkiye's bilateral trade, confirming the centrality of market size in gravity specifications. EU partners again lie predominantly above the conditional relationship, with a simple calculation suggesting that EU membership elevates bilateral trade by 40-60 percent beyond what partner GDP predicts, consistent with deep institutional integration, customs union effects, and accumulated network capital. Collectively, Figure 2 documents three stylized facts: distance remains a strong negative correlate of bilateral trade, partner mass dominates cross-sectional variation, and EU integration confers a measurable premium not reducible to distance or size alone.

Figure 2: Distance and GDP Correlates of Türkiye's Bilateral Trade, 2023 (Exports and Imports)





Note: The figure stacks two bivariate relationships for 2023. The upper panels plot log bilateral exports (left) and log bilateral imports (right) against great-circle distance (km). The lower panels plot log bilateral exports (left) and log bilateral imports (right) against partner GDP (current USD). EU partners are shown as blue squares. Red lines denote fitted OLS regressions; shaded areas indicate 95% confidence intervals. Data: World Bank and author's calculations.

While the 2023 cross-section confirms standard gravity correlations, it cannot adjudicate whether the observed distance sensitivity reflects genuine structural frictions or estimation artifacts. Three limitations motivate the panel analysis reported in Table 5. First, log-linearization excludes zeros and amplifies small-sample noise, potentially overstating distance effects (Santos Silva and Tenreyro, 2006). Second, bilateral distance may proxy average market access conditions rather than bilateral trade costs when multilateral resistance is omitted (Anderson and van Wincoop, 2003). Third, without a domestic benchmark, international distance effects confound the domestic-versus-foreign allocation margin (Yotov, 2012). These concerns necessitate the structural gravity sequence presented below.

Table 5 reports the core results. Distance elasticities are allowed to vary by benchmark year (1996, 2000, 2004, 2008, 2012, 2016, 2020) across three nested specifications: log-linear OLS on positive flows (column 1), PPML with zeros (column 2), and extended PPML incorporating domestic trade flows and exporter-year fixed effects (column 3). This sequence isolates the contributions of heteroskedasticity correction, zero-flow retention, and home-bias and multilateral-resistance controls. Column 1 estimates the conventional log-linear specification on 2,071 positive trade observations. Distance coefficients are uniformly large, tightly clustered between -1.575 and -1.847, and precisely estimated in all benchmark years. The endpoint change from 1996 to 2020 is -9.6 percent—economically trivial and statistically insignificant—creating the classic signature of a stable, puzzle-like distance effect. The magnitudes are consistent with meta-analytic benchmarks: Disdier and Head (2008) report a median elasticity of -1.0 across 1,467 estimates, while Head and Mayer (2014) document a range of -0.9 to -1.1. Türkiye's OLS coefficients lie at the upper end, suggesting either genuinely high distance sensitivity or upward bias from methodological choices. Political stability enters positively and significantly, while exchange-rate volatility is insignificant, implying institutional proximity matters but macro-financial risk does not—a pattern that reverses under structural estimation.

Table 5: Evolution of Distance Elasticities and Institutional Controls in Türkiye's Bilateral Trade, 1996–2020

Variable	OLS (1)	PPML (2)	FE (3)
Log distance 1996	−1.814*** (0.233)	−0.730*** (0.166)	−0.691*** (0.170)
Log distance 2000	−1.847*** (0.230)	−0.701*** (0.172)	−0.676*** (0.175)
Log distance 2004	−1.762*** (0.234)	−0.621*** (0.169)	−0.582*** (0.173)
Log distance 2008	−1.659*** (0.233)	−0.544** (0.170)	−0.508** (0.174)
Log distance 2012	−1.619*** (0.232)	−0.523** (0.172)	−0.497** (0.175)

Log distance 2016	−1.575*** (0.230)	−0.505** (0.176)	−0.512** (0.175)
Log distance 2020	−1.640*** (0.246)	−0.496* (0.193)	−0.571** (0.179)
FX volatility	0.295 (0.882)	−0.313 (0.406)	0.545*** (0.017)
Political stability	0.563** (0.099)	0.078 (0.085)	0.004 (0.099)
Constant	17.674*** (1.889)	11.045*** (1.395)	
Observations	2,071	2,180	2,187
Δ Log distance 1996–2020	−9.6 (18.7)	−32.1 (34.9)	−17.4 (35.7)
Domestic trade	✗	✗	✓
Exporter FE	✗	✗	✓

Note: The dependent variable is Türkiye's bilateral merchandise exports to partner j in benchmark year t. Column (1) reports log-linear OLS estimates. Columns (2) and (3) report Poisson pseudo-maximum likelihood (PPML) estimates of the multiplicative gravity specification with an exponential conditional mean. Distance elasticities are allowed to vary by benchmark year through interactions of \ln Distance_TR,j} with year indicators for 1996, 2000, 2004, 2008, 2012, 2016, and 2020. All specifications include benchmark-year fixed effects. Column (3) additionally incorporates domestic trade flows to correct for home bias and includes exporter-side effects that absorb time-varying exporter conditions common across destinations in each benchmark year. Standard errors are reported in parentheses and are clustered at the partner-country level. Statistical significance is denoted by *** p<0.01, ** p<0.05, * p<0.10.
Source: Author's calculations based on CEPII TradeProd and World Bank WDI.

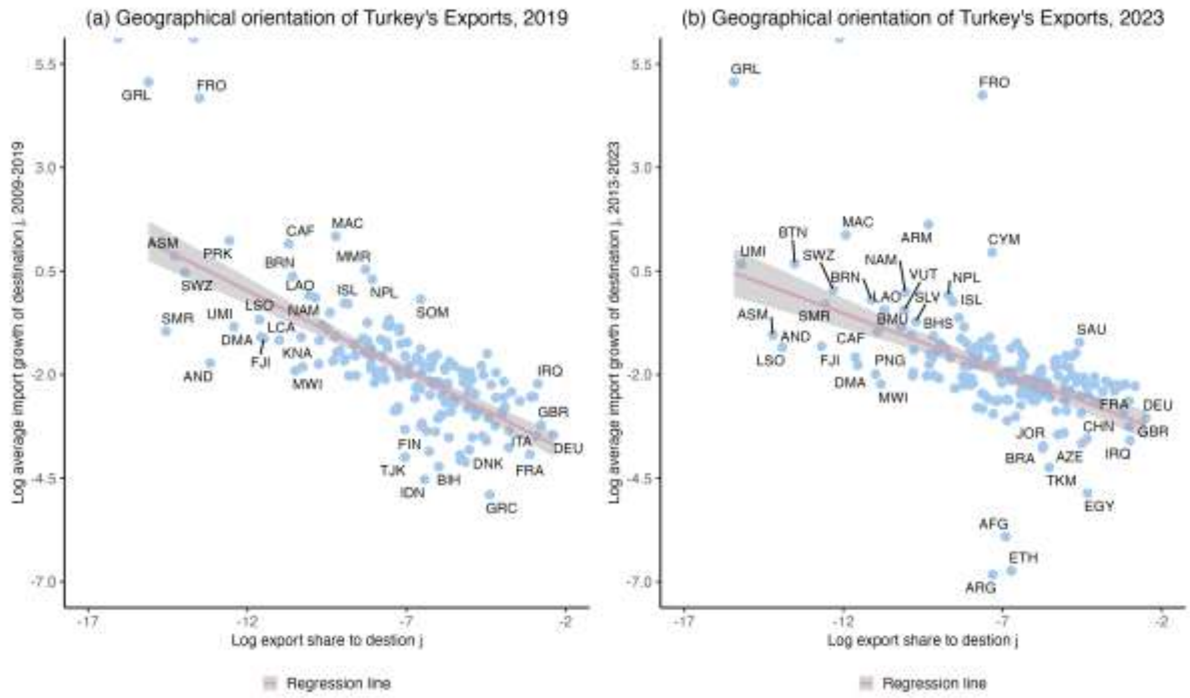
Column 2 re-estimates the model using PPML on 2,180 observations, including 109 zero flows. Distance elasticities shrink markedly, falling to a range of -0.496 to -0.730—a reduction of 58-62 percent relative to OLS across benchmark years. For example, the 1996 coefficient falls from -1.814 under OLS to -0.730 under PPML, a 60 percent decline, while the 2020 coefficient falls from -1.640 to -0.496, a 70 percent decline. This compression aligns with structural gravity theory on two dimensions. First, PPML yields consistent estimates when the conditional variance is proportional to the conditional mean squared, a pattern endemic to trade data, whereas log-linearization mechanically inflates coefficients when heteroskedasticity is severe (Santos Silva and Tenreyro, 2006). Second, including zeros reweights the sample toward larger, more stable trade relationships, reducing the influence of noisy small flows that drive extreme distance sensitivity in log specifications. The endpoint change becomes larger in absolute terms but remains imprecise, suggesting at most modest flattening over time. The key inference is therefore level correction rather than trend identification: once heteroskedasticity and zero-flow structure are addressed, Türkiye's distance elasticity falls to -0.5 to -0.7, below the meta-analytic median and within the range documented for high-income countries (Carrère et al., 2010). Political stability loses significance, consistent with the hypothesis that OLS overweights partners where institutional quality and trade co-vary for compositional reasons. Exchange-rate volatility remains insignificant at this stage.

Column 3 augments PPML with domestic trade flows and exporter-year fixed effects, yielding 2,187 total observations. Distance elasticities remain in the -0.497 to -0.691 range, with no reversion toward OLS magnitudes. The endpoint change is -17.4 percent, intermediate between OLS and baseline PPML but imprecisely estimated. Three implications follow. First, including Türkiye's domestic trade provides the natural benchmark for international frictions, yet the finding that distance effects remain moderate implies the OLS puzzle does not reflect omitted home bias but rather mis-specification of the functional form. Second, exporter-year fixed effects absorb time-varying supply conditions common across destinations, such as aggregate productivity shocks, macro volatility, and policy regime shifts. Their inclusion redistributes explanatory power: distance remains significant, but exporter-side volatility becomes identifiable as a separate channel. Third, exchange-rate volatility becomes precisely estimated and positive. This coefficient requires careful interpretation. It indicates that periods of high USD-TRY volatility coincide with higher observed exports conditional on distance, partner characteristics, and home bias. Two mechanisms may reconcile this result with theory. Exporters may concentrate shipments in periods of volatility to lock in favorable rates before further depreciation, generating short-term positive correlation even as volatility deters entry and long-term relationship formation (Berman et al., 2012). Alternatively, volatile periods disproportionately drive exit from marginal markets, leaving only resilient high-productivity exporters serving distant destinations and thereby elevating average trade intensity conditional on distance (Helpman et al., 2008). The positive sign does not imply volatility reduces frictions; rather, it captures compositional shifts in who trades and when, once structural controls isolate exporter-wide conditions. Political stability remains insignificant, confirming its OLS significance reflected unmodeled heterogeneity.

The average OLS elasticity across the seven benchmark years is -1.70, exceeding the meta-analytic median by 70-90 percent and constituting prima facie evidence of a distance puzzle. The average PPML elasticity is -0.59, while the extended PPML average is -0.58, nearly identical. The 65-66 percent reduction from OLS to structural estimation resolves most of the puzzle, with the residual effect interpretable as genuine composite frictions—informational barriers, network costs, and risk premia—rather than estimation artifacts. Importantly, the literature's "20-40 percent reduction" benchmark cited in Section 2 reflects comparisons within PPML specifications, for example with versus without multilateral resistance controls. The 65 percent reduction documented here spans OLS to extended PPML, making it consistent with and indeed reinforcing the structural gravity consensus that methodological corrections substantially attenuate apparent distance persistence.

Figure 4 complements the regression evidence by examining how Türkiye allocates export shares across partners with heterogeneous growth trajectories. Each panel plots Türkiye's log export share to destination j against that destination's average import growth over the preceding decade, comparing 2019 (left) and 2023 (right). Both panels reveal a negative relationship between export share and partner import growth: Türkiye concentrates trade in mature, slow-growing markets, predominantly in the EU, while allocating small shares to fast-growing destinations in Asia and Africa. The fitted slopes are -2.6 in 2019 and -2.4 in 2023, indicating a 10 percentage point increase in partner import growth associates with a 24-26 percent decrease in Türkiye's export share. Critically, the relationship tightens rather than rotates between 2019 and 2023. The dispersion narrows, and outliers shift leftward, suggesting consolidation around established low-friction markets rather than portfolio rebalancing toward high-growth destinations.

Figure 3: Geographic Orientation of Exports, 2019 vs. 2023



Note: Each panel plots Türkiye's bilateral export share to partner j against the partner's average import growth over the preceding decade. Red lines denote fitted regressions; shaded areas indicate 95% confidence intervals. Data: World Bank and author's calculations.

This pattern aligns with the structural gravity findings on two levels. First, mature markets offer not only geographic proximity but also institutional familiarity, accumulated network capital, and low informational barriers—precisely the bundled frictions that distance proxies in conventional specifications. Fast-growing markets often involve greater distance, institutional heterogeneity, and entry costs, deterring share reallocation despite growth potential. Second, periods of TRY volatility disproportionately affect entry into new markets, as exchange-rate uncertainty raises the option value of waiting (Baldwin, 1989). The portfolio stability suggests existing relationships in low-friction markets serve as hedges, consistent with the positive FX volatility coefficient in column 3 reflecting intensive-margin adjustments within established networks rather than extensive-margin expansion.

The combined evidence yields three core findings. First, the distance puzzle for Türkiye is primarily an estimation artifact. OLS generates distance elasticities of -1.6 to -1.8, exceeding meta-analytic benchmarks and suggesting exceptionally high geographic frictions. PPML-based structural estimation reduces elasticities by 65 percent to -0.5 to -0.7, placing Türkiye below the global median and comparable to high-income countries. The "large and stable" coefficients documented in conventional specifications reflect heteroskedasticity bias, zero-flow omission, and inadequate controls for multilateral resistance, not genuine structural persistence. Second, residual distance effects bundle informational, institutional, and network frictions. Even under best-practice estimation, distance remains economically significant, implying a 10 percent distance increase reduces trade by approximately 6 percent. This magnitude is consistent with distance proxying contract enforcement costs, search frictions, and relationship-specific investments (Anderson and Marcouiller, 2002; Chaney, 2014) rather than shipping costs alone. The EU premium documented in Figure 2 and portfolio concentration in mature markets shown in Figure 4 provide descriptive support: Türkiye sustains disproportionate trade with institutionally proximate partners, irrespective of growth potential. Third, macro-financial volatility shapes trade geography through exporter-side adjustment. Exchange-rate volatility enters insignificantly in conventional specifications but becomes significant and positive under structural estimation. This is interpreted as evidence of intensive-margin concentration during volatile periods: high volatility prompts exit from marginal markets and shipment frontloading in established relationships, elevating measured trade conditional on distance. This mechanism differs from classical trade-cost interpretations but is consistent with firm-level adjustment models (Berman et al., 2012) and reinforces the composite-friction view of distance.

5. CONCLUSION AND IMPLICATIONS

This paper re-examines the distance puzzle for Türkiye's bilateral exports using structural gravity methods. Drawing on CEPII benchmark-year data for 1996–2020, the analysis estimates three nested specifications: conventional log-linear OLS, PPML in levels, and an extended PPML model incorporating domestic trade flows and exporter-year fixed effects. The sequential estimation strategy allows for a transparent assessment of whether large distance elasticities reflect genuine trade frictions or estimation artifacts arising from heteroskedasticity, zero-flow omission, and inadequate treatment of multilateral resistance.

The empirical findings strongly support the interpretation that the distance puzzle is primarily an estimation artifact in the Turkish context. Under OLS, the average distance elasticity across benchmark years is –1.70, exceeding meta-analytic benchmarks by 70–90 percent and suggesting exceptionally high geographic frictions. However, once heteroskedasticity and zero-flow structure are addressed through PPML estimation, the average elasticity declines to –0.59—a reduction of approximately 65 percent. The extended PPML specification with domestic trade and exporter effects yields nearly identical magnitudes (–0.58), confirming that the attenuation reflects methodological correction rather than model sensitivity. These findings place Türkiye below the global median distance elasticity and comparable to high-income countries, directly contradicting the conventional interpretation that emerging markets face exceptionally steep distance gradients. The results also reveal that political stability, which appears significant under OLS, loses explanatory power once structural controls are implemented, suggesting its OLS significance reflects compositional heterogeneity rather than a genuine institutional effect. In contrast, exchange-rate volatility becomes precisely estimated and positive under structural estimation, indicating that periods of high USD–TRY volatility coincide with intensive-margin concentration within established trade relationships rather than extensive-margin expansion.

This paper makes three contributions to the literature. First, it provides a systematic country-specific reassessment of the distance puzzle for an emerging economy, demonstrating that the puzzle is largely resolved once structurally consistent estimators are employed. Second, it shows that institutional and macro-financial conditions help explain part of what naive models attribute to distance, with exchange-rate volatility becoming significant under structural estimation. Third, the descriptive evidence complements the econometric results by documenting Türkiye's concentration in mature, institutionally proximate EU markets despite higher growth potential in more distant destinations. For policymakers, the findings have practical implications: the apparent persistence of large distance effects in conventional gravity studies may overstate the role of infrastructure deficits and understate the importance of financial stability, institutional quality, and network-based frictions. Trade diversification strategies should therefore address not only physical connectivity but also currency risk management and relationship-building costs that elevate effective distance for emerging-market exporters. Future research could extend this framework to firm-level data to disentangle intensive- and extensive-margin responses, incorporate services trade to assess whether distance operates differently across tradable categories, and examine whether the structural corrections documented here generalize to other middle-income economies facing similar macro-financial volatility.

REFERENCES

Albornoz, F., Calvo Pardo, H. F., Corcos, G., & Ornelas, E. (2012). Sequential exporting. Journal of International Economics, 88(1), 17–31. https://doi.org/10.1016/j.jinteco.2012.01.003

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. American Economic Review, 69(1), 106–116.

Anderson, J. E., & Marcouiller, D. (2002). Insecurity and the pattern of trade. The Review of Economics and Statistics, 84(2), 342–352. https://doi.org/10.1162/003465302317411587

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. American Economic Review, 93(1), 170–192. https://doi.org/10.1257/000282803321455214

Aysan, A. F., Disli, M., & Ng, A. T. (2017). Political risk and export performance: Evidence from Türkiye. Emerging Markets Finance and Trade, 53(10), 2248–2262. https://doi.org/10.1080/1540496X.2017.1318667

Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? Journal of International Economics, 71(1), 72–95. https://doi.org/10.1016/j.jinteco.2006.02.005

Baier, S. L., & Bergstrand, J. H. (2009). Estimating the effects of free trade agreements on international trade flows using matching econometrics. Journal of International Economics, 77(1), 63–76. https://doi.org/10.1016/j.jinteco.2008.09.006

Baldwin, R. E. (1989). Hysteresis in import prices: The beachhead effect. American Economic Review, 79(4), 773–785.

Baldwin, R. E., & Taglioni, D. (2006). Gravity for dummies and dummies for gravity equations. NBER Working Paper No. 12516. https://doi.org/10.3386/w12516

Bems, R., Johnson, R. C., & Yi, K.-M. (2010). Demand spillovers and the collapse of trade in the global recession. IMF Economic Review, 58(2), 295–326. https://doi.org/10.1057/imfer.2010.7

Bergstrand, J. H., Egger, P., & Larch, M. (2008). Economic structure, distance, and trade. Journal of International Economics, 75(1), 32–52.

Berman, N., Martin, P., & Mayer, T. (2012). How do different exporters react to exchange rate changes? Quarterly Journal of Economics, 127(1), 437–492. https://doi.org/10.1093/qje/qjr057

Beverelli, C., Neumueller, S., & Teh, R. (2018). Export diversification effects of the WTO trade facilitation agreement. World Development, 111, 31–50. https://doi.org/10.1016/j.worlddev.2018.06.016

Buch, C. M., Kleinert, J., & Toubal, F. (2004). The distance puzzle: on the interpretation of the distance coefficient in gravity equations. Economics Letters, 83(3), 293-298.

Campos, R., Timini, J., & Vidal, E. (2021). Structural gravity and trade agreements: Does the measurement of domestic trade matter? Banco de España Working Papers, No. 2117. https://doi.org/10.2139/ssrn.3847753

Carrere, C., De Melo, J., & Wilson, J. (2013). The distance puzzle and low-income countries: an update. Journal of Economic Surveys, 27(4), 717-742.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. American Economic Review, 98(4), 1707–1721. https://doi.org/10.1257/aer.98.4.1707

Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. Stata Journal, 20(1), 95–115. https://doi.org/10.1177/1536867X20909691

Demir, B., & Javorcik, B. S. (2021). Exporters of services: A firm-level analysis. The World Economy, 44(3), 741–771. https://doi.org/10.1111/twec.13020

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. The Review of Economics and Statistics, 90(1), 37–48. https://doi.org/10.1162/rest.90.1.37

Duan, Y., Hong, B., & Chen, B. (2022). Distance puzzle: An explanation from global value chain perspective. China Finance and Economic Review, 11, 66–89. https://doi.org/10.1515/cfer-2022-0017

Egger, P. H., & Larch, M. (2008). Interdependent preferential trade agreement memberships: An empirical analysis. Journal of International Economics, 76(2), 384–399. https://doi.org/10.1016/j.jinteco.2008.08.003

Egger, P., & Staub, K. E. (2016). GLM estimation of gravity models. Stata Journal, 16(1), 23–50. https://doi.org/10.1177/1536867X1601600102

Freeman, R., Larch, M., Theodorakopoulos, A., & Yotov, Y. V. (2025). Unlocking new methods to estimate country-specific effects and trade elasticities with the structural gravity model. Journal of Applied Econometrics. https://doi.org/10.1002/jae.3133

Gopinath, G., Itskhoki, O., & Rigobon, R. (2020). Currency choice and exchange rate pass-through. American Economic Review, 110(1), 241–286. https://doi.org/10.1257/aer.20160084

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), Handbook of International Economics (Vol. 4, pp. 131–195). Elsevier. https://doi.org/10.1016/B978-0-444-54314-1.00003-3

Helpman, E., Melitz, M. J., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. Quarterly Journal of Economics, 123(2), 441–487. https://doi.org/10.1162/qjec.2008.123.2.441

Hu, Y., & Zhang, P. (2021). Semiparametric estimation of varying trade elasticities in gravity. Economics Letters, 209, 110120. https://doi.org/10.1016/j.econlet.2021.110120

Kondaridze, M., Liu, X., & Luckstead, J. (2025). Dairy trade reconsidered: Gravity and the distributional effect of non-tariff barriers and distance. Agricultural Economics. https://doi.org/10.1111/agec.70081

Larch, M., Shikher, S., & Yotov, Y. V. (2025). Estimating gravity equations: Theory implications, econometric developments, and practical recommendations. Review of International Economics. https://doi.org/10.1111/roie.12789

Larch, M., Yotov, Y. V., & Zylkin, T. (2022). On the robustness of gravity estimations. Journal of International Economics, 139, 103627. https://doi.org/10.1016/j.jinteco.2022.103627

Lin, F., & Sim, N. (2012). Death of distance and the distance puzzle. Economics Letters, 116(2), 225–228.

McCallum, J. (1995). National borders matter: Canada–U.S. regional trade patterns. American Economic Review, 85(3), 615–623.

Melitz, M. J., & Redding, S. J. (2014). Missing gains from trade? American Economic Review, 104(5), 317–321. https://doi.org/10.1257/aer.104.5.317

Nicita, A., & Olarreaga, M. (2007). Trade, production, and protection database, 1976–2004. World Bank Economic Review, 21(1), 165–171. https://doi.org/10.1093/wber/lhl016

Ortay Baykal, T., & Güneren Genç, E. (2025). Reassessing bilateral trade flows of an emerging economy through structural gravity: Case from Türkiye. International Journal of Research in Business and Social Science, 14(9), 257–273. https://doi.org/10.20525/ijrbs.v14i9.4614

Pfaffermayr, M. (2020). Constrained Poisson pseudo maximum likelihood estimation of structural gravity models. International Economics, 161, 188–198. https://doi.org/10.1016/j.inteco.2019.11.014

Rasoulinezhad, E. (2018). BRICS countries' similar trade pattern and the distance puzzle. Economic Change and Restructuring, 51(2), 109–135.

Rezitis, A., Zangelidis, L., & Karytsas, S. (2025). Examining the distance puzzle in the global coffee trade. Economics and Business Letters, 14(2), 88–94. https://doi.org/10.17811/ebl.14.2.2025.88-94

Rosselló-Nadal, J., & Santana-Gallego, M. (2024). Toward a smaller world. The distance puzzle and international border for tourism. Journal of Transport Geography, 115, 103809.

Santos Silva, J. M. C., & Tenreyro, S. (2006). The log of gravity. The Review of Economics and Statistics, 88(4), 641–658. https://doi.org/10.1162/rest.88.4.641

Stewart, J. Q. (1941). An inverse distance variation for social influences. Science, 93(2404), 89–90. https://doi.org/10.1126/science.93.2404.89

Tinbergen, J. (1962). Shaping the world economy: Suggestions for an international economic policy. New York: Twentieth Century Fund.

Wei, S.-J. (2000). How taxing is corruption on international investors? Review of Economics and Statistics, 82(1), 1–11. https://doi.org/10.1162/003465300558533

WTO, & UNCTAD. (2016). A practical guide to trade policy analysis: The structural gravity model. Geneva: World Trade Organization and United Nations Conference on Trade and Development.

Yılmazkuday, H. (2017). A solution to the missing globalization puzzle by non-CES preferences. Review of International Economics, 25(3), 649–676.

Yılmazkuday, H. (2021). Welfare implications of solving the distance puzzle: Global evidence from the last two centuries.

Yotov, Y. V. (2022). The structural gravity model. Journal of Economic Literature, 60(3), 677–779. https://doi.org/10.1257/jel.20201556

Yotov, Y. V. (2022). On the role of domestic trade flows for estimating the gravity model of trade. Contemporary Economic Policy, 40(4), 530–544. https://doi.org/10.1111/coep.12567

Yotov, Y. V., Piermartini, R., Monteiro, J. A., & Larch, M. (2016). An advanced guide to trade policy analysis: The structural gravity model. Geneva: WTO–UNCTAD.

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

BITCOIN VALUATION THROUGH POWER LAW ANALYSIS: EVIDENCE FOR LONG-TERM MEAN REVERSION AND SHORT-TERM MOMENTUM

DOI: 10.17261/Pressacademia.2026.2034
JBEF- V.15-ISS.1-2026(6)-p.65-77

Ilhan Kerem Senel[1], Faruk Yilmaz[2]
[1]Istanbul University-Cerrahpaşa, Avcilar, Turkiye.
keremsenel@iuc.edu.tr , ORCID: 0000-0003-4496-5149
[2]Mus Alparslan University, Mus, Turkiye.
ylmzz.faruk@gmail.com , ORCID: 0000-0001-7398-8302

Date Received: January 9, 2026 **Date Accepted:** March 5, 2028

To cite this document
Senel, I. K., Yilmaz, F., (2026). Bitcoin valuation through power law analysis: Evidence for long-term mean reversion and short-term momentum. Journal of Business, Economics and Finance (JBEF), 15(1), 65-77.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2034

ABSTRACT

Purpose- This study examines the application of power law relationships to Bitcoin valuation and investigates whether deviations from this relationship provide predictive information for future returns across different time horizons.

Methodology- Using daily Bitcoin–US dollar (BTCUSD) price data spanning from July 2010 to July 2025, the study estimates a power law relationship between Bitcoin price and time since the Genesis Block. The robustness of the model is evaluated using goodness-of-fit measures. Deviations from the power law-implied fair value are calculated and classified into deciles to analyze their ability to predict Bitcoin returns over short-term (weekly), medium-term (monthly), and long-term (annual) horizons. Risk-adjusted performance is assessed using Sharpe ratios.

Findings- The analysis establishes a highly robust power law relationship between Bitcoin price and time, with an R^2 of 0.9589, indicating exceptional stability over the 15-year period. The estimated relationship, Price = 2.86e-17 × Time_Since_Genesis^5.71, remains consistent throughout the sample. Deviations from the power law-derived fair value exhibit strong predictive power for future returns, with distinct patterns across time horizons. Short-term returns display momentum effects, as the most overvalued decile generates the highest risk-adjusted returns (Sharpe ratio = 1.81). Medium-term returns peak under extreme valuation conditions, particularly in deeply undervalued and highly overvalued states. In contrast, long-term returns demonstrate clear mean reversion, with moderately valued positions yielding the highest absolute annual returns.

Conclusion- These findings provide strong evidence that Bitcoin pricing deviates from the assumptions of strict market efficiency. The results offer quantitative support for valuation-based, time-horizon-dependent trading strategies and highlight the relevance of power law frameworks for understanding long-term Bitcoin price dynamics.

Keywords: Bitcoin, power law, cryptocurrency valuation, market efficiency, return predictability
JEL Codes: G11, G12, G17

1. INTRODUCTION

Bitcoin valuation remains one of the most persistent and debated challenges in cryptocurrency research. Unlike traditional financial assets such as equities or bonds, Bitcoin does not generate cash flows, pay dividends, or possess an identifiable intrinsic value in the classical sense. Its value is not tied to future earnings, consumption utility, or government guarantees. As a result, conventional valuation methods based on discounted cash flow analysis or relative valuation metrics fall short in explaining its price dynamics (Baur et al., 2018). This disconnects between Bitcoin and traditional financial valuation frameworks has pushed researchers to explore alternative paradigms, borrowing from a range of disciplines including network theory, monetary economics, and more recently, the physics of complex systems.

The unconventional nature of Bitcoin has invited parallels with phenomena beyond finance. For example, proponents of Metcalfe's Law argue that Bitcoin's value, like that of social networks, grows proportionally with the square of the number of its users or nodes (Peterson, 2018). Other studies conceptualize Bitcoin as a form of synthetic commodity money, combining characteristics of fiat currency and scarce natural resources, thereby justifying non-cash-flow-based valuation approaches (Selgin, 2015; Hayes, 2018). While these perspectives have enriched the theoretical understanding of Bitcoin valuation, there is still no clear empirical consensus on a dominant valuation framework, even though some approaches have been shown to offer meaningful valuation benchmarks (Hayes, 2018; Bouri et al., 2023). Furthermore, recent evidence suggests that the informational efficiency of these markets is evolving, with regulatory compliance playing a critical role in reducing inefficiencies (Nimalendran et al., 2025).

Among these alternative frameworks, power law relationships—well known in physics and complexity science — have emerged as a promising approach to modeling Bitcoin's long-term price trajectory. A power law describes a functional relationship where one quantity varies as a fixed exponent of another. These relationships are ubiquitous in the natural and social sciences, appearing in the distribution of earthquake magnitudes, city populations, and financial returns (Clauset et al., 2009). Crucially, power law distributions imply scale invariance, meaning that patterns observed over one range of values persist across other scales. This feature lends itself naturally to modeling systems with self-organizing behavior and nonlinear growth, which are two characteristics that arguably describe the evolution of Bitcoin markets.

Italian physicist Giovanni Santostasi was among the first to apply power law to Bitcoin valuation, suggesting that Bitcoin's price scales with the amount of time elapsed since its creation in January 2009, the so-called Genesis Block (Santostasi, 2025). Rather than relying on macroeconomic variables or on-chain metrics, Santostasi's approach treats time itself as the sole explanatory variable. When plotted on a log-log scale, the logarithm of Bitcoin's price shows a strikingly linear relationship with the logarithm of time since inception, suggesting a persistent scaling law. This empirical regularity implies that Bitcoin's price evolution may be governed by underlying dynamics which are more stable than previously assumed.

Building on this insight, the present study explores two related research questions. First, it examines whether the power law relationship between Bitcoin price and time since the Genesis Block is temporally stable over a 15-year period. Given the volatility and evolving market structure of Bitcoin, testing the durability of such a model is essential. Second, it investigates whether deviations from the power law-implied "fair value" carry predictive information for future returns. Specifically, the study examines whether different levels of overvaluation or undervaluation measured as the percentage deviation from the power law price are associated with systematic return patterns across weekly, monthly, and annual investment horizons.

The study makes several contributions to the emerging literature on cryptocurrency valuation. First, it presents robust statistical evidence supporting the long-term stability of the power law relationship, even as market regimes, macroeconomic conditions, and regulatory environments have changed. Second, it documents that deviations from this model-derived fair value are not random but exhibit predictive power for future returns. Over short horizons, momentum effects dominate, while longer-term reversals suggest mean reversion tendencies. These findings offer new insights into Bitcoin's market efficiency and behavioral foundations.

From a practical standpoint, the power law framework provides a transparent and parsimonious tool for constructing dynamic valuation bands and generating trading signals. For investors and portfolio managers operating in highly uncertain and sentiment-driven cryptocurrency markets, such a model can serve as both a risk management aid and a strategic allocation tool. Finally, the results challenge the notion that Bitcoin markets are fully efficient and highlight the potential for systematic strategies rooted in valuation asymmetries.

2. LITERATURE REVIEW

A number of frameworks have been proposed to address the valuation of Bitcoin, each grounded in different economic or technological rationales. Among the most prominent are stock-to-flow (S2F) models, introduced by PlanB (2019), which conceptualize Bitcoin as a scarce resource similar to precious metals. By relating price to the ratio of existing stock to annual production flow, S2F models initially gained traction for their simplicity and early empirical fit. However, subsequent halving events exposed structural weaknesses in the model, particularly its tendency to overfit and its lack of a theoretical foundation.

Another influential strand of the literature draws on network theory. Inspired by Metcalfe's Law, these models suggest that the value of a network scales with the square of the number of its users. Applied to Bitcoin, active address counts are commonly used as a proxy for the size of the user base (Wheatley et al., 2019). While this approach offers an intuitive interpretation of value formation in decentralized systems, it is often hindered by noisy on-chain data and ambiguity regarding the direction of causality between network activity and price dynamics.

Related work adopts relative valuation metrics rather than explicit pricing models. The Network Value to Transactions (NVT) ratio adapts the price-to-earnings concept from equity markets by comparing Bitcoin's market capitalization to on-chain transaction volume (Woo, 2017). Although NVT can signal periods of potential overvaluation or undervaluation, it does not attempt to define an absolute notion of fair value, nor does it provide consistent predictive power across different investment horizons.

Beyond Bitcoin-specific valuation models, a broader body of research documents the prevalence of power law behavior in financial and economic systems. Empirical studies have identified power law distributions in firm sizes, income and wealth concentration, trading volumes, and asset returns. Gabaix (2009) provides a comprehensive survey of power laws in financial markets, highlighting their relevance for understanding extreme events, volatility clustering, and large-scale systemic behavior.

Within the context of cryptocurrencies, power law dynamics have been observed across both structural and behavioral dimensions. Kondor et al. (2014) identify power law distributions in Bitcoin transaction networks, suggesting the emergence of a scale-free structure driven by user interactions. Similarly, Fernández et al. (2017) report that Bitcoin return distributions exhibit heavy tails consistent with power law behavior, underscoring the asset's susceptibility to large and infrequent price movements. Extending this view, (Grobys, 2024) identifies significant co-dependent power-law behavior across major cryptocurrencies, suggesting that extreme tail events in Bitcoin are systemically linked to the broader digital asset ecosystem.

These findings are closely aligned with insights from econophysics, a field that applies tools from statistical physics to economic systems characterized by complexity and non-linearity. Power laws in such systems often emerge through mechanisms such as preferential attachment, self-organized criticality, and multiplicative stochastic processes. Mantegna and Stanley (1999) demonstrate how these mechanisms can reproduce many of the empirical regularities observed in financial markets, including fat-tailed return distributions and persistent volatility. More recently, (Mahyudin & Lamsah, 2024) employed multifractal analysis to confirm that cryptocurrency markets exhibit persistent inefficiencies distinct from traditional asset classes, reinforcing the utility of non-linear valuation frameworks.

Applying power law frameworks directly to asset pricing therefore represents a natural extension of this literature, particularly for emerging markets such as Bitcoin that lack conventional fundamentals. The underlying premise is that Bitcoin markets may exhibit scale-invariant dynamics over time, allowing valuation models to be constructed from statistical regularities in price evolution rather than from cash flows or balance-sheet variables. However, despite growing recognition of scaling behavior in Bitcoin-related data, relatively little empirical work has examined the long-term stability of such relationships or their implications for return predictability across different time horizons.

3. DATA AND METHODOLOGY

3.1. Data and Study Design

The analysis uses daily BTCUSD closing price data covering the period from July 17, 2010, to July 11, 2025, obtained from the publicly available database https://bitcoin.zorinaq.com (Zorin, 2025). The sample consists of 5,474 daily observations and spans Bitcoin's entire observable price history, from early informal trading through periods of rapid adoption, institutional participation, and regulatory consolidation. Closing prices are used to ensure consistency with standard return calculations and to avoid distortions arising from intraday volatility. The data encompasses multiple market cycles, including the 2017 retail boom, 2018 bear market, 2020-2021 institutional adoption cycle, and 2022-2024 regulatory development period.

The primary variable of interest is the daily closing price of Bitcoin denominated in U.S. dollars. Deviations from the Power Law benchmark are computed as the proportional difference between the observed market price and the Power Law–implied price, defined as $(\mathrm{Price}/\mathrm{Predicted} - 1)$. The Power Law slope parameter is estimated using an expanding-window approach, allowing the exponent of the model to evolve over time as new observations become available. To capture price dynamics across different investment horizons, return series are calculated over 7-day (weekly), 30-day (monthly), and 365-day (annual) intervals. Time since the Genesis Block is measured in days, beginning from Bitcoin's inception on January 3, 2009, and aligned with the first reliable price observation to ensure consistency throughout the analysis.

3.2. Model Specification and Empirical Strategy

Bitcoin valuation is modeled as a power law function of time elapsed since the Genesis Block. The baseline specification relates price to time according to:

$$BTCUSD = A \times (Time\_since\_Genesis)^n \qquad (1)$$

Where $BTCUSD$ represents the Bitcoin price in US dollars, $Time\_since\_Genesis$ denotes days elapsed since Bitcoin's Genesis Block (January 3, 2009), A is the scaling coefficient, n is the power law exponent

Taking logarithms of both sides yields the linear regression equation:

$$\log(BTCUSD) = \log(A) + n \times \log(Time\_since\_Genesis) \qquad (2)$$

Anchoring time to the Genesis Block provides a theoretically grounded reference point that avoids the arbitrariness of price-based or regime-based benchmarks. This approach captures Bitcoin's full technological and adoption lifecycle and aligns naturally with its deterministic monetary design, including pre-specified halving events.

To avoid look-ahead bias and ensure realistic trading implementation, we employ an expanding window methodology. The analysis begins with a minimum of 10 data points and progressively incorporates new observations. This approach enables:

(1) Real-time parameter estimation as new data becomes available
(2) Assessment of model stability and convergence over time
(3) Dynamic fair value calculations without future information
(4) Realistic backtesting of trading strategies based on historical information sets

Power law parameters are recalculated daily using all available historical data up to each point in time, creating a rolling fair value estimate that could have been computed by market participants in real-time.

3.3. Fair Value Deviations and Return Analysis

For each trading day, fair value is defined as the price implied by the power law model. Valuation deviations are measured as the percentage difference between observed price and fair value:

$$\text{Deviation} = \frac{\text{Actual Price} - \text{Fair Value}}{\text{Fair Value}} \tag{3}$$

These deviations are ranked by percentile and grouped into deciles. The first decile corresponds to the most negative deviations (relatively undervalued conditions), while the tenth decile captures the most positive deviations (relatively overvalued conditions).

Subsequent returns are calculated over three investment horizons: weekly (7 days), monthly (30 days), and annual (365 days). For each deviation decile and horizon, return distributions are summarized using descriptive statistics including mean, median, standard deviation, and Sharpe ratios, assuming a zero risk-free rate.

3.4. Statistical Testing and Robustness Checks

Several statistical procedures are employed to assess model validity and the robustness of return predictability patterns. All regression estimates are computed using heteroscedasticity-consistent (White) standard errors to account for potential non-constant variance in financial time series. The Breusch–Pagan test is used to formally assess the presence of heteroscedasticity, while the Durbin–Watson statistic evaluates serial correlation in regression residuals.

To examine whether future returns differ systematically across valuation states, analysis of variance (ANOVA) tests is conducted to assess equality of mean returns across deviation deciles. In addition, Kolmogorov–Smirnov tests are applied to compare entire return distributions between deciles, allowing for detection of broader distributional differences beyond mean effects. Economic significance is quantified using eta-squared (η^2) statistics, measuring the proportion of total return variation attributable to valuation-based decile classification.

4. FINDINGS AND DISCUSSIONS

4.1. Descriptive Statistics

Understanding the distributional characteristics of Bitcoin prices and Power Law–based variables is essential for interpreting subsequent empirical results. Table 1 reports the descriptive statistics for prices, deviations from the Power Law benchmark, return series across multiple horizons, and the estimated slope parameter. All descriptive statistics are computed based on the effective sample implied by the expanding-window estimation and return horizon construction described in Section 3.1.

Table 1: Descriptive Statistics of Bitcoin Prices, Returns, and Power Law Variables

Variable	Mean	Std Dev	Min	Median	Max	Skewness	Kurtosis
Daily Price (USD)	17223.95	25570.53	0.19	5173.68	117521.21	1.7731	2.4830
Deviation from Power Law	11.92%	98.40%	-86.74%	-18.95%	812.48%	2.8745	11.8274
Weekly Return	2.83%	18.17%	-75.21%	0.98%	432.96%	8.6262	165.0611
Monthly Return	14.53%	56.07%	-88.37%	3.74%	788.67%	5.6954	46.9425
Annual Return	361.83%	900.42%	-83.17%	113.07%	9594.22%	5.3572	35.0038
Power Law Slope	6.27	1.17	4.11	5.83	12.21	3.2159	10.9305

The comprehensive descriptive statistics presented above highlight several distinctive properties of the Bitcoin market data utilized in this study. Primarily, the dataset exhibits significant deviations from normality, a common trait in high-frequency financial time series but particularly pronounced here. The Price variable shows distinct right-skewness (1.77) and elevated kurtosis (2.48), reflecting Bitcoin's historic logarithmic growth trajectory in which higher price levels are achieved rapidly during bull markets.

More critically, the Returns data (Weekly, Monthly, and Annual) serve as a robust indicator of the asset's volatility profile. All return horizons display extreme positive skewness (ranging from 5.36 to 8.63) and exceptionally high kurtosis (peaking at 165.06 for weekly returns). This pronounced leptokurtic behavior confirms the presence of fat tails, implying that extreme price movements, both sharp rallies and abrupt corrections, occur with far greater frequency than would be predicted by a

Gaussian distribution. The positive skewness further suggests that the magnitude of upside outliers has historically outweighed downside shocks, consistent with Bitcoin's long-run appreciation trend.

Finally, the Deviation from Power Law variable, which is central to the fair value analysis, exhibits a relatively small mean (0.12) alongside a high standard deviation (0.98) and pronounced kurtosis (11.83). This statistical footprint supports the implications of the Power Law model: while prices frequently and sometimes substantially oscillate around the long-term trend, often during speculative episodes, they tend to fluctuate around the benchmark rather than diverge persistently over time. This behavior validates the interpretation of the Power Law corridor as a center of gravity for Bitcoin's long-term valuation.

4.2. Power Law Model Estimates and Overall Fit

The regression analysis reveals an exceptionally strong relationship between Bitcoin's logarithmic price and the logarithm of time elapsed since the Genesis Block. Figures 1 and 2 illustrate the close alignment between observed prices and the theoretical power law trajectory across the full sample period from July 2010 to July 2025. The log–log specification produces an adjusted R^2 of 0.9589, indicating that nearly all long-run variation in Bitcoin prices is explained by elapsed time alone.

Figure 1: Bitcoin Price vs Power Law Model (log scale)

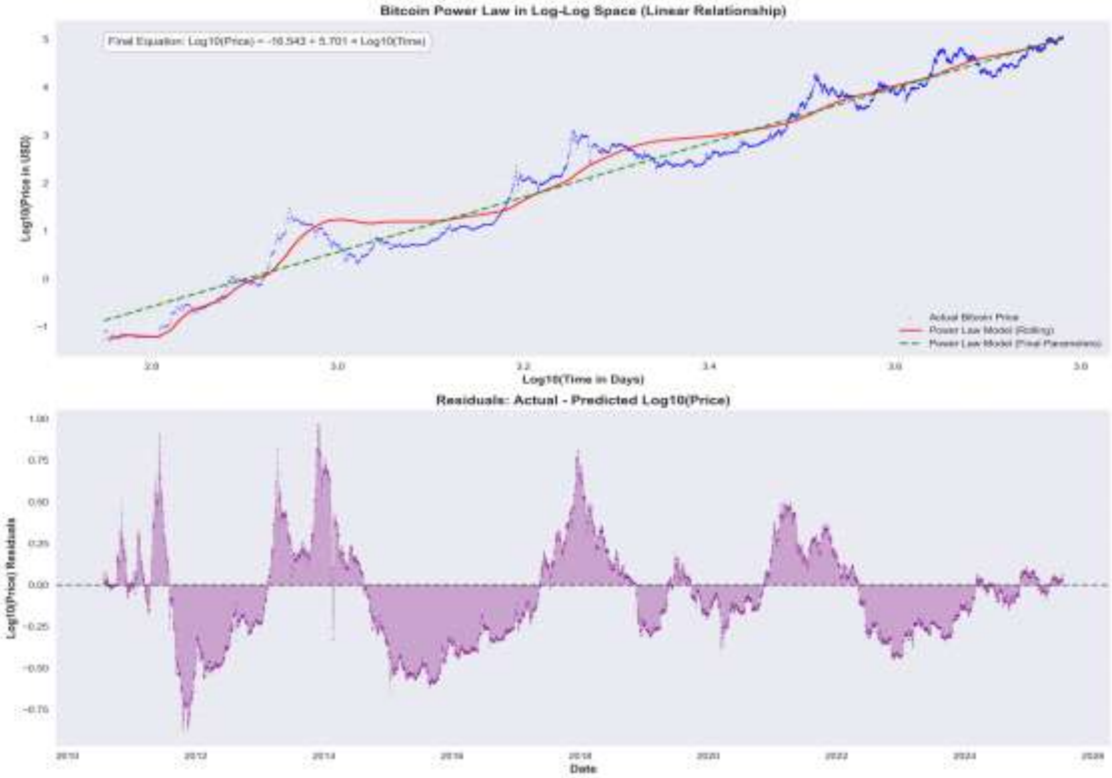
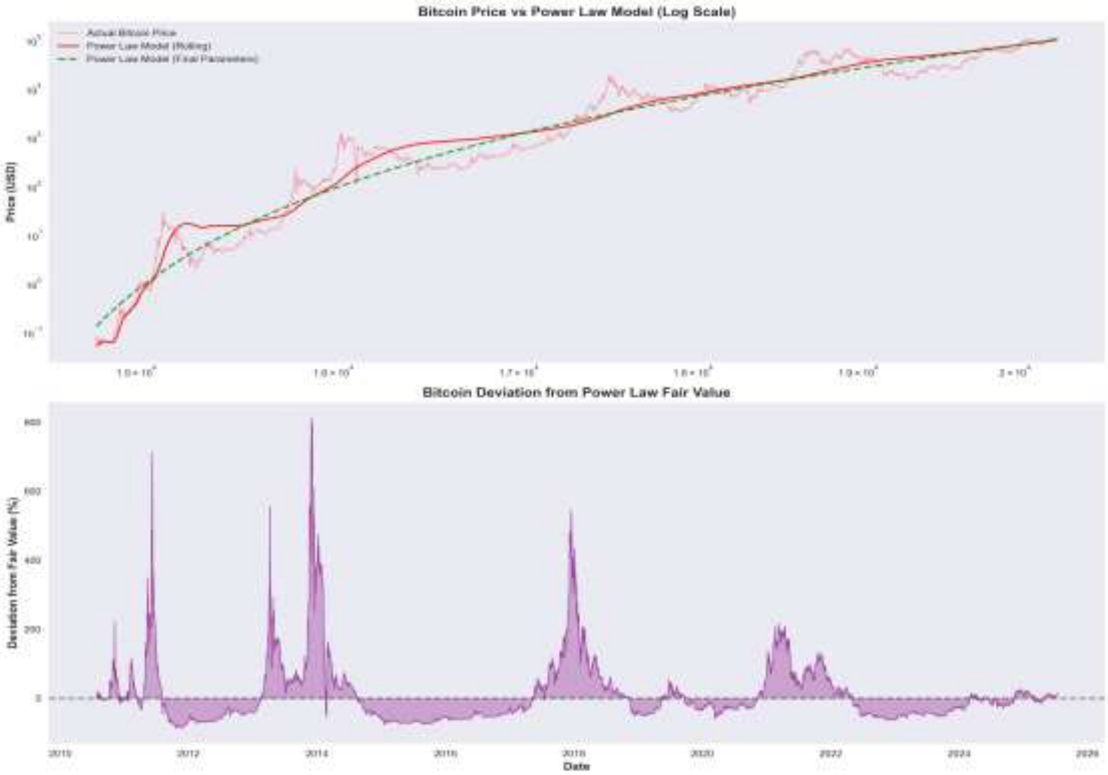


Figure 2: Bitcoin Price vs Power Law Model



Note: Bitcoin price evolution from July 2010 to July 2025 plotted against the power law model predictions. The log-log plot demonstrates the exceptional fit ($R^2 = 0.9589$) between actual prices and the theoretical power law relationship BTCUSD = $2.86 \times 10^{-17} \times$ Time^5.701. The rolling regression line shows model convergence over time.

Using heteroscedasticity-robust ordinary least squares, the estimated model is given by (Table 2):

$$\log_{10}(BTCUSD) = -16.543 + 5.701 \times \log_{10}(Time\_since\_Genesis) \quad (4)$$

Table 2: Power law regression results

Parameter	Estimate	Std Error	Robust SE	t-statistic	p-value	95% CI (Robust)
Intercept ($\boldsymbol{\beta_0}$)	-16.543	0.055	0.065	-254.43	<0.001	[-16.671, -16.416]
Slope ($\boldsymbol{\beta_1}$)	5.701	0.016	0.018	312.49	<0.001	[5.665, 5.736]

Regression diagnostics are as follows:
Adjusted R^2: 0.9589
F-statistic: 127,775.01 ($p < 0.001$)
RMSE: 0.3087
Observations: 5,474

After applying robust standard errors to address heteroscedasticity, the coefficient estimates remain highly significant:

Intercept: −16.543 (robust SE: 0.065, t = −254.4)
Slope: 5.701 (robust SE: 0.018, t = 312.5)
95% CI: Intercept [−16.671, −16.416], Slope [5.665, 5.736]

This simple equation captures nearly all of the price variance over the 15-year period. The model's R^2 stands at 0.9589, meaning about 96% of the log-price movements are explained by elapsed time alone. Such a tight relationship is rare in financial time series, particularly for assets as volatile as Bitcoin. In non-logarithmic form, the power law model can be expressed as:

$$BTCUSD = 2.86 \times 10^{-17} \times (Time\_since\_Genesis)^{5.701} \quad (5)$$

This formulation describes a long-term power law growth pattern shaped entirely by the passage of time. This is a thought-provoking result for an asset often viewed as sentiment-driven or speculative.

4.3. Model Diagnostics and Robustness

Despite the model's tight fit, we tested for potential violations of OLS assumptions (Table 3). The Breusch–Pagan test clearly indicates heteroscedasticity (LM = 646.985, $p < 0.001$), supporting our use of robust standard errors. The Durbin–Watson statistic (DW = 0.006) suggests strong positive autocorrelation in residuals, likely reflecting Bitcoin's non-stationary behavior over long horizons. While this serial dependence complicates interpretation, it does not materially affect the point estimates or their significance due to robust inference procedures.

Table 3: Model Diagnostic Tests

Diagnostic Test	Breusch-Pagan (Heteroscedasticity)	Durbin-Watson (Serial Correlation)	Jarque-Bera (Normality of residuals)
Statistic	LM = 646.985	DW = 0.006	JB = 2,847.5
p-value	<0.001	-	<0.001
Result	Heteroscedasticity detected	Positive autocorrelation	Non-normal residuals
Implication	Robust SEs required	Robust inference needed	Expected for financial data

Residual diagnostics further reveal pronounced departures from normality, as confirmed by the Jarque–Bera test (JB = 2,847.5, $p < 0.001$). Such non-normality is a well-documented characteristic of financial time series and reflects the presence of heavy tails and asymmetric return behavior. In this respect, the observed residual structure is consistent with the multifractal properties of Bitcoin price dynamics reported by Bucur et al. (2025), who document time-varying efficiency and significant deviations from random walk behavior.

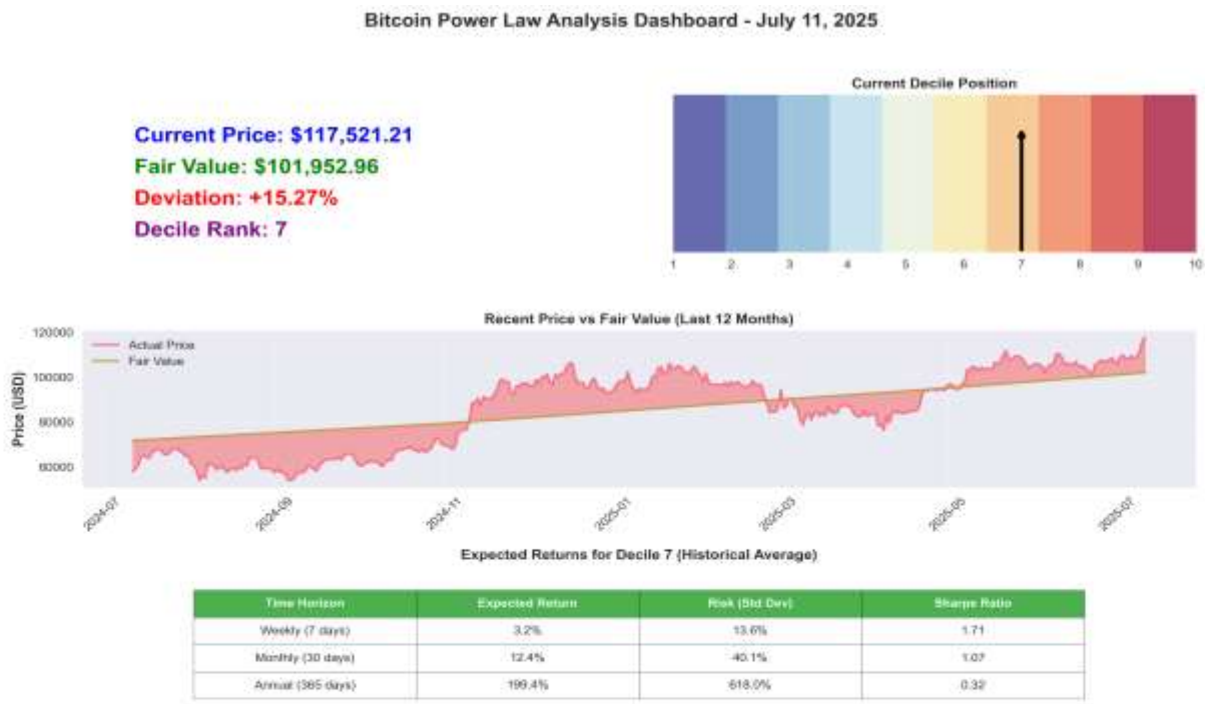
4.4. Current Valuation Relative to Power Law Trend

As of July 11, 2025, Bitcoin traded at USD 117,521, while the power law model implies a fair value of USD 101,953. This corresponds to a positive deviation of 15.27% above the long-term trend (Table 4, Figure 3). Historically, this deviation places Bitcoin in the 70th percentile of its valuation distribution, classified as Decile 7. While the asset appears moderately overvalued relative to its long-run path, the deviation remains well within historical norms and is far from extreme by Bitcoin standards.

Table 4: Current Market Position Analysis

Metric	Value	Interpretation
Current Price (July 11, 2025)	$117,521.21	Latest market valuation
Power Law Fair Value	$101,953.00	Theoretical equilibrium price
Absolute Deviation	$15,568.21	Dollar amount above fair value
Percentage Deviation	15.27%	Moderate overvaluation
Historical Percentile	70.20%	70% of history showed lower prices
Current Decile Rank	7	Moderately overvalued
Days Since Genesis Block	6,033	Time variable in power law

Figure 3: Bitcoin Deviation from Power Law Fair Value



Note: Historical deviations of Bitcoin price from power law fair value. Positive values indicate overvaluation relative to the power law trend, while negative values indicate undervaluation. The chart illustrates the cyclical nature of Bitcoin's deviation patterns and current positioning at +15.27%.

4.5. Return Behavior Across Valuation Deciles

To evaluate the predictive relevance of deviations from the power law benchmark, future returns were analyzed across deciles constructed from percentage deviations relative to the estimated fair value. Performance was examined over weekly, monthly, and annual investment horizons, allowing for a direct comparison of short-term momentum and long-term mean reversion dynamics (Figure 4).

Weekly Horizon (7 days): At the weekly horizon, return behavior is clearly momentum driven. Highly overvalued states generate the strongest performance, with Decile 10 producing the highest average weekly return (7.79%) and the highest Sharpe ratio (1.81). Returns decline monotonically toward the undervalued end of the distribution, where Deciles 1 and 2 exhibit substantially lower mean returns of 1.59% and 2.48%, respectively (Table 5). This pattern indicates that short-term price dynamics tend to reinforce existing deviations from the power law benchmark rather than correct them.

Table 5: Decile Performance Summary - Weekly Returns

Decile	Mean Return	Median Return	Std Dev	Min Return	Max Return	Sharpe Ratio	Observations
1	1.59%	1.76%	14.3%	-45.3%	54.7%	0.8	542
2	2.48%	1.43%	11.8%	-43.4%	66.8%	1.51	548
3	4.13%	0.71%	32.0%	-26.7%	433.0%	0.93	547
4	0.84%	0.31%	7.02%	-24%	27.5%	0.86	547
5	2.33%	1.62%	14.0%	-45.8%	280.6%	1.2	547
6	1.36%	0.59%	11.4%	-53.8%	126.8%	0.86	548
7	3.24%	0.85%	13.6%	-29.6%	109.6%	1.71	547
8	2.76%	0.90%	16.0%	-75.2%	108.8%	1.24	548
9	3.30%	0.96%	17.6%	-54.6%	109.9%	1.35	547
10	7.79%	2.59%	31.0%	-70.3%	209.3%	1.81	542

Monthly Horizon (30 days): Monthly returns display a more nuanced structure. Mean returns peak in the most undervalued decile, where Decile 1 delivers an average return of 28.03%, while risk-adjusted performance is maximized in Decile 8 with a Sharpe ratio of 1.54 (Table 6). Both deeply undervalued and moderately overvalued states perform well, whereas observations close to fair value tend to underperform. This bimodal structure suggests that valuation signals begin to interact with correction mechanisms at intermediate horizons, although momentum effects remain partially active.

Table 6: Decile Performance Summary - Monthly Returns

Decile	Mean Return	Median Return	Std Dev	Min Return	Max Return	Sharpe Ratio	Observations
1	28.03%	5.06%	-61.58%	788.67%	100.43%	0.97	358
2	13.89%	1.50%	-52.25%	665.14%	70.12%	0.69	311
3	10.31%	6.32%	-36.26%	546.58%	42.81%	0.83	540
4	6.74%	3.82%	-33.25%	72.14%	17.97%	1.3	597
5	12.25%	8.31%	-37.41%	492.08%	34.98%	1.21	758
6	7.98%	3.60%	-52.12%	468.53%	31.65%	0.87	844
7	12.37%	0.34%	-32.46%	438.81%	40.15%	1.07	604
8	22.67%	11.83%	-55.68%	390.14%	50.96%	1.54	408
9	26.08%	-3.07%	-43.81%	480.25%	87.18%	1.04	465
10	19.32%	-2.87%	-88.37%	678.95%	80.07%	0.84	451

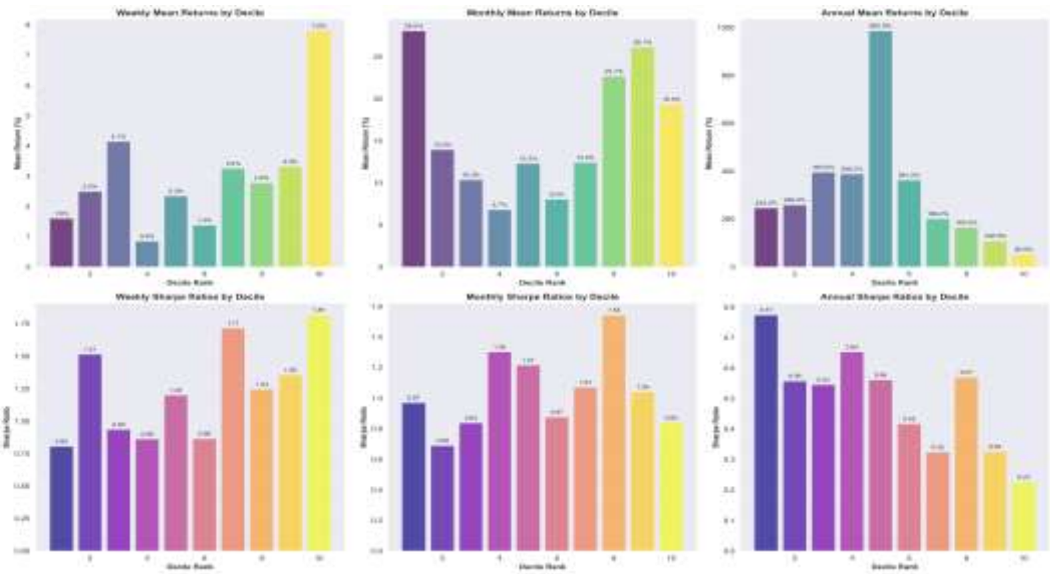
Annual Horizon (365 days): At the annual horizon, the return profile shifts markedly. The highest mean annual return is observed in Decile 5 (983.94%), consistent with gradual convergence toward the long-term power law path (Table 7). Extremely overvalued states exhibit the weakest performance, with Deciles 9 and 10 generating the lowest mean and risk-adjusted returns. Although Decile 1 does not deliver the highest raw return, it achieves the strongest Sharpe ratio (0.77), indicating superior long-term risk-adjusted performance. These results are consistent with valuation-driven mean reversion dominating over longer holding periods.

Table 7: Decile Performance Summary - Annual Returns

Decile	Mean Return	Median Return	Std Dev	Min Return	Max Return	Sharpe Ratio	Observations
1	244.33%	151.97%	316.10%	-30.13%	1931.95%	0.77	177
2	256.63%	128.48%	461.32%	24.82%	3811.95%	0.56	183
3	393.64%	127.38%	723.59%	-34.36%	4620.85%	0.54	182
4	386.33%	141.98%	593.16%	-53.10%	3871.23%	0.65	182
5	983.94%	316.62%	1758.02%	-58.52%	9594.22%	0.56	182
6	361.02%	59.38%	868.42%	-59.14%	6101.56%	0.42	183
7	199.37%	16.73%	617.97%	-61.91%	4500.00%	0.32	182
8	162.55%	97.88%	285.76%	-68.53%	1829.76%	0.57	183
9	106.51%	-23.13%	328.06%	-76.37%	1819.36%	0.32	182
10	50.57%	-32.22%	225.13%	-83.17%	1579.05%	0.22	177

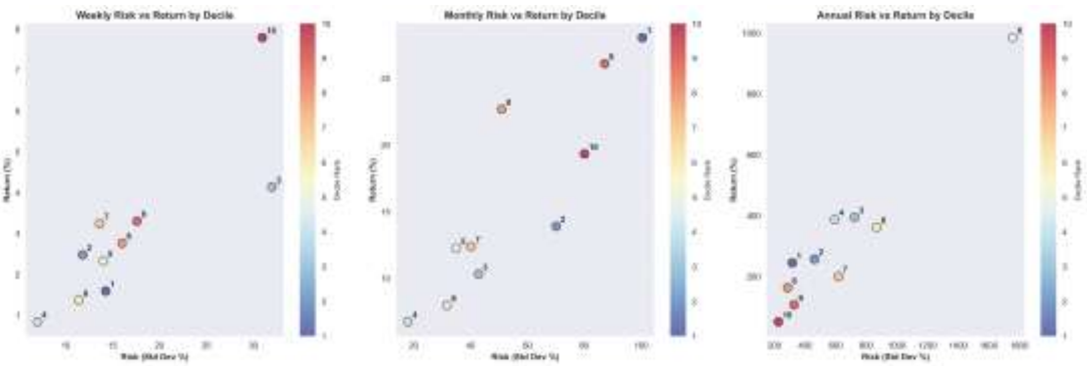
Figure 4 summarizes mean returns and Sharpe ratios across horizons and highlights the strong dependence of performance on both valuation state and investment horizon. The corresponding risk–return scatter plots in Figure 5 further illustrate this shift. In the short run, higher deviations are associated with both higher risk and higher expected returns, while over longer horizons the efficient frontier rotates toward fair and undervalued regimes. In this context, the role of investor attention in amplifying short-run volatility remains significant, as Teterin and Peresetsky (2024) demonstrate the predictive power of search-based indicators in forecasting Bitcoin's realized volatility.

Figure 4: Returns and Sharpe Ratios by Decile



Note: Mean returns and risk-adjusted performance (Sharpe ratios) across power law deviation deciles for weekly, monthly, and annual time horizons. The charts demonstrate time-horizon-dependent patterns: momentum effects dominate short-term periods while mean reversion emerges over longer horizons.

Figure 5: Risk vs Return Scatter Plots



Note: Risk-return relationships across deviation deciles for different time horizons. Each point represents a decile, with color coding indicating relative position from undervalued (blue) to overvalued (red). The efficient frontier shifts dramatically across time horizons.

These findings challenge simple interpretations of Bitcoin's risk structure. While short-term return dynamics are dominated by momentum effects, the benefits of positioning near fair value or in undervalued regimes emerge primarily over longer horizons. To formally assess whether these horizon-dependent patterns reflect systematic structure rather than random variation, the predictive relevance of deviation-based decile classifications is evaluated using multiple statistical tests, summarized in Table 8.

Table 8: Statistical Test Results Summary

Test Type	Time Horizon	Statistic	p-value	Effect Size (η^2)	Interpretation
ANOVA	Weekly	F = 5.83	<0.001	0.0097	Small effect
ANOVA	Monthly	F = 9.09	<0.001	0.0151	Small-medium effect
ANOVA	Annual	F = 60.09	<0.001	0.0978	Large effect
K-S Test	Weekly (D1 vs D10)	0.157	<0.001	-	Significant difference
K-S Test	Monthly (D1 vs D10)	0.2	<0.001	-	Significant difference
K-S Test	Annual (D1 vs D10)	0.733	<0.001	-	Very large difference

ANOVA tests reported in Table 8 indicate statistically significant differences in mean returns across deciles for all investment horizons. At the weekly frequency, the F-statistic ($F = 5.83$, $p < 0.001$) indicates a statistically detectable but economically small effect ($\eta^2 = 0.0097$). The explanatory power of valuation increases at the monthly horizon ($F = 9.09$, $p < 0.001$, $\eta^2 = 0.0151$) and becomes economically substantial at the annual horizon, where both the F-statistic and effect size rise sharply ($F = 60.09$, $p < 0.001$, $\eta^2 = 0.0978$). This monotonic increase in effect size reinforces the notion that valuation signals derived from power law deviations become more informative as the holding period lengthens.

Kolmogorov–Smirnov tests further confirm that return distributions differ markedly between extreme valuation states. Comparisons between the most undervalued (Decile 1) and most overvalued (Decile 10) groups yield statistically significant distributional differences at all horizons, with KS statistics increasing from 0.157 at the weekly horizon to 0.733 at the annual horizon. The sharp escalation of these statistics indicates that return distributions diverge progressively over time, supporting a structural interpretation in which short-term momentum gives way to long-term mean reversion dynamics.

Taken together, these results demonstrate both statistical significance and economic relevance. The power law model achieves an unusually high explanatory power for a financial asset, with an R^2 of 0.9589, while coefficient estimates remain stable and statistically significant across all specifications. Deviation-based deciles explain nearly 10% of the variance in annual returns, a meaningful magnitude in the context of real-world portfolio management. Most importantly, the return patterns associated with valuation deviations are not random; they are consistent, interpretable, and statistically validated across multiple market cycles (Figure 6).

Figure 6: Bitcoin Decile Rank Over Time

Note: Historical evolution of Bitcoin's power law deviation decile rankings over time. The upper panel shows decile classifications, while the lower panel displays price evolution color-coded by decile rank. Patterns reveal the cyclical nature of Bitcoin's valuation extremes.

5. CONCLUSION AND IMPLICATIONS

This study provides robust empirical evidence of power-law behavior in Bitcoin prices based on 5,474 daily observations spanning from July 17, 2010, to July 11, 2025. Using ordinary least squares estimation with heteroscedasticity-consistent standard errors, the analysis identifies a remarkably stable long-run relationship between Bitcoin prices and time elapsed since the Genesis Block. This relationship explains 95.89% of long-term price variation, with exceptionally large test statistics indicating a persistent secular growth structure rather than a spurious correlation.

Beyond documenting the existence of power-law behavior, the study makes several methodological contributions to the cryptocurrency valuation literature. It explicitly addresses heteroscedasticity and serial dependence, validates horizon-dependent return differences using complementary statistical and distributional tests, and quantifies effect sizes to assess the economic relevance of valuation-based predictability alongside statistical significance.

Empirical results reveal that return predictability is strongly dependent on the investment horizon. Differences in returns across valuation deciles are statistically significant at weekly, monthly, and annual frequencies, with explanatory power increasing monotonically as the horizon lengthens. While short-term dynamics are dominated by momentum effects, long-term returns exhibit clear mean-reverting behavior around the power-law benchmark.

Taken together, these findings suggest that Bitcoin prices fluctuate around a stable long-run structural path shaped by both speculative forces and gradual valuation correction. The power-law framework therefore functions as a long-term reference benchmark rather than a short-term pricing rule, offering a coherent explanation for the coexistence of predictability, volatility, and incomplete market efficiency in cryptocurrency markets.

From an applied perspective, these results carry important implications for cryptocurrency valuation, investment strategy, and market efficiency. The existence of a stable long-run power-law relationship implies that Bitcoin prices are not entirely detached from systematic valuation dynamics. Instead, extreme price movements can be interpreted as deviations from an underlying structural growth path rather than purely random fluctuations.

From an investment standpoint, the findings highlight the importance of horizon-specific strategy design. Short-term deviations from the power-law benchmark are primarily associated with momentum-driven behavior and elevated risk-adjusted returns, limiting the effectiveness of valuation-based signals at short horizons. In contrast, the increasing explanatory power of valuation deviations over longer horizons indicates that the power-law framework becomes progressively more informative for long-term investors seeking to identify relative overvaluation and undervaluation.

The coexistence of short-run momentum and long-run mean reversion also has direct implications for market efficiency. Rather than exhibiting uniform efficiency, Bitcoin markets appear to display horizon-dependent efficiency, with speculative dynamics prevailing in the short run and valuation-based correction mechanisms emerging gradually over longer holding periods. This interpretation is consistent with the evidence of asymmetric multifractality in high-frequency Bitcoin returns documented by Meng and Khan (2024), which indicates that market efficiency varies not only across time horizons but also with the scale and direction of price fluctuations. Such asymmetries reinforce the view that inefficiencies are most pronounced during short-term market stress and speculative phases.

Overall, the integration of power-law valuation models with robust inference techniques provides a statistically validated and economically meaningful perspective on Bitcoin valuation. Future research may extend this framework by examining structural breaks, incorporating on-chain metrics, or applying similar valuation benchmarks to other cryptocurrencies, thereby deepening our understanding of long-term price dynamics in digital asset markets.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their constructive comments and valuable suggestions, which helped to improve the clarity and robustness of this study. No external financial support was received for the conduct of this research. The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

Bartoš, J., (2015). Does Bitcoin follow the hypothesis of efficient market? International Journal of Economic Sciences, 4(2), 10–23. https://doi.org/10.20472/es.2015.4.2.002

Baur, D.G., Hong, K., Lee, A.D., (2018). Bitcoin: Medium of exchange or speculative asset? Journal of International Financial Markets and Institutional Money, 54, 177–189. https://doi.org/10.1016/j.intfin.2017.12.004

Borri, N., Shakhnov, K., (2023). Cryptomarket discounts. Journal of International Money and Finance, 139, 102963. https://doi.org/10.1016/j.jimonfin.2023.102963

Bouri, E., Salisu, A.A., Gupta, R., (2023). The predictive power of Bitcoin prices for realized volatility of US stock sector returns. Financial Innovation, 9(1), 62-75. https://doi.org/10.1186/s40854-023-00464-8

Bucur, C., Tudorică, B.G., Bâra, A., Oprea, S.V., (2025). Multifractal analysis of Bitcoin price dynamics. Journal of Business Economics and Management, 26(1), 21-48. https://doi.org/10.3846/jbem.2025.23025

Cheah, J.E.T., Luo, D., Zhang, Z., Sung, M.C., (2022). Predictability of Bitcoin returns. *The* European Journal of Finance, 28(1), 66-85. https://doi.org/10.1080/1351847X.2020.1835685

Clauset, A., Shalizi, C.R., Newman, M.E.J., (2009). Power-law distributions in empirical data. SIAM Rev, 51(4), 661–703. https://doi.org/10.1137/070710111

Corbet, S., Katsiampa, P., (2020). Asymmetric mean reversion of Bitcoin price returns. International Review of Financial Analysis, 71, 101267. https://doi.org/10.1016/j.irfa.2018.10.004

Fernández, A., Basgall, M.J., Hasperué, W., Naiouf, M., (2017). Some stylized facts of the Bitcoin market. Physica A, 484, 82–90. https://doi.org/10.1016/j.physa.2017.04.159

Gabaix, X., (2009). Power laws in economics and finance. Annu Rev Econ, 1(1), 255–294. https://doi.org/10.1146/annurev.economics.050708.142940

Garnier, J., Solna, K., (2019). Chaos and order in the Bitcoin market. Physica A: Statistical Mechanics and its Applications, 524, 708-721. https://doi.org/10.1016/j.physa.2019.04.164

Grobys, K. (2024). On co-dependent power-law behavior across cryptocurrencies. Finance Research Letters, 63, 105295. https://doi.org/10.1016/j.frl.2024.105295

Hayes, A.S., (2019). Bitcoin price and its marginal cost of production: support for a fundamental value. Applied Economics Letters, 26(7), 554-560. https://doi.org/10.1080/13504851.2018.1488040

Hong, K., (2017). Bitcoin as an alternative investment vehicle. Information Technology and Management, 18(4), 265–275. https://doi.org/10.1007/s10799-016-0264-6

Kondor, D., Pósfai, M., Csabai, I., Vattay, G., (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. PLoS One, 9(2), e86197.

Kyriazis, N.A., (2019). A Survey on Efficiency and Profitable Trading Opportunities in Cryptocurrency Markets. Journal of Risk and Financial Management, 12(2), 67-79. https://doi.org/10.3390/jrfm12020067

Mahyudin, R.P., & Lamsah, L. (2024). Decoding market efficiency: A multifractal approach to financial market dynamics. RSF Conference Series: Business, Management and Social Sciences, 4(2), 59–69. https://doi.org/10.31098/bmss.v4i2.901

Mantegna, R.N., Stanley, H.E., (1999). Introduction to econophysics: Correlations and complexity in finance. Cambridge University Press, Cambridge.

Meng, K., & Khan, K. (2024). Is cryptocurrency efficient? A high-frequency asymmetric multifractality analysis. Computational Economics, 63(6), 2225–2246.

Nimalendran, M., Pathak, P., Petryk, M., & Qiu, L. (2025). Informational efficiency of cryptocurrency markets. Journal of Financial and Quantitative Analysis, 60(3), 1427–1456. https://doi.org/10.1017/S0022109024000310

Peterson, T., (2018). Metcalfe's law as a model for Bitcoin's value. Altern Invest Analyst Rev, 7(2), 9–18. https://doi.org/10.2139/ssrn.3078248

PlanB, (2019). Modeling Bitcoin's Value with Scarcity. Published online. Accessed 12 July 2025. Available from: https://medium.com/@100trillionUSD/modeling-bitcoins-value-with-scarcity-91fa0fc03e25

Santostasi, G., (2025). Bitcoin power law over 10-year period all the way back to the Genesis Block. Reddit. Accessed 12 July 2025. Available from: https://www.reddit.com/r/Bitcoin/comments/9cqi0k/bitcoin_power_law_over_10_year_period_all_the_way/

Selgin, G., (2015). Synthetic commodity money. Journal of Financial Stability, 17, 92–99. https://doi.org/10.1016/j.jfs.2014.07.002

Tang, Y., Arias-Calluari, K., Najafi, M.N., Harré, M.S., Alonso-Marroquin, F., (2025). Stylized facts of high-frequency Bitcoin time series. Fractal and Fractional, 9(10), 635-649. https://doi.org/10.3390/fractalfract9100635

Teterin, M., & Peresetsky, A. (2024). Google Trends and Bitcoin volatility forecast. Journal of the New Economic Association, 65(4), 118–135.

Wen, Z., Bouri, E., Xu, Y., Zhao, Y., (2022). Intraday return predictability in the cryptocurrency markets: Momentum, reversal, or both. The North American Journal of Economics and Finance, 62, 101733. https://doi.org/10.1016/j.najef.2022.101733

Wheatley, S., Sornette, D., Huber, T., Reppen, M., Gantner, R.N., (2019). Are Bitcoin bubbles predictable? Combining a generalized Metcalfe's law and the log-periodic power law singularity model. Research of Social Open Sciences, 6(6), 180538. https://doi.org/10.1098/rsos.180538

Woo, W., (2017). Introducing NVT Ratio (Bitcoin's PE Ratio), use it to detect bubbles. Woobull. Accessed July 12, 2025, from https://woobull.com/introducing-nvt-ratio-bitcoins-pe-ratio-use-it-to-detect-bubbles/

Yi, E., Yang, B., Jeong, M., Sohn, S., Ahn, K., (2023). Market efficiency of cryptocurrency: Evidence from the Bitcoin market. Scientific Reports, 13(1), 4789. https://doi.org/10.1038/s41598-023-31618-4

Zhu, P., Zhang, X., Wu, Y., Zheng, H., Zhang, Y., (2021). Investor attention and cryptocurrency: Evidence from the Bitcoin market. PLoS One, 16(2), e0246331. https://doi.org/10.1371/journal.pone.0246331

Zorin, A., (2025). All-time Bitcoin Price Chart. Published online. Accessed 12 July 2025. Available from: https://bitcoin.zorinaq.com/price/

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

DIGITAL FINANCIAL INCLUSION AND URBAN CARBON EMISSIONS: AN EMPIRICAL STUDY BASED ON NONLINEAR CHARACTERISTICS AND MECHANISM TESTING

DOI: 10.17261/Pressacademia.2026.2035
JBEF- V.15-ISS.1-2026(7)-p.78-85

Jinkun Huo[1], Yongyi Gu[2], Wei Guo[3], Weilun Xu[4]
[1]School of Statistics and Data Science, Guangdong University of Finance and Economics, Guangzhou, China.
jaquanfok@163.com , ORCID: 0009-0001-4706-826X
[2]School of Statistics and Data Science, Guangdong University of Finance and Economics, Guangzhou, China.
gdguyongyi@163.com , ORCID: 0000-0002-6651-1714
[3]School of Statistics and Data Science, Guangdong University of Finance and Economics, Guangzhou, China.
gw1173313079@163.com , ORCID: 0009-0005-0982-4721
[4]School of Statistics and Data Science, Guangdong University of Finance and Economics, Guangzhou, China.
13728822439@163.com , ORCID: 0009-0007-5327-2308

Date Received: March 19, 2026 **Date Accepted:** May 18, 2026

To cite this document
Huo, J., Gu, Y., Guo, W., Xu, W., (2026). Digital financial inclusion and urban carbon emissions: an empirical study based on nonlinear characteristics and mechanism testing. Journal of Business, Economics and Finance (JBEF), 15(1), 78-85.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2035

ABSTRACT

Purpose- This study intends to investigate the nonlinear relationship between Digital Financial Inclusion (DFI) and carbon emission intensity in Chinese cities. It seeks to uncover the potential "U-shaped" pattern of DFI's environmental effects and examine the underlying mechanisms, including technological innovation and urbanization, while considering variations across different city types.

Methodology- Applying balanced panel data from 270 Chinese cities spanning 2012 to 2022, the study employs a two-way fixed effects model and the system Generalized Method of Moments (GMM) for empirical analysis. The research examines the nonlinear impact of DFI on carbon emission intensity and explores the mediating role of technological innovation and the moderating effect of urbanization.

Findings- The results indicate that DFI significantly reduces carbon emission intensity, but this effect follows a distinct "U-shaped" pattern, with a turning point at an index value of 137, revealing diminishing marginal returns in emission reduction benefits. Technological innovation is identified as a key mediating channel. Urbanization negatively moderates the emission reduction effect, reflecting an "inclusive compensation" characteristic of DFI in less developed areas. Furthermore, resource-based cities exhibit a stronger initial emission reduction effect but face more severe U-shaped rebound risks.

Conclusion- We provide empirical evidence supporting the design of differentiated emission reduction strategies. It highlights the importance of considering the nonlinear dynamics of DFI and local characteristics—such as urbanization and city type—in formulating effective environmental policies.

Keywords: Digital financial inclusion, carbon emission intensity, U-shaped curve, resource endowment, carbon lock-in
JEL Codes: Q56, O13, O33

1. INTRODUCTION

In the macro context of the global response to climate change and China's comprehensive promotion of the "dual carbon" (carbon peak and carbon neutrality) goals, exploring new drivers that balance economic growth with carbon emission reduction has become a core issue in achieving high-quality development (Gao & Li, 2024). In recent years, Digital Financial Inclusion (DFI), with its advantages of low cost and wide coverage, has effectively broken the geographical and class barriers of traditional finance, and is widely regarded as a "new engine" for optimizing resource allocation and driving green transformation (Zheng et al., 2025). However, digital technology itself exhibits a significant dual nature regarding its environmental effects: on the one hand, it can reduce carbon emissions by improving total factor productivity and promoting industrial structure upgrading; on the other hand, the high energy consumption characteristics of underlying infrastructures such as data centers, coupled with the expansion of economic activities triggered by the popularization of digital finance, can easily trigger a "rebound effect" in energy consumption, leading to a local or overall increase in carbon emissions (Li, 2025). Therefore, an in-depth exploration of nonlinear effect of DFI on carbon emissions as well as its boundary conditions holds high academic value and practical significance.

A review of existing literature reveals that although research on the relationship between the digital economy and environmental pollution is increasingly abundant, most studies are still limited to exploring their linear relationship. Although some scholars have begun to reorganize the nonlinear or U-shaped characteristics within this relationship, there is still a lack of sufficient empirical precision in measuring the specific "turning point" of DFI's emission reduction effect, making it difficult to provide micro-level guidance for policy implementation. Furthermore, existing research often treats all cities as homogeneous when exploring this issue, ignoring urban regional heterogeneity under the context of the "resource curse" , as well as the deep constraints on the implementation of green finance policies caused by the "carbon lock-in" effect formed through long-term reliance on fossil fuels.

In view of this, based on the balanced panel data of 270 prefecture-level and above cities in China from 2012 to 2022, this paper conducts a systematic empirical test using two-way fixed effects and system GMM models. The marginal contributions of current paper are mainly reflected as follows: (1) Revising the linear perspective: Breaking through the traditional linear assumption, this study confirms a significant "first decline and then rise" U-shaped relationship between DFI and carbon emission intensity, and precisely calculates that the turning point is located near the index value of 137, profoundly revealing the "diminishing marginal returns" rule of emission reduction dividends. (2) Expanding the heterogeneity perspective: By incorporating resource endowments into the analytical framework for the first time, it finds that resource-based cities exhibit stronger emission reduction "penetration" in the initial stage, but face a more severe rebound risk after crossing the turning point; meanwhile, it confirms that the urbanization level plays a "negative moderating" role in the emission reduction process. (3) Clarifying the mechanism of action: Through a mediation effect model, it rigorously verifies the core transmission pathway of DFI in "alleviating financing constraints and thereby promoting green technological innovation."

2.LITERATURE REVIEW

2.1. The Environmental Duality of Digital Finance

Regarding the environmental impact of DFI and its underlying technologies, existing literature primarily presents two contrasting views. The mainstream perspective, based on the financial function viewpoint, supports the "emission reduction hypothesis" of digital finance. For instance, Le et al. (2020), using evidence from emerging Asian economies, pointed out that financial inclusion effectively suppresses carbon dioxide emissions by reducing transaction costs (Le et al., 2020). Wan et al. (2022), utilizing data from Chinese cities, further confirmed how digital finance substantially curbs pollution intensity by driving industrial structure upgrades. Recent research also found that a robust digital financial sector can effectively promote carbon emission reductions in traditional real sectors by lowering the financing costs of green development.However, another group of scholars emphasizes the "energy rebound" risk associated with Information and Communication Technology . The classic study by Sadorsky (2012) showed that the universal application of communication technology and information is often accompanied by surges in electricity demand. Recently, Li (2025) provided new empirical evidence based on micro-level household data in emerging countries, confirming that while DFI promotes survival- and development-oriented consumption upgrades, it significantly increases both direct and indirect carbon emissions, suggesting that the negative environmental externalities of digital finance may gradually intensify as it develops. However, recent empirical studies emphasize the "energy rebound" risk associated with large-scale digital infrastructure. Pu et al. (2025) warned that while digital technology innovation greatly decreases carbon intensity over the near term, the widespread deployment of new digital technologies inevitably triggers a long-term carbon rebound effect, potentially offsetting the initial efficiency gains.

2.2. Nonlinear Characteristics of the Digital Economy and Carbon Emissions

Based on the aforementioned contradictions, recent studies have begun to focus on nonlinear characteristics. Li & Wang (2022) were the first to find that the relationship within the carbon emissions and digital economy may exhibit a 'first decline and then rise' U-shaped pattern. Concurrently, recent research by Kang et al. (2025) indicates that the emission reduction impact of DFI demonstrates notable heterogeneity depending on city type, alongside cross-regional spatial spillover effects. However, the measurement of the exact turning point in this regard remains imprecise. Further clarifying the specific turning point of DFI's emission reduction effect and its intrinsic driving mechanisms is of great significance for grasping the optimal environmental window for digital finance development, which is exactly the direction this paper seeks to explore deeply.

2.3. Resource Endowments and the "Carbon Lock-in" Effect

Furthermore, when exploring the emission reduction effects of digital finance, existing research has overlooked the heterogeneity under the context of the 'resource curse'. As noted by Badeeb, Lean, & Clark (2017), resource-based economies face severe transformation deadlocks. Unruh (2000) defined this as the 'carbon lock-in' effect, where fossil-fuel-dependent infrastructures and institutions create systematic barriers to green transitions. Therefore, exploring the role of DFI in breaking the "high-carbon lock-in" of resource-based cities, as well as the potential rebound risks it may trigger, is a critical gap that urgently needs to be addressed in the current literature.

3. THEORETICAL ANALYSIS AND RESEARCH DESIGN

3.1. Theoretical Mechanism and Research Hypothesis

The interplay of DFI on urban carbon emissions is complex. The present paper establishes a theoretical framework from three dimensions: direct effects, nonlinear characteristics, and transmission channels.

3.1.1. Direct Emission Reduction Effect

From a financial function perspective, DFI achieves carbon emission reduction through two pathways. First, it mitigates information asymmetry. Traditional financial institutions often allocate credit to green small and medium sized enterprises because of high-risk control costs. Through leveraging big data to refine credit profiles, DFI reduces the cost of financial service access, directing funds to low-energy, high-efficiency green industries and optimizing resource allocation (Stiglitz & Weiss, 1981). Second, it replaces physical services with contactless alternatives. Micro-level evidence indicates that smartphone banking and digital financial platforms significantly reduce the need for physical travel and offline branch operations, which directly lowers the household and corporate carbon footprints (Li, 2025).

Hypothesis 1 (H1): The development of Digital Financial Inclusion can generally reduce urban carbon emission intensity.

3.1.2. U-shaped Nonlinear Characteristics

Recent literature increasingly corroborates this nonlinear dynamic. For instance, studies examining G20 countries have identified curvilinear or N-shaped relationships between financial inclusion and CO2 emissions, emphasizing that emission reduction effects often stabilize or reverse at advanced levels of financial development (Shaheen, 2025). Furthermore, recent empirical evidence from Chinese cities verifies a U-shaped relationship between digital finance and carbon emission efficiency, primarily driven by the escalating energy demands of digital infrastructure in later stages (Chen, n.d.).

While DFI inherently reduces emissions, its emission reduction benefits may follow the law of diminishing marginal utility and be constrained by the Jevons Paradox. In the early development phase, DFI adoption drives the elimination of outdated production capacity through technology spillover effects, yielding extremely high marginal emission reductions (Lange et al., 2020). However, as development enters deeper waters, two factors may weaken or even reverse emission reduction effects: First, the high energy consumption lock-in of infrastructure. Data centers, 5G base stations, and blockchain computing power are inherently energy-intensive industries, and their scaled-up electricity consumption cannot be ignored. Second, the rebound effect from scale expansion. Financial facilitation reduces production and consumption costs, potentially stimulating excessive expansion of total output and consumption, thereby offsetting the emission reduction benefits from improved energy efficiency per unit.

Hypothesis 2 (H2): The impact of Digital Financial Inclusion on carbon emission intensity follows a U-shaped pattern, initially suppressing emissions before gradually recovering.

3.1.3. The Mediating Mechanism of Technological Innovation

The endogenous growth theory asserts that technological progress serves as the fundamental driver to overcome environmental constraints (Romer, 1990). Green technology innovation, characterized by its long development cycles, high risks, and limited collateral, faces severe external financing constraints (Feng et al., 2022). DFI provides stable funding for enterprises to upgrade clean production equipment and develop green processes by expanding financing channels and reducing costs, thereby curbing carbon emissions at the source.This mechanism is strongly supported by recent international and regional evidence, which highlights that technological innovation acts as a crucial moderating and mediating variable. By alleviating corporate financial constraints, digital finance bridges the gap between financial inclusion and environmental sustainability, directly stimulating investments in cleaner technologies and reducing regional CO2 emissions (He & Jiang, 2024).

Hypothesis 3 (H3): Investment in technological innovation serves as the key intermediary channel for Digital Financial Inclusion to achieve emission reduction effects.

3.2. Data Sources and Variable Selection

This paper employs balanced panel data on 270 prefecture-level and above cities in China for the years 2012–2022. Carbon emission figures are rooted in the China Carbon Accounting Database—a source based on nighttime light inversion with proven high accuracy—and DFI Index is obtained from Peking University's Digital Finance Research Center. The main variables are defined as follows:

Dependent variable: Urban carbon emission intensity (Carbon), measured by carbon dioxide emissions per unit of GDP, mirroring the green efficiency of economic growth.

Core explanatory variables: Digital Financial Inclusion Index (Carbon) and its squared term ($\ln DIF^2$). To eliminate heteroscedasticity, the index was log-transformed.

Mediating variable: Technology investment ($\ln Tech$), measured by local fiscal expenditure on science and technology. Moderating variable: Urbanization rate (Urb), calculated as the percentage of the population that are urban permanent residents.

Control variables: To eliminate confounding factors, industrial structure (Ind, proportion of the secondary industry) and foreign direct investment ($\ln FDI$) were selected as control variables.

Table 1 presents descriptive statistics of primary variables. The dependent variable, carbon emission intensity (Carbon), has a mean of 3.361 with a large standard deviation, indicating significant disparities in low-carbon development levels across cities. DFI Index (($\ln DIF$), which is primary explanatory variable, ranges from 4.017 to 5.375, reflecting the exponential growth of digital finance during the study period. Furthermore, all control variables exhibit statistical characteristics within reasonable ranges, with no outliers detected, thus meeting the fundamental requirements for subsequent regression analysis.

Table 1: Descriptive Statistics of Primary Variables

Variable quantity	Observed number	Mean value	Standard deviation	Minimum	Median	Maximum
Carbon	2,970	3.361	3.254	0.064	2.461	28.852
ln_DIF	2,970	5.274	0.365	4.017	5.375	5.889
ln_DIF_sq	2,970	27.949	3.741	16.134	28.894	34.681
Industrial structure	2,970	44.468	10.720	11.700	44.985	81.820
Urbanization rate	2,970	0.539	0.166	0.112	0.531	1.000
ln_FDI	2,970	-4.949	1.455	-9.197	-4.637	-2.185
ln_Tech	2,970	-4.417	0.936	-7.390	-4.395	-1.575

3.3. Model Specification

For a more comprehensive verification of the aforementioned hypothesis, we construct a multi-level econometric model system based on York et al., 2003 and Balli & Sørensen, 2013.

3.3.1. Benchmark and Nonlinear Models

To test H1 and H2, we utilize a bidirectional fixed effects model containing quadratic terms:

$$Carbon_{it} = \alpha_0 + \alpha_1 \ln DIF_{it} + \alpha_2 (\ln DIF_{it})^2 + \lambda Controls_{it} + \mu_i + \delta_t + \varepsilon_{it}. \tag{1}$$

Here, μ_i represents the city-specific fixed effect, and δ_t indicates the time-specific fixed effect. If α_1<0 and α_2>0, it confirms the U-shaped characteristic. Based on the coefficient, we can calculate the inflection point value $K = -\alpha_1/(2\alpha_2)$.

3.3.2. Mechanism and Regulatory Effect Model

To validate H3, a mediation effect model was constructed using the stepwise regression method:

$$\ln Tech_{it} = \beta_0 + \beta_1 \ln DIF_{it} + \lambda Controls_{it} + \mu_i + \delta_t + \varepsilon_{it}. \tag{2}$$

To examine the moderating effect of urbanization, an interaction term is introduced into equation (1):

$$Carbon_{it} = \cdots + \gamma(\ln DIF_{it} \times Urb_{it}) + \cdots \tag{3}$$

3.3.3. Dynamic Panel Model

To account for the carbon lock-in inertia in emissions, the dynamic panel model incorporates a lagged first-period dependent variable, thereby constructing the dynamic panel model:

$$Carbon_{it} = \rho Carbon_{it-1} + \theta_1 \ln DIF_{it} + \theta_2 (\ln DIF_{it})^2 + \cdots + \varepsilon_{it}. \tag{4}$$

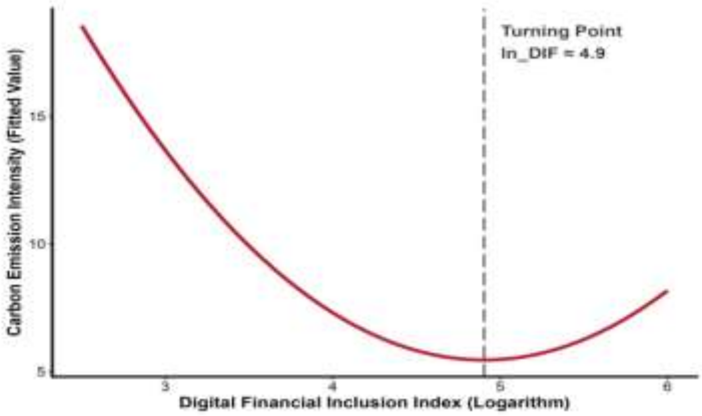
The two-step System-GMM approach is used to estimate equation (4) to address the potential endogeneity issue.

4. EMPIRICAL RESULTS AND ANALYSIS

4.1. The Analysis of Nonlinear Characteristics of Regression

Table 2 shows the core empirical results. Column (1) indicates that without the squared term, the regression coefficient of $\ln DIF$ is -2.594, statistically meaningful at the 1% level. It means that, from the perspective of overall average effects, a 1% increase in digital finance reduces carbon emission intensity by approximately 2.59%, validating H1. After introducing the square term in column (2), the coefficient of the first-order term is -22.245 and the coefficient of the second-order term is 2.268, both of which are significant and have opposite signs (negative, positive). This confirms that the relationship between DFI and carbon emissions follows a significant "U-shaped" pattern rather than a simple linear one (see Fig. 1). This finding aligns with Zhou et al. (2024), who similarly discovered a nonlinear effect of digital finance on carbon performance characterized by initial suppression followed by promotion (Zhou & Wang, 2024). This suggests that as digital finance development deepens, its marginal emission reduction effect gradually diminishes. Calculations indicate the inflection point of the curve occurs at ln DIF ≈ 4.92 (corresponding to an original index of approximately 137).

Figure 1: U-shaped Fitting Curve Graph



Using the inflection point value of 4.92 as a threshold, the sample was divided into "early development" and "mature phase" for grouped regression analysis (Table 2, columns 3-4). The results were highly enlightening: In the early development phase (ln DIF ≤4.92), the regression coefficient reached -2.381 ($p<0.01$), indicating that DFI effectively filled the gap in traditional finance during the initial stage, demonstrating strong emission reduction momentum. In the mature phase (ln DIF>4.92), the regression coefficient plummeted to -0.580 and lost statistical significance. This strongly demonstrates that after crossing the inflection point, relying solely on digital finance expansion faces emission reduction bottlenecks, with potential rebound risks triggered by infrastructure energy consumption.

Table 2: Baseline Regression and Stage-based Heterogeneity Results

Variables	(1) Baseline(Linear)	(2) Non-linear(U-Shape)	(3) Early Stage(Left of Turn)	(4) Mature Stage(Right of Turn)
ln_DIF	-2.594***	-22.245***	-2.381***	-0.58
	-0.29	-1.913	-0.776	-0.548
ln_DIF²		2.268***		
		-0.218		
Ind	-0.061***	-0.068***	-0.059***	-0.042***
	-0.003	-0.003	-0.004	-0.009
Urb	-0.584**	-0.473*	-0.612**	-0.115
	-0.21	-0.217	-0.231	-0.402
ln_FDI	0.023	0.008	0.015	0.033
	-0.015	-0.015	-0.019	-0.024
ln_Tech	-0.124***	-0.192***	-0.155***	-0.201***
	-0.028	-0.03	-0.035	-0.051
Constant	8.452***	45.320***	9.102***	3.220*
	-1.201	-3.55	-1.54	-1.89
Observations	2970	2970	1450	1520
R-squared	0.145	0.184	0.152	0.098
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

4.2. Transmission Mechanism and Regulatory Effects

The mechanism test results (detailed tables omitted due to space constraints) demonstrate that in the regression with technological input ($\ln \mathrm{T}\,\mathrm{ech}$) as the dependent variable, the coefficient of $\ln \mathrm{D}\,\mathrm{IF}$is dramatically positive (0.352, $\mathrm{p} < 0.01$). Combined with the considerable negative influence of $\ln \mathrm{T}\,\mathrm{ech}$ on carbon emissions in the baseline regression (-0.192, Table 2, Column 1), the mediation effect criterion confirms that "promoting green technological innovation" serves as a crucial channel for DFI to exert its emission reduction effects, thereby validating Hypothesis H3. The moderating effect analysis (Table 3, Column 1) implies that the coefficient of the interaction term $\ln \mathrm{D}\,\mathrm{IF} \times \mathrm{Urb}$ is 1.777 ($\mathrm{p} < 0.01$), opposite in sign to the main effect, indicating that urbanization exerts a negative moderating effect on emission reduction. This finding highlights DFI's "inclusive compensation" characteristic: in underdeveloped regions with low urbanization rates where financial exclusion is severe, the introduction of DFI effectively fills the financial service gap, providing critical support for local low-carbon development with higher marginal emission reduction benefits. In highly urbanized areas, however, the emission reduction dividends of DFI are partially diluted due to crowding effects and energy demand rigidity.

4.3. Heterogeneity and Robustness Analysis

Considering the spatial differences in China's resource endowment, we further disaggregated the sample into resource-based and non-resource-based cities, with the results shown in Table 3, columns 2–3. For resource-based cities, the absolute value of the first-order coefficient is 38.416, substantially larger than the corresponding figure of 7.365 for non-resource-based cities. It demonstrates that DFI has a stronger structural correction effect in breaking the "high-carbon lock-in" of resource-based cities (Sachs & Warner, 2001). However, its second-order coefficient is also larger (3.928), suggesting that although resource-based cities achieve significant initial emission reductions, they also face a steeper U-shaped rebound risk, necessitating vigilance against energy rebound during the digitalization process (Wang et al., 2024). Finally, to address endogeneity and verify the robustness of the results, we employ the systematic GMM method (Table 3, column 4). The results reveal that the one-period lagged carbon emission coefficient ($\mathrm{L.\,Carbon}$) is as high as 0.930 ($\mathrm{p} < 0.01$), confirming the strong path dependence of carbon emissions. After controlling for this dynamic inertia and endogeneity, the U-shaped characteristics of the core explanatory variables remain robust. Additionally, robustness tests using one-period lagged explanatory variables and excluding municipal samples all support the above conclusions.

Table 3: Mechanism, Moderation, and Advanced Heterogeneity Analysis

Variables	(1) Moderation(Interaction)	(2) ResourceCities	(3) Non-ResourceCities	(4) Dynamic(System GMM)
L.Carbon(Lagged 1)				0.930***
				-0.001
ln_DIF	-2.084***	-38.416***	-7.365	4.674***
	-0.313	-4.034	-1.924	-0.223
ln_DIF²		3.928***	0.690*	-0.495***
		-0.445	-0.22	-0.022
ln_DIF × Urb	1.777***			
	-0.275			
Controls	YES	YES	YES	YES
Constant	6.551***	89.201***	15.332***	-
	-1.502	-8.12	-4.33	-
Observations	2970	1265	1705	2700
R-squared / AR(2)	0.164	0.235	0.157	0.45 (p-val)
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "Controls" include Ind, Urb, ln_FDI, ln_Tech.

5. CONCLUSIONS

5.1. Main Conclusions

This study systematically investigates the nonlinear impact and underlying mechanisms of DFI on carbon emission intensity, utilizing balanced panel data from 270 Chinese cities spanning 2012 to 2022. The main empirical findings are fourfold:

(1) Nonlinear Duality: A significant "U-shaped" relationship exists between DFI and carbon emission intensity, with the inflection point occurring at a digital finance index of approximately 137. This indicates that while digital finance significantly curbs emissions initially, its marginal emission reduction dividend follows the law of "diminishing returns." This finding aligns with recent literature indicating that the fast growth of digital infrastructure and data centers eventually introduces severe energy-intensive burdens that can offset early emission reduction gains (Dong et al., 2025).

(2) Transmission Mechanism: Alleviating financing constraints to promote green technological innovation acts as the critical intermediary channel. Digital finance effectively bridges the funding gap for environmental sustainability, accelerating the clean transformation of enterprises and directly lowering carbon intensity (Jiang et al., 2024).

(3) Heterogeneity and Moderation: The urbanization rate exerts a negative moderating effect on emission reductions, confirming the "inclusive compensation" feature of DFI. Furthermore, resource-based cities exhibit a stronger structural correction effect in the early stages but face a much steeper U-shaped rebound risk subsequently (Kang et al., 2025).

(4) Dynamic Characteristics: Urban carbon emissions demonstrate significant path dependence and a robust "carbon lock-in" effect, emphasizing the immense inertia of historical high-carbon development models.

5.2. Policy Implications

Based on the empirical findings and the urgent requirements of the global "dual carbon" goals, we propose the differentiated policy recommendations as follows:

(1) Implement Stage-Specific Digital Finance Strategies: Policymakers must abandon "one-size-fits-all" approaches and adopt threshold-based governance. For regions with a DFI index below 137, local governments should accelerate the deployment of digital financial infrastructure to fully unleash its "inclusive compensation" dividend and lower regional carbon abatement costs. Conversely, for mature regions that have crossed the inflection point, policies must pivot from scale expansion to quality enhancement. This involves strictly regulating the energy consumption of new digital infrastructure (e.g. 5G base stations and data centers) to flatten the U-shaped rebound curve.

(2) Construct Targeted "Digital-Green" Financial Conduits: Given that green technological innovation is the core mediator, financial institutions should design specialized digital credit products that strictly ring-fence funds for green R&D and clean production. By utilizing big data and blockchain for end-to-end capital tracking, policymakers can accelerate the market exit of high-pollution and high-emission enterprises while preventing digital funds from flowing into the scale-expansion of carbon-intensive activities.

(3) Design Asymmetric Low-Carbon Transitions for Resource-Based Cities: Recognizing their steep rebound risks, resource-based cities must capitalize on the early "structural correction" window provided by digital finance. Local governments in these areas should mandate strict renewable energy purchasing quotas for newly established digital platforms and couple digital integration directly with industrial decarbonization metrics to permanently break the "carbon lock-in."

(4)Foster Spatially Synergistic Governance: Because highly urbanized areas experience crowding effects that dilute emission reduction benefits, national digital economy strategies should intentionally direct digital financial resources toward less developed, lower-urbanization regions where the marginal emission reduction effects are strongest. Cultivating this spatial spillover effect will maximize the marginal carbon productivity of digital finance at a macroeconomic level.

REFERENCES

Badeeb, R. A., Lean, H. H., & Clark, J. (2017). The evolution of the natural resource curse thesis: A critical literature survey. *Resources Policy*, 51, 123-134.

Balli, H. O., & Sørensen, B. E. (2013). Interaction effects in econometrics. *Empirical Economics*, 45(1), 583-603.

Chen, Y. The nonlinear impact of digital finance on carbon emission efficiency: Evidence from Chinese cities. *Frontiers in Environmental Science*, 14, 1739323.

Dong, X., Lv, S., & Wu, S. (2025). The green promise of digital finance: Exploring its effects and mechanisms on carbon reduction. *China Economic Review*, 90, 102374.

Feng, S., Zhang, R., & Li, G. (2022). Environmental decentralization, digital finance and green technology innovation. *Structural Change and Economic Dynamics*, 61, 70-83.

Gao, X., & Li, J. (2024). China's digital economy: A dual mission of carbon-emission reduction and efficiency enhancement. *Sustainability*, 16(6), 2351.

He, Q., & Jiang, H. (2024). Digital inclusive finance, digital technology innovation, and carbon emission intensity. *Sustainability*, 16(15), 6407.

Jiang, Y., Zhao, R., & Qin, G. (2024). How does digital finance reduce carbon emissions intensity? Evidence from chain mediation effect of production technology innovation and green technology innovation. *Heliyon*, 10(9), 373-399.

Kang, J., Duan, S. Y., Yan, S. X., Xue, S. Y., Li, Y. H., Li, N., ... & Liu, G. H. (2025). Impact of Digital Inclusive Finance on Carbon Emissions under the Dual-Carbon Goal: Evidence from Nine Regions in the Yellow River basin of China. *Water Economics and Policy*, 2440011.

Lange, S., Pohl, J., & Santarius, T. (2020). Digitalization and energy consumption. Does ICT reduce energy demand? *Ecological Economics*, 176, 106760.

Le, T. H., Le, H. C., & Taghizadeh-Hesary, F. (2020). Does financial inclusion impact CO2 emissions? Evidence from Asia. *Finance Research Letters*, 34, 101451.

Li, Y. (2025). How does digital financial inclusion affect households' CO2? Micro-evidence from an emerging country. *Journal of Economics and Business*, 133, 106222.

Li, Z., & Wang, J. (2022). The dynamic impact of digital economy on carbon emission reduction: evidence city-level empirical data in China. *Journal of Cleaner Production*, 351, 131570.

Pu, Z., Qian, Y., & Liu, R. (2025). Is digital technology innovation a panacea for carbon reduction?. *Technological and Economic Development of Economy*, 31(2), 496-524.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), 71-102.

Sadorsky, P. (2012). Information communication technology and electricity consumption in emerging economies. *Energy Policy*, 48, 130-136.

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, 45(4-6), 827-838.

Shaheen, R. (2025). Investigating the relationship between financial inclusion and CO2 emissions in G20 countries: a quantile-on-quantile approach. *Social Responsibility Journal*, 21(10), 2152-2172.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3), 393-410.

Unruh, G. C. (2000). Understanding carbon lock-in. *Energy Policy*, 28(12), 817-830.

Wan, J., Pu, Z., & Tavera, C. (2023). The impact of digital finance on pollutants emission: evidence from Chinese cities. *Environmental Science and Pollution Research*, 30(15), 42923-42942.

Wang, S., Zhou, A., Zhang, Y., Yang, G., & Peng, X. (2024). Enhancing carbon emission performance in resource-based cities through technological innovation under China's digital economic development strategy. *Frontiers in Environmental Science*, 12, 1397063.

York, R., Rosa, E. A., & Dietz, T. (2003). STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. *Ecological Economics*, 46(3), 351-365.

Zheng, F., Chen, S., & Wang, X. (2025). How the impact and mechanisms of digital financial inclusion on agricultural carbon emission intensity: new evidence from a double machine learning model. *Frontiers in Environmental Science*, 13, 1549623.

Zhou, B., & Wang, Y. L. (2024). The nonlinear effects of digital finance on carbon performance: Evidence from China. *Journal of Innovation & Knowledge*, 9(2), 100484.

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

DETERMINANTS OF FOOD SECURITY AND COPING STRATEGIES AMONG RURAL HOUSEHOLDS IN BAKA-DAWULA ARI WOREDA, SOUTHERN, ETHIOPIA

DOI: 10.17261/Pressacademia.2026.2036
JBEF- V.15-ISS.1-2026(8)-p.86-99

Abdi Alemayehu Angasu[1], Belete Meka Gerta[2]
[1]Jinka University, Department of Economics, Jinka, Ethiopia.
alemayabdi@gmail.com, ORCID: 0009-0002-2831-6624
[2]Ethiopian Civil Service University, Department of Department of Development Management, Addis Ababa, Ethiopia.
mekabeletebbb@gmail.com, ORCID: 0009-0009-7982-1904

Date Received: March 21, 2026 **Date Accepted:** May 20, 2026

To cite this document
Angasu, A.A., Gerta, B.M., (2026). Determinants of food security and coping strategies among rural households in Baka-Dawula Ari Woreda, Southern, Ethiopia. Journal of Business, Economics and Finance (JBEF), 15(1), 86-99.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2036

ABSTRACT
Purpose- Food is any substance that people eat and drink to maintain healthy and productive life as well as growth. While food security is when all peoples access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life. However, now a day food security is one of the main confuse factors of economic growth of developing county in general, Ethiopia in particular. Therefore, the main aim of the study was to examine the determinants of food security and coping strategies among rural households' in Baka Dawula Ari district, Ari Zone, Southern Ethiopia regional state.
Methodology- The study employed quantitative research approach and explanatory research design. To collect data from the sample respondents of 269, multistage random sampling techniques used. For the data analysis, both descriptive statistics and econometric models particularly the logistic regression model used.
Findings- The results of logistic regression analysis indicated that age household head, education level, access to extension services, households participate in off farm activities, households ownership of oxen, livestock ownership and cultivable land size were positive and significantly influencing household food security in the study area. while, family size is negative and significantly influencing household food security in the study area. The survey result of coping strategies state that majority of household was used mechanisms such as reduced number of meals eaten in a day, selling small animals, relay on causal labor, borrow money/food from relatives and selling firewood and charcoal were used in the period of shortage of food.
Conclusion- Based on the findings, the local government development strategies need to encompass education programs; attempt short-term training center should experience in a strategic and organized to improve household education. Additionally, facilitate starting income to participate on off-farm activities, improving the quality of the land through improved soil and nutrient management to increase food production, and supply better veterinary services to improve livestock markets to achieve household food security.

Keywords: Household Food Balance Model, Food security, coping strategies, Baka Dawula Ari district, binary logit model
JEL Codes: Q18, Q12, C25

1. INTRODUCTION

1.1. Background of the Study

In the worldwide food insecurity remains a major challenge, mainly among the rural areas of developing nation. Women and children are most vulnerable to this phenomenon (Ridwan et al. 2020). Hunger and undernourishment are the main challenges of today's world, and 960 million people are hungry and undernourished (Mesfin et al. 2021). As a result, food security continues to be the greatest problem of economic development and adverse effect on public health on the globe up today (Gizachew et al. 2023). Food security exists when all people have physical, social and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life (Abduselam, 2017; Gizachew et al. 2023). As population of a county at all-time are available food, access food, nutritionally adequate food in terms of quantity and acceptable within the given culture the county is said to be food secure (Seid and Biruk, 2019). Additionally, households assumed food secure when they produce adequate essential food to encounter their daily needs or when they have enough income to purchase food from the market (Workineh, 2024).

The concept of food security encompasses multiple dimensions such as food availability, food access, food utilization and food stability. Food availability refers to the existence of food stocks for consumption. Household food access is the ability to acquire sufficient quality and quantities of food to meet all household members' nutritional requirements. Access to food is determined by physical and financial resources, as well as by social and political factors. Utilization of food depends on how food is used, whether food has sufficient nutrients, and a balanced diet can be maintained (FAO, 2014). Food security issues become a critical concern and top priority for developing countries, particularly Ethiopia (Girma, et al., 2023). From Sub-Sahara African countries, Ethiopia is one of the poorest and greatest foods insecure counties (Desta and Negussie, 2017). Ethiopia is one of the countries with the most food insecurity and famines in the developing world (Adimasu et al, 2019). In Ethiopia, food insecurity is the most serious problems (Abebaw and Mesele, 2022) and development challenge after some economic recovery has shown in the county (Arragaw and Argaw, 2024). Additionally, under nutrition are significant problem of the county economic growth (Nigusu and Shewadinber, 2022).

In rural household level food security status of Ethiopian country is found to be a most horrible stage and the county confused by different factors (Wondim et al., 2022). For instance, a study conducted by Amanuel (2025) indicated that gender, educational level, farmland size, livestock holding, access to credit, improved seed, and social capital positively affect households' food security status, whereas family size, market distance, and natural shocks negatively influence food security. Additionally, education status, off-farming activities, livestock ownership, family size, farm size, number of oxen, expenditure on agricultural technology, agro ecology zone and distance from market center are statistically significant determinants of food security in terms of diet quantity (Tamene and Ermias, 2023).

As a result, several efforts have made to recover the general challenges of food security; as it is a main difficult in the rural areas of Ethiopia (Aweke et al 2022). The country's population has affected through chronic and transitory food insecurity and in the time, community becoming an increasingly severe living condition (Adimasu et al, 2019). Additionally, poor households are insufficient purchasing power to improve food security to the establishment of food coping mechanisms to alleviate insufficiency (Wuryaningsih et al. 2022). Food insecure households engaged in different coping strategies with the respective level of food insecurity from intake inedible, low-quality foods to the greatest severe of migrating and begging for food (Gizachew et al. 2023). The coping strategies is mechanisms in the time of food shortages, including the sale of livestock, productive assets, the receipt food aid, the participation in small trade, the reduction in the number of meals, temporary migration (Girma, et al., 2023).

Mostly used coping strategies rural household in Ethiopia are agricultural employment, sailing livestock, sale of wood and charcoal, small trading, residues crop and reduction of food consumption (Abdukerim et al., 2022). Additionally, household was coping food shortage by relying on less preferred and cheapest food and borrowing food to utilizing foods cope up to food shortage and starvation (Mesfin et al. 2021). In addition, reducing the number of meals, working as daily laborers, borrowing money, migrating for seasonal work, relied on food aid, dropped children from school, or sent them to live with relatives are most used strategies (Tadese and Yabsira, 2025).

1.2. Statement of the Problem

Ethiopian rural areas security at household food level issues has remained a challenging goal until today (Ahmed et al., 2018). There are numerous of influences household to put into the problem of food insecurity in rural Ethiopia. Among them sex of household head, landholding size, livestock and off farm activity are important factors (Tesfahun, 2022; Girma, et al., 2023). Additionally, family size, dependent ratio, cultivated land size, numbers of oxen and fertilizer use are factors affect household food security (Ahmed et al., 2018; Girma, et al., 2023). In addition, household head age, education level, income and grant types are other factors of household food security (Mazenda et al., 2022; Girma, et al., 2023).

Therefore, existence of various studies have been conducted on investigating the determinants of food security by employ household food insecurity access scale as an indicator of food security status measurement (Ahmed et al., 2018; Mojela et al., 2018; Adimasu et al. 2019; Mazenda et al., 2022) and giving little attention to the availability and utilization of food. As a result, those studies concentrated on the diet quantity aspect of food security and ignored the diet quality aspect of food security (Abayineh and Belay, 2017 and Abebaw and Mesele, 2022). In addition, many similar studies (Gebremariam et al., 2019; Girma, et al., 2023 and Adane et al., 2023) have been done focused on food production or availability giving little attention for access and utilization of food. However, when repeatedly exposed to recurrent drought and famine the total production is persistently inadequate to cover food requirement of the population.

Additionally, the previous studies (Abayineh and Belay, 2017; Wondim et al., 2022; Ahmed et al., 2018; Alemseged et al., 2018; Hailu, 2022; Amanuel, 2025; Abebaw and Mesele, 2022) has been overlook coping strategy by which household or community members used to meet their relief and recovery food insecurity situation. However, food security status and its coping mechanism are complex issues, and which are not the same from Household to Household within environment (Tesfahun, 2022). On the other hand, some studies (Mazenda et al., 2022; Adane et al., 2023; Tadese and Yabsira, 2025) conducted in urban case study rather than rural household food security. However, food security at the household level in the rural areas of Ethiopia has remained a challenging goal until today (Ahmed, et al., 2018).

Moreover, Household Food Balance Model the ability to establish access to productive resources such as land, livestock, agricultural inputs and family labor combined to produce food or cash. Additionally, it identified pillars of food security such as availability, accessibility, and utilization of food (Meskerem and Degefa, 2015). Therefore, there is lack of study conducted in considering the rural farm households in the case of studies. Thus, against the above background, this study was utilizing by Household Food Balance Model (HFBM) to quantify the driving force of food security to examine the determinants of rural households' food security and enhancing coping mechanism of rural households in Baka Dawula Ari district, Ari Zone, Southern Ethiopia regional state.

1.3. Objectives of the Study

The general objective of this study was to examine the determinants of rural households' food security and their coping strategies in Baka Dawula Ari district, Ari Zone, Southern Ethiopia regional state. Specifically, this study aimed to examine the rural households' food security status in the study area, investigate the determinant factors that affecting rural household food security in the study area and identify the coping methods used by rural households in the study area.

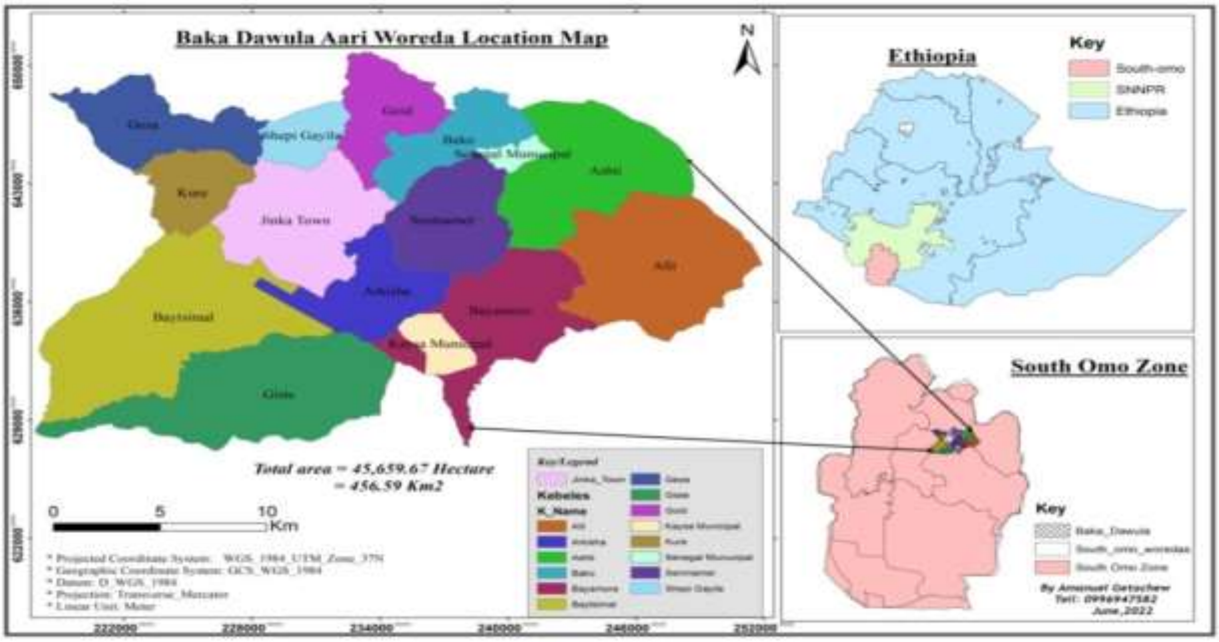
2. RESEARCH METHODOLOGY

2.1. Description of the Study Area

Geographically, Baka Dawulla Ari District, which is one of the four districts in Ari zone of Southern Ethiopian Regional state. It is bordered on the southeast and southwest by Salamago distrct specially Mago National Park, on the North by South Ari district, on the North-east by Uba Debretsehay district, most of its eastern boundary is defined by the course of the Kako River. The total population of the Baka Dawulla Ari district is estimated to be 49,623 (88.7 % rural and 11.3 % urban). In addition, the gender distribution is (male 24,663 and female 24,960). The district is sub divided into twelve (12) rural kebeles (Ari Zone Agriculture and Rural Development Office Report, 2023). Astronomical location of the district is between 4043' to 6046'North latitude and 35079' to 360.06'East longitude. It is located about in the south, 590 km from Addis Ababa by Sellam ber and 212 km from the regional city of Southern Ethiopian state Wolaita Soddo. Ari Zone has four woredas and two municipality city administrations (Ari Zone Agriculture and Rural Development Office Report, 2023).

Agro-ecologically, the landmass of the district lies between (1500–2500 Meters above sea level altitude. Annual mean minimum and maximum temperature respectively are 12ºc and 21.7 ºc. The amount of rainfall it received ranges from 977mm to 1300 mm (Ari Zone Agriculture and Rural Development Office Report, 2023). Rainfall distribution is seasonal that means the rainy seasons are Belg (February to April) and the main rainy season Meher (July to September). According to the districts Agriculture and Natural Resource Office (2023), the total area of the district estimated to be 49,659.67 hectare. The major land uses patterns are private holding (Farming), communal (grazing) and forest land. The current land use coverage's 20,258.06 (41) ha are cultivated, grazing lands is 3,121(6%) ha, Forestlands are 11,518.44 (23) ha, Marshy land is 2,375 (5%) ha and the rest 12,387.17 (25%) ha land is used for different institutions and residence of study area people.

Figure 1: Location Map of the Study Area



The farming practice in the district is mixed farming systems which are crop production and livestock keeping. The major means of livelihood in the area are subsistence rain fed agriculture, traditional weaving and involvement in off-farm activities.

The major of the district is cropping production were Maize, Wheat, Teff, Sorghum. Potato, Barely, Bean and Haricot Bean (local name Adongare) are common double-time harvesting crops per year. Additionally, coffee, chat and banana farm is more source of income of the community (Ari Zone Agriculture and Natural resource office, 2023).

2.2. Sampling Technique and Sample Size

The determination of sample size of this study is based on the formula of Yamane (1967); which is given by:

$$n = \frac{N}{1+N(e^{2)}} \tag{1}$$

Where; n= sample size; N= population size and e=level of precession; e = 0.05.

$$n = \frac{959}{1+959(0.05)^2} = 282$$

However, data collected and analyzed in response to answer research objectives based on the responses of those selected samples with the help of a questionnaire. Therefore, a sample of the study 282 questionnaires were distributed to the respondents and out of it valid questionnaires 269 were received back properly, while 13 (4.61%) were incomplete. Thus, the questionnaire return rate was 95.39 percent, which is adequate to come up with valid study findings.

A multi-stage sampling technique was utilizing for select the household heads in the study. First, Baka Dawulla Ari woreda purposively selected from four rural woredas under Ari zone. In the second stage, 12 rural kebeles under the Baka Dawulla Ari woreda were stratifies into three strata groups (Kolla, Woyina Dega and Dega agro-ecological zones) based on their agro-ecological characteristics. Then, four rural kebeles selected based on population density and agro-ecology zones randomly. Finally, based on the total number of households in each kebeles, the number of households to be included in the study from each kebele was selecting by using proportionate random sampling technique.

Table 1: Proportionate sample size distribution

Name of kebele	Total Household in each kebele	Sample Household from each kebele	Respondent Rate
Arkisha	286	84	81
Baitsimal	232	68	68
Goid	228	67	62
Sanmamer	213	63	58
Total	959	282	269

Source: Ari Zone Agriculture and Natural resource office, 2023

2.3. Methods of Data Collection

The study was collected data from sample households by using a focus group discussion and structured questionnaire. The questionnaires were used to generate quantitative and qualitative data regarding household demographic, socioeconomic characteristics, determinants and management practices understandings of household's food security status, and coping strategies during food shortfall. To gather the information from selected respondent like Sex, Age, household size, educational status, marital status, cultivable land size, livestock and oxen owner, the access to credit, access to extension service, adaptations of improved technologies, quantity of food item and main reasons to food shortage was collected using a structured closed and open ended questions.

2.4. Method of Data Analysis

In this study, descriptive and econometric data analysis methods were used. The descriptive statistics of this study explained the socioeconomic, institutional and demographic characteristics of the farm households in the study area. Descriptive statistical tools such as mean, percentage, frequency, standard error, standard deviation, minimum, and maximum were analyzing after the data collected, edited, coded, and labelled. The chi-square (χ2) and Student's t-test were used to test the statistical significance of the dummy and continuous variables.

2.4.1. Measurement of Food Security Status

The study was using Household Food Balance Model identified from the theoretical and empirical literatures. To identify the food secure and insecure households, household food balance sheet was employed. In the calculation of kilocalories intake, the amounts of calorie available to a household were determined through an equation termed as household food balance model (Eq. (1)), which was later used for different studies Abayineh and Belay (2017) and Hailu, (2022).

Household food balance model is expressed as:

$$NG_i = (PR_i + PU_i + FW_i + RG_i) - (CL_i + CS_i + TM_i + SE_i + PB_i) \tag{2}$$

Where, the index i runs for 1, 2,…, 269 sample household of the study,

NG_i is net grain food available for household i, PR_i is total grain produced by household i, PU_i is total grain purchased by household i, FW_i is total grain obtained through food-for-work by household i, RG_i is total relief grain food received by household i, CL_i is post-harvest crop losses to household i, CS_i total crop utilized for seed by household i, TM_i is total marketed output by household i, SE_i is grain used for social events by household i and PB_i is repayment of grain borrowed by household i.

Finally, food security in the present study was measured into the following four steps. The first, net grain accessible for each household in kilogram (NG_i) was changed into corresponding total kilocalories using conversion factors of Ethiopia. Second, the food supply at the household level calculated in first step was used to compute calories available per person per day for each household. Third, following Ethiopian food security strategy, 2,100 kcal calories per person per day were used as a measure of calories required to enable an adult to live a healthy and moderately active life. Then, using 2,100 kcal calories as cut off point, a household whose available daily per capita calories is getting 2,100 kilocalorie and above are food secured (which take 1), while a household who get less than 2,100 kcal calories was considered as food insecure (which take 0).

2.4.2. Econometric Model Specification

Three models have been proposed for estimating binary choice models: the linear probability model, logit, and Probit models represented by a linear probability function, logistic distribution, and normal distribution function, respectively (Gujarati, 2013). In principle, one can substitute the probit model for logistic model, as their formulations are quite comparable; the main difference is that the logistic model has slightly flatter tails than the cumulative normal distribution, i.e., the probit curve approaches the axes more quickly than the logistic curve (Gujirati, 1995). On this score, the logit model is generally used in preference to probit. It also noted that the logistic distribution has an advantage over the others in the analysis of dichotomous outcome variables, because it is extremely flexible and easily used model from the mathematical point of view and results in meaningful interpretations (Gujarati, 2013). Therefore, the logistic model was selected for this study.

The Gujarati (2013) logit model is expressed as follows by

$$P_i = E\left(Y = {}^{1}/_{X_i}\right) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_i)}} \tag{3}$$

For ease of exposition, Eq. (2) can be expressed as:

$$P_i = \frac{1}{1+e^{-Z_i}} \tag{4}$$

Where, $Z_i = \beta_0 + \beta_1 X_i$. If P_i, is the probability of being food secure, then the probability of being food insecure is given by $1 - P_i$, which is expressed as follows by Eq. (4):

$$1 - P_i = \frac{1}{1+e^{Z_i}} \tag{5}$$

Therefore, take the ratio of the probability of an event happening (P_i) to the probability of an event not happening ($1 - P_i$) and the resulting ratio is called odds ratio, this can be written as Eq. (5):

$$\frac{P_i}{1-P_i} = \frac{1+e^{Z_i}}{1+e^{-Z_i}} = e^{Z_i} \tag{6}$$

Where, ${}^{P_i}/_{1-p_i}$ is simply the odds ratio in favor of food security; the ratio of the probability that the household will food secure to the probability that it will food insecure.

Taking the natural log of Eq. (5) above, it is possible to arrive at a log of odds ratio, which is linear not only inX_i, but also in the parameters.

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta_0 + \beta_1 X_i \tag{7}$$

Where, L_i is log of odds ratio; P_i is the probability of being food secure ranging from zero to one; Z_i is a function of n-explanatory variables (X_i) and is expressed as Eq. (7):

$$Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots .. + \beta_n X_n + \varepsilon_i \tag{8}$$

Where, β_0 is the intercept or constant term; β_1, β_2……β_n are the slope of the equation in the model (parameters to be estimated).

Finally, an empirical model for the determinants of rural households' food security was specified as follow:

$$HFS_i = \beta_0 + \beta_1 AGE_i + \beta_2 SEX_i + \beta_3 CRP_i + \beta_4 EDU_i + \beta_5 FMZ_i + \beta_6 DPR_i + \beta_7 OFA_i + \beta_8 CLS_i + \beta_9 TLU_i + \beta_{10} NOX_i + \beta_{11} FTU_i + \beta_{12} AES_i + \beta_{13} LTS_i + \varepsilon_i \tag{9}$$

Where, HFS_i is Household Food security status, AGE_i is Age of Household Head, SEX_i is Sex of the household head, CRP_i is Credit Participation, EDU_i is Education level, FMZ_i is Family size, DPR_i is Dependency ratio, OFA_i is Off–farm activities, CLS_i is Cultivated land Size, TLU_i is Tropical livestock unit, NOX_i is Number of oxen owned, FTU_i is Fertilizer uses and AES_i is Access to Extension Service, LTS_i is Land tenure security.

Table 2: Description of Variables and Expected Relationship

Covariates	Description of Variables	Measurement	Expected Sign
HFS_i	Dummy variable Food security status	1 if secure, &0 otherwise	
AGE_i	Age of Household Head	Continuous	+
SEX_i	Sex of the household head	1 if male, &0 otherwise	±
CRP_i	Credit Participation	1 if participated, &0 otherwise	+
EDU_i	Education level	Continuous	+
FMZ_i	Family size	Continuous	−
DPR_i	Dependency ratio	Continuous	−
OFA_i	Off–farm activities	1 if participated, &0 otherwise	+
LTS_i	Land tenure security	1 if tenure, &0 otherwise	+
CLS_i	Cultivated land Size	Continuous	+
TLU_i	Tropical livestock unit	Continuous	+
NOX_i	Number of oxen owned	Continuous	+
FTU_i	Fertilizer uses	1 for users, & 0 otherwise	+
AES_i	Access to Extension Service	1 if access, & 0 otherwise	+

3. RESULTS AND DISCUSSION

3.1. Household Food Security Status

The study measure household food security status by using household food balance model, 2,100 kcal calories per person per day were used as a measure of calories required as cut off point, a household whose get daily per capita calories are food secured and household whose get less than cut off point are food insecurity status. Based on this procedure, 114 sample households were to be able to meet the minimum subsistence requirement. The mean dietary energy available for food secured households was 2512.679 kcal calories per person per day adult equivalent, while 1275.17 kcal calories per person per day adult equivalent for the insecure group households. It showed that the mean of dietary energy supply for food secure households was larger than that of food insecure groups. Their mean difference between the two groups was statistically significant at (p<0.01).

Table 3: household food security status

Households	Mean	Std. Dev.	Minimum	Maximum	T-value
Food secure (n= 114)	2512.679	626.3377	2140.349	7523.336	-23.2423***
Food insecure (n= 155)	1275.17	187.0951	1030.312	2070.897	
Total (n=269)	1799.616	748.9207	1030.312	7523.336	

Source: Survey data by authors

3.2. Descriptive Result of Categorical Explanatory Variables

The results of the descriptive statistics of categorical variables such as sex, household head, credit access, and land tenure security and fertilizer user revealed that there are insignificant difference among food insecure and food insecure households. However, the two groups found to differ significantly in their access to extension services and off-farm participation households are significant difference between food secure and food-insecure households.

Based on descriptive statistics, out of the food secure households, 70 (61.4%) households were received an extension service, while the rest of 44 (38.6%) did not receive extension services. On the other hand, about 70.32% of the food insecure household did not receive extension services while 29.68% of insecure households had access to extension services. Furthermore, this imply that more food insecure household not get extension services the chi-square test for this variable shows a significant difference between food secure and food-insecure households (χ2 = 26.96).

Table 4:Descriptive result for dummy variables

Variables	Food Secured (N=114)		Food insecured (N=155)		Chi-square test
	Frequency	Percent	Frequency	Percent	
Sex (if male)	102	89.47	139	89.68	0.0029
Credit access (if yes)	44	38.60	58	37.42	0.0387

Land tenure security (if secured)	104	91.23	131	84.52	3.0450
Off-farm income (if yes)	51	44.74	47	30.32	5.8932**
Extension contact (if yes)	70	61.40	46	29.68	26.9571***
Fertilizer user (if yes)	52	45.61	83	53.55	1.6541

Sources: own survey Note: *p<0.05, **p<0.01 & ***p<0.001

As indicated descriptive statistics result, out of the food secure sample households, the proportion of households are off-farm activities participant was 44.74 percent while, only about 30.32 percent of households food insecurity were participate on off-farm activity. The result implies that most food insecure household was not participating on off-farm activity. The chi-square value also confirmed a significant difference between food-secure and insecure households with respect to access to non-farm activities ($\chi2 = 5.89$).

3.3. Description Result of Continuous Variables

Moreover, the descriptive results for continuous revealed variables such as age of household head, education level of households, family size, land holding, cultivated land size, number of livestock and number of Oxen is significant difference between food secure and food-insecure households. The survey result indicated that on average the proportion of education food secure household heads were larger than the proportion of education food insecure household heads. Finally, the t-test is this variable is statistically significant difference between the two groups in terms of years of study at 10% level of significance.

Table 5: Descriptive results for continuous variables

Variables	Food secured	Food insecured	T-test
	Mean	Mean	
Age of household head	49.79	46.56	-3.1406***
Education level	2.03	2.27	1.6100*
Family size	7.59	8.08	2.6526***
Land holding	3.45	3.09	-1.8386**
Cultivated land size	2.13	1.26	-3.5318***
Dependency ratio	1.13	1.03	-0.6768
Number of livestock	22.28	13.68	-6.9923***
Number of Oxen	4.33	2.53	-7.6841***

Sources: own survey Note: *p<0.05, **p<0.01 & ***p<0.001

Moreover, the survey result shows, the mean 3.45 household are holding of land were food secured household heads, while household heads 3.09 household were land holding. The results of the statistics of the t-test revealed a statistically significant difference between the two groups in terms of the holding of land in the study area (t-value = -1.8386). Additionally, the mean cultivated land size are 2.13 hectares were cultivated by food secured household heads, while food insecure household heads cultivate 1.26 hectares cultivated land size. The results of the statistics of the t-test revealed a statistically significant difference between the two groups in terms of the cultivated land size in the study area (t-value = -3.5318). This result support that farmers who have larger cultivated land size are more likely to be food secure than those who cultivated smaller land size due to the fact that there is high possibility to produce more food.

According to the results of the descriptive statistics, the average family size for secured food secured and food insecure households was 7.59 and 8.08, respectively. The statistics of the t-test show statistically significant differences in the family size for food secure and insecure households in the study area (t-value = 2.6526).

Additionally, oxen ownership is also an important variable in the study areas that almost entirely rely on traditional farming methods, thereby significantly affecting households crop production. Oxen the sole provider of draft power and determinant of on time land preparation, is not uniformly distributed between food secure and insecure households. The average number of oxen owned by food secure was 4.33, while the average number of oxen owned by food insecure was 2.53, respectively. The result of the t-test indicated a significant difference between food security and food insecurity in oxen ownership at a 1% significance level (t-value = -7.6841).

Livestock play an imperative role to the households' food security through increasing purchasing power of household to buy food from market. Livestock considered as a means of food security and means of coping mechanism during crop failure. Number of Livestock (Excluding oxen) was difference between food secure and insecure households. On average households food secure have 22.28 livestock, while food insecure households have 13.68 livestock. Furthermore, the result of the t-test confirmed statistically significant mean differences between the two groups at 1% level of significance (t-value = 1%).

3.4. Econometric Results

The results of logistic regression showed that variable such as age household head, education level, access to extension services, households participate in off farm activities, households ownership of oxen, livestock ownership and cultivable land size were positive and significantly influencing household food security in the study area. On the other hand, family size is negative and significantly influencing household food security in the study area.

Table 6: The Binary Logistic Regression Model Result of Food Security Determinants Result

Variables	Coefficient	Std. Err.	Z	P>|z|	Odds ratio
Age household head	.0450687	.0212341	2.12	0.034**	1.0461
Sex	-.60387	.5944209	-1.02	0.310	.5466918
Educational level	.5870828	.140122	4.19	0.000***	1.798734
Family size	-.2502614	.1237289	-2.02	0.043**	.7785972
Access to extension services	1.689036	.3820694	4.42	0.000***	5.414261
Land tenure security	.7178342	.6371967	1.13	0.260	2.049989
Off farm participate	1.119204	.3850455	2.91	0.004***	3.062416
Number of Oxen	.3458699	.1115266	3.10	0.002***	1.413219
Number of livestock	.0614106	.0229687	2.67	0.008***	1.063335
Credit access	.0770811	.3819348	0.20	0.840	1.08013
Dependency ratio	.1963739	.1692149	1.16	0.246	1.216982
Cultivated land size	.4331343	.2217647	1.95	0.051*	1.542083
Fertilizer user	-.4812896	.3887653	-1.24	0.216	.6179859
_cons	-7.239629	1.764592	-4.10	0.000***	.0007176
Number of obs = 269 Log likelihood = -111.3008	LR chi2(13) = 144.04 Pseudo R2 = 0.3929		Prob > chi2 = 0.0000		

Notes: *** p < 0.01, ** p < 0.05, * p < 0.1; Dependent variable: =1 if the household is food secured, 0 otherwise.

Age household head: The age household head was found to be positive and statistically significant at five percent level of statistical significant. The positive relationship implies that, the oldest households is more life experience to improve food production to access food. The implication is that, keeping other factor constant, as the age household head increase by one year the probability of household being food secures increases by a factor of 1.0461. The results were similar to study conducted by Seid and Biruk (2019); Abebaw and Mesele (2022); Wondim et al. (2022).

Family size: The coefficient of family size was found to be negative and statistically significant at five percent level of statistical significant. This indicates household that has more family size probably food insecure from those small family size household head. Citrus paribus, a small family member to household probability of increase food secure by 77.85 percent and statistically significant at 5 percent significant level. This finding was in line with studies by Desta and Negussie (2017); Mojela et al (2018); Hailu (2022); Aweke et al. (2022); Abebaw and Mesele (2022).

Education level of the household head: The household educational status was found to be positive and statistically significant at one percent level of statistical significant. This result implies that, household heads with relatively better education are more likely to be food secure than those headed by illiterate household heads. Thus, as year of school increase by one year the probability of household to be food secure increase by 1.798 holding all other factors constant. Therefore, education levels become important factors to improve food security status. This goes in line with some previous studies (Mazenda et al., 2022; Gizachew et al. 2023; Abebaw and Mesele, 2022; Amanuel, 2025) which showed statistically significant and positive relationship between level of household head education and the probability of being food secure.

Access to extension services: The coefficient of access to extension services was found to be positive and statistically significant at one percent level of statistical significant. This indicates household that can access extension service probably food secure from those who did not get an extension service household head. Citrus paribus, a percentage increase in extension services to household probability of increase food secure by 5.414 percent and statistically significant at 1 percent significant level. This finding was in line with studies by Abayineh and Belay (2017); Adimasu et al. (2019); Wondim et al., (2022) and Girma et al (2023).

Off farm participate households: The study result showed that the effect of off-farm income on household food security was positive and statistically significant at one percent level of significance. This implies that off-farm activities are important activities done by rural households get extra income to supplement their livelihoods. Households who engaged in off-farm activities are less risk-averse than farmers without sources of off-farm income are. The result of odds ratio state that as off-farm income by one Ethiopian Birr (ETB) the probability of household in favor of being food secures to increase by 3.062 percent other factors remains constant. This result is consistent with studies conducted by Seid and Biruk (2019); Abebaw and Mesele (2022); Wondim et al. (2022); Hailu (2022); Nigusu and Shewadinber (2022). As study Abebaw and Mesele (2022) expanding the access to off-farm activities to increase household income access food from market to improve rural food security status.

Number of oxen owned: The result of the model revealed that this variable found to be statistically significant at one percent probability level and has positive association with household food security status of rural households in study area. Oxen are important economic assets that can support households attain higher production by cultivating their land effectively and on time. Households that has more oxen is most probably being food secure. The result indicates that, as number of oxen increase by one the probability of household to be food secure increases by a 1.413 citrus paribus. This finding was in line with study conducted by Meskerem and Degefa (2015); Gebremariam et al. (2019); Tamene and Ermias (2023).

Livestock ownership: It refers to the total number of livestock measured in terms of tropical livestock units owned by the head of the household. Livestock is a source of income through the sale of livestock and livestock products, as well as a source of supplementary food. Furthermore, livestock can use as a coping strategy in the event of crop failure or other disasters. Households with greater livestock holdings are to be more food secure than those without. The study result confirmed that the effect of livestock holdings on household food security was positive and statistically significant at one percent level of significance. The odds ratio in favor an increase in livestock ownership by one being food secure was increased by 1.063 on average citrus paribus. The result is consistent with the theory and most of the findings such as the study by Desta and Negussie (2017); Alemseged et al. (2018); Gebremariam et al. (2019); Hailu (2022) and Girma, et al. (2023).

Cultivated farm landing size: Cultivated land size was positively and significantly associated to food security status of the household at ten percent probability level of significance. The positive relationship implies that, the households with more opportunities cultivate land size is improve food production to access food. The implication is that, keeping other factor constant, as the cultivated land size increase by one hectare the probability of household being food secures increases by a factor of 0.618. This result supported by the findings of Adimasu et al. (2019); Gebremariam et al. (2019); Aweke et al. (2022); Girma et al (2023); Tamene and Ermias (2023).

3.5. Households Coping Strategies to Food Shortage

Coping strategies are mechanism that household's choice in order to attain food in the period of shock has occurred at household level by household itself rather than external body. Most coping mechanism based on the household's ability and constraints as well as the availability of opportunities. Accordingly, the survey result of coping strategies majority of household was used mechanisms such as reduced number of meals eaten in a day, selling small animals like hen, sheep, goats, relay on causal labor, borrow money/food from relatives and selling firewood and charcoal were used regularly and occasionally in the period of shortage of food.

Table 7: Coping strategies of households to food insecurity

Possible coping mechanisms	Regularly		Occasionally		Never	
	Frequency	%	Frequency	%	Frequency	%
Selling small animals (hen, sheep, goats)	33	12.27	140	52.04	96	35.69
Relay on causal labor	47	17.47	124	46.10	98	36.43
Reduced number of meals eaten in a day	36	13.38	156	57.99	77	28.62
Selling firewood and charcoal	54	20.07	64	23.79	151	56.13
Borrow money/food from relatives	42	15.61	81	30.11	146	54.28
Reducing the amount of food served	16	5.95	58	21.56	195	72.49
Selling large animals (ox, bull, cow, etc)	14	5.20	67	24.91	188	69.89
Rented out land to buy food	7	2.60	105	39.03	157	58.36
Sent children to stay with relatives	12	4.46	56	20.82	201	74.72
Rely on less expensive foods	32	11.90	68	25.28	169	62.83
Reduce spending on non-food items	13	4.83	57	21.19	199	73.98
Harvest immature crop	9	3.35	71	26.39	189	70.26
Household members migrate to work	20	7.43	44	16.36	205	76.21

Selling small animals: As the survey result of coping strategies, revealed household in the study area was selling small animals like hen, sheep, and goats in the period of shortage of food. About 12.27 percent of household was use this strategy regularly, while 52.04 percent was occasionally selling small animals in the period of shock has occurred at household level. On the other hand, 35.69 percent of household was never using this strategy. A study conducted by Arragaw and Argaw (2024) the sale of livestock are the dominant coping strategies reported by 84% of households.

Relay on causal labor: this strategy is another important copy strategy in the period of shock in the study area. As survey, result indicated 124 (46.10%) household occasionally of relay on causal labor and 47 (17.47%) regularly relay on causal labor in the period of shock. Working as daily laborers, borrowing money, migrating for seasonal work, and selling livestock or household assets are coping with food shortages; many households used strategies (Tadese and Yabsira, 2025).

Reduced number of meals eaten in a day: this strategy is most important mechanism in study area. The survey result state that majority (57.99%) of household occasionally use reduced number of meals eaten in a day mechanism in the period of shortage of food. Additionally, about 13.38 percent of household were used regularly. However, 77 (28.62%) household never reduced number of meals eaten in a day mechanism in the period of shortage of food. As a study conducted by Mojela et al. (2018), coping strategy index results showed that reduce number of meals eaten in a day used by 60% household in the period of shock.

Selling firewood and charcoal: selling charcoal is the other strategy that the households in the study area practice when there is food shortage in a household level. Form the total sampled households 151(56.13%) households responded that they never sell charcoal when there is food shortage in the household level. The result reveal that 64 (23.79%) households occasionally practice a strategy when the household face food shortage and 54 (20.07%) households regularly sell firewood and charcoal to cope with the problem of food shortage.

Borrow money/food from relatives: borrowing in other important mechanism in the period of food shortage. About 54.28 percent of sample household in the study was never borrowing money/food from relatives in the period of shock. On the other hand, 15.61 percent of household was borrowing money/food from relatives regularly; while 30.11 percent was occasionally, borrowing money/food from relatives in the period of shock has occurred at household level. The study conducted by Mesfin et al. (2021) community was coping food shortage by borrowing food to utilizing less preferred foods to cope up food shortage and starvation.

Reducing the amount of Food served: this mechanism was also another mechanism in the period of food crisis. A survey, result indicated that 58 (21.56%) household occasionally reduce the amount of food served and 16 (5.95%) regularly reduce the amount of food served in the period of food shortage. As a study by Gizachew et al. (2023) fewer-quality foods is most severe mechanism to improve during food shortage.

Selling large animals: large animals like ox, bull, cow, are also important factors to solve the problem of food crisis. Based on survey result, about 5.20 percent of household was use this strategy regularly, while 24.91 percent was occasionally selling large animals in the period of shock has occurred at household level. On the other hand, 69.89 percent of household was never using this strategy.

Rented out land to buy food: this mechanism was also another mechanism that rural household used in the period of food crisis. Renting land mechanism used occasionally by 39.03% household, while about 2.60% of household was used regularly at the period of food shortage in the study area.

Sent children to stay with relatives: sent children to stay with relatives is another copy strategy that rural household used in the period of shock in the study area. As survey, result indicated 56 (20.82%) household occasionally sent children to stay with relatives and 12 (4.46%) regularly sent children to stay with relatives in the period of shock. Some households also dropped children from school, or sent them to live with relatives (Tadese and Yabsira, 2025).

Rely on less expensive foods: rely on less expensive and less preferred food is other copy strategy that rural household used in the period of food shortage. About 62.83 percent of sample household in the study was never rely on less expensive and less preferred food in the period of shock. On the other hand, 11.90 percent of household was rely on less expensive and less preferred food regularly; while 25.28 percent was occasionally, rely on less expensive and less preferred food in the period of shock has occurred at household level. As a study conducted by Mojela et al., (2018) rely on less expensive and preferred food has been used by 86% of the population to coping food crisis.

Reduce spending on non-food items: As the survey result of coping strategies, revealed household in the study area was, reduce spending on non-food items in the period of shortage of food. About 4.83 percent of household was use this strategy regularly, while 21.19 percent was occasionally reduce spending on non-food items in the period of shock has occurred at household level. On the other hand, 73.98 percent of household was never using this strategy.

Harvest immature crop: harvesting immature crop is one of the mechanism that practiced by the households in the study area during food shortage in the household level. From the total sampled households, 189(70.26%) households responded that they never harvest immature crops, 71(26.39%) households occasionally harvest immature crop during food shortage in the household level and the remaining 9(3.35%) households regularly harvest immature crop to overcome the problem of food crises.

Household members migrate to work: Household members migrate to work find income to improve food crisis. About 76.21 percent of sample household in the study were never household members migrate to work in the period of shock. On the other hand, 7.43 percent of household was household members migrate to work regularly; while 16.36 percent was occasionally, household members migrate to work in the period of shock has occurred at household level. As a studies conducted by Tadese and Yabsira (2025); Gizachew et al. (2023), the most severe of coping strategy is migrating and begging for food.

4. CONCLUSION AND RECOMMENDATION

4.1. Conclusion

Food security remains an issue in Ethiopia particularly in the rural households. It is one of the greatest challenges for today's population and future generations. Hence, the main objective of the study was to examine the determinants of rural households' food security and their coping strategies in Baka Dawolla Ari district, Ari Zone, Southern Ethiopia regional state. The descriptive analysis of the study revealed that about 57.62 percent and 42.38 percent of the sample households were food insecure and food secure, respectively. Description result revealed that male and female household headed in the food secure sample households were 89.47 and 10.53 percent, respectively, while about 89.68 percent and 10.32 percent of the food insecure households were male and female, respectively.

On the other hand, average proportion of education level of household heads that can be food secure was larger than the proportion of educational level food insecure household heads. From all respondents of household heads, about 37.92 percent of them had access to credit, whereas 62.08 percent from those who had no access to credit service. About 38.92% household heads food secured had access to credit where as about 37.42 percent were access to credit from household heads food in secured and there is no statistical difference in age between the two groups.

Additionally, descriptive analysis of the study revealed that farmers who have larger cultivated land size are more likely to be food secure than those who cultivated smaller land size due to the fact that there is high possibility to produce more food. In addition, about 61.40% of the food secured and 29.68% of food insecure households had access to extension services. This implies those households that accesses to extension services are greatest probability to be food secure. According to the survey results of the study from the sample households' food security 44.74 percent of the household members participate on off-farm activity where, only about 30.32 percent of household's food insecure were participate and about 69.68 percent did not involve in off-farm activities.

Moreover, oxen ownership is also an important variable in the study areas that almost entirely rely on traditional farming methods, there by significantly affecting household's crop production. Oxen the sole provider of draft power and determinant of on time land preparation, is not uniformly distributed between food secure and insecure households. The average number of oxen owned by food secure households is larger than the food insecure. Additionally, livestock play an imperative role to the households' food security through increasing purchasing power of household to buy food from market. Number of livestock (excluding oxen) was on average food secure households have 22.28 livestock while food insecure households have 13.68 livestock.

As the survey of coping strategies result revealed that occasionally majority of household use reduced number of meals eaten in a day (57.99%) mechanism in the period of shortage of food. Additionally, the second and third copy mechanisms used by sampled household in study area are selling small animals like hen, sheep, goats and relay on causal labor by 52.04% and 46.10%, respectively. In addition, copy mechanisms used occasionally by sampled household in study area are selling firewood and charcoal (23.79%); borrow money/food from relatives (30.11%), reducing the amount of food served (21.56%), and selling large animals (24.91%) are important copy mechanisms. Moreover, rented out land to buy food (39.03%), sent children to stay with relatives (20.82%), rely on less expensive foods (25.28%), reduce spending on non-food items (21.19%), harvest immature crop (26.39%) and household members migrate to work (16.36%) are copy strategy used occasionally by household in the study area.

Additionally, regularly sample household used mechanism such as selling small animals (12.27%), relay on causal labor (17.47%), reduced number of meals eaten in a day (13.38%), selling firewood and charcoal (20.07%); borrow money/food from relatives (15.61%), reducing the amount of food served (5.95%), and selling large animals (5.20%) are important copy mechanisms. Moreover, rented out land to buy food (2.60%), sent children to stay with relatives (4.46%), rely on less expensive foods (11.90%), reduce spending on non-food items (4.83%), harvest immature crop (3.35%) and household members migrate to work (7.43%) are copy strategy used regularly by household in the study area.

The results of logistic regression analysis indicated that age household head, education level, access to extension services, households participate in off farm activities, households' ownership of oxen, livestock ownership and cultivable land size were positive and significantly influencing household food security in the study area. On the other hand, family size is negative and significantly influencing household food security in the study area.

4.2. Recommendations

Based on the findings of the study, the following recommendations are forwarded to improve household food security in study area. Livestock holding is important factor affecting food security positively because livestock is an important source of wealth that could contribute to food security in the study area. Therefore, the livestock sector should strengthen through rearing, establishment of better management, veterinary services and improve livestock markets.

As household education level is positively affecting food security, in order to bring food security at the household level, the woreda development strategies need to encompass education programs to the smallholders.

Additionally, an extension service is boost food security of household. Therefore, short-term trainings that attempted in farmers training center should experience in a strategic and organized way whenever necessary.

Cultivated land was to be directly related and positively to food security of households in the study area. As a result, emerging farming training to reduce land fragmentation and improve production. Therefore, the possible measures that can undertake to achieve this strategy include improving the quality of the land through improved soil and nutrient management, promotion of labor-intensive technologies, using improved seed and creation of labor-intensive rural employment opportunities in the short-to-intermediate terms.

When off farm activities is an important factor of household food security, as a result, give training on off-farm activities and facilitate starting income to improve rural households get extra income to supplement their livelihoods.

REFERENCES

Abera, N., and Mekonin, S. (2022). The status and determinants of rural household food insecurity in North Shewa Zone, Oromia Region, Ethiopia. Turkish Journal of Food and Agriculture Sciences, 4(1), 6-12. Doi:10.53663/turjfas.1020187

Abi, M., and Tolossa, D. (2015). Household food security status and ıts determinants in Girar Jarso Woreda, North Shewa Zone of Oromia Region, Ethiopia. Journal of Sustainable Development in Africa, 17(7), 118-137.

Adamu, T., Y., and Tsehay, Y., A. (2025). Food security among urban households: status, determinant factors and coping strategies evidence from Chiro Town, West Hararghe Zone, and Oromia Region, Ethiopia. Frontiers, 5(3), 77-93.https://doi.org/10.11648/j.frontiers.20250503.11

Agidew, A,. A., and Singh, K., N. (2018). Determinants of food insecurity in the rural farm households in South Wollo Zone ofEthiopia: the case of the Teleyayen subwatershed. Agidew and Singh Agricultural and Food Economics, 6 (10), 1-23. https://doi.org/10.1186/s40100-018-0106-4

Alemayehu, A. (2025). Determinants of Ethiopia export performances: evidence from an auto regressive distributive lag model. International Journal of Oil, Gas and Coal Engineering. 13(5), 84-93. https://doi.org/10.11648/j.ogce.20251305.12.

Alemayehu, A., and Tesfaye, A. (2024). Food security determinants and coping strategies in central highlands of Ethiopia. Wiley Advances in Agriculture, 5(1), 1-13, https://doi.org/10.1155/2024/8361245

Amare, A., and Simane, B. (2017). Assessment of household food security in the face of climate change and variability in the upper Blue-Nile of Ethiopia. Journal of Agricultural Science and Technology, 7(4), 285-300. 10.17265/2161-6264/2017.04.006

Ari Zone Agriculture and Rural Development Office. Agricultural Report. 2023; Accessed on October 30, 2023.

Asefawu, G., S. (2022). Seasonal migration and household food security status in the drought-prone areas of Northeast Ethiopia. Environmental Challenges, 8(1), 1-13. 100566. https:// doi. org/ 10. 1016/j. envc. 2022. 100566.

Awoke, W., Eniyew, K., Agitew, G., and Meseret, B. (2022). Determinants of food security status of household in Central and North Gondar Zone, Ethiopia, Cogent Social Sciences, 8(1), 2040138, 10.1080/23311886.2022.2040138

Aysheshim, A., Yayehand, D., and Mulugeta, M. (2022). Determinants of Household Food Security in the Benishangul-Gumuz Region, Western Ethiopia. Ethiopian Journal of Developmetn Research, 44 (2), 55-75. doi: 10.3329/sja.v14i2.31257

Dawit, D., and Zeray, N. (2017). Determinants of rural household food security in Wolaita Zone: The Case of Humbo Woreda. Journal of Poverty, Investment and Development, 32(1), 65-82. 10.11648/j.ijae.20190406.16

Dube, A., K., Haji, J., and Zemedu, L. (2018). Determinants of food insecurity and coping strategies of rural households: The case of Shalla District, West Arsi Zone, Oromia Region, Ethiopia. Journal of Development and Agricultural Economics, 10(6), 200-212. 10.5897/JDAE2018.0933

FAO (2014). The state of food insecurity in the world 2014. Strengthening the enabling environment for food security and nutrition. Food and Agriculture Organization, Rome. Available online at http://www.fao.org/3/a-i4030e.pdf

Fikire, A,. H., and Zegeye, M., B. (2022). Determinants of rural household food security status in North Shewa Zone, Amhara Region, Ethiopia. Scientific World Journal, 11(1), 1-8. https://doi.org/10.1155/2022/9561063

Fikre, T. (2022). Determinants of rural household food ınsecurity and local coping strategies: the case of Hoko District in Sidama Region, Ethiopia. International Journal of Novel Research in Life Sciences, 9(2), 51-58.

Gebremariam M., Ayalew Z., Gebremariam B. (2019). Household food ınsecurity and coping strategies: a perspective from Lasta Wereda, North Wollo Zone of Amhara, Ethiopia. Journal of Scientific and Innovative Research, 8(2), 59-63.

Gerezgiher, A., Alamirew, B. and Zaid, N. (2018). Determinants of rural household food ınsecurity ın Northern Ethiopia: an empirical study.

Journal of Sustainable Development in Africa, 20(1), 32-69.

Gujarati D. N. (2013). Basic Econometrics, 4th ed., McGraw-Hill Co.

Gujirati, D. N. (1995). Basic Econometrics, 3rd ed.. New York: McGraw-Hill Co.

Guyalo, A., K. (2025) Food security and its determinants among rural households in Gambella region, Ethiopia. Cogent Economics & Finance, 13 (1), 2484651, 10.1080/23322039.2025.2484651

Haılu, H., and Tolossa, D. (2022) Impacts of Adoptıng ımproved wheat varıetıes on household food securıty ın Gırar Jarso Dıstrıct, Ethiopia. Review of Agricultural and Applied Economics, 25(1), 73-86. 10.15414/raae.2022.25.01.73-86

Hailu MK. (2022). Assessment of household food security and its determinants in Ethiopia. Journal of Agricultural Research Advances, 4(4), 20-26.

Igene, C. A., Belonwu, N. E., Elusaiwe, M., and Amah J.U. (2025). Coping strategies of food security: a perspective of female-headed households in Esan West Local Government Area. Journal of Arid Agriculture, 26(1), 103 – 113. https://doi.org/10.63659/jaa.v26i1.72

Kassahun, B.Y. and Hamer C.H. (2020). Effects of climate variability in South Ari Woreda, South Omo Zone, Ethiopia. International Journal of Scientific Research in Multidisciplinary Studies, 6(12) 42-52.

Ketema, S., & Megento, T. (2021). Rural households' food security status in arsi zone, Oromia National Regional State, Ethiopia. Journal of Sustainable Rural Development, 5(2), 189-204. https://dorl.net/dor/20.1001.1.25383876.2021.5.2.2.8

Mazenda, A., Manzi. P., Mushayanyama. T. and Ngarava, S. (2022). Household level determinants of food security in the City of Tshwane, South Africa. Food Research, 6(6), 184-192. https://doi.org/10.26656

Mebrie, T., and Ashagrie, E. (2023) Rural household food security status and its determinants in Libokemkem woreda of the Amhara region, North Western Ethiopia. Journal of Poverty and Social Justice, 31(2), 252–275, DOI: 10.1332/175982721X16760300561102

Melese, M., Tilahun, M., and Alemu, M., (2021). Household food insecurity and coping strategies in Southern Ethiopia. Agriculture & Food Security, 10(23), 1-12. https://doi.org/10.1186/s40066-021-00296-8

Moges, A., and Melkamu, M. (2020). Determinants of rural household food ınsecurity status and coping strategies in case of Konso Woreda, Segen Area People's Zone of Southern Ethiopia. International Journal of Business and Economics Research, 9(4), 202-206. 10.11648/j.ijber.20200904.16

Mohamed, A. A. (2017). Food security situation in Ethiopia: a review study. International Journal of Health Economics and Policy, 2(3), 86-96. 10.11648/j.hep.20170203.11.

Mojela T, Hlongwane J, and Ledwaba L (2018). Household food security and coping strategies: a case study of Tembisa Township of Ekurhuleni Municipality, Gauteng Province, South Africa. The World's Largest Open Access Agricultural & Applied Economics Digital Library, 58-77. http://ageconsearch.umn.edu

Mota, A., A., Lachore, S., T., and Handiso, Y., H. (2019). Assessment of food insecurity and its determinants in the rural households in Damot Gale Woreda, Wolaita zone, southern Ethiopia. Agriculture and Food Security, 8(11), 2-11. https://doi.org/10.1186/s40066-019-0254-0

Mukaila, R., Falola, A., and Omotesho, O., A. (2020). Food security status: ıts drivers and coping strategies among vegetable farming households. Cercetări Agronomice în Moldova, 53(4), 414-425. 10.46909/cerce-2020-035

Mumed, A., A., Bonso, A., B,. and Habib, A., H. (2022). Review on food in security status and ıts major coping strategies in rural Ethiopia. International Journal of Health Economics and Policy, 8(1), 10-15. 10.11648/j.hep.20230801.12

Oyato, M., T., Borano, G., B., Tariku, K., A., & Zekariyas, E., H. (2024). Household food insecurity predictors and coping strategies in Basketo Special District, Southern Ethiopia: using the double hurdle model. Cogent Social Sciences, 10(1),1-14, 10.1080/23311886.2024.2362287

Sani, S., and Kemaw, K. (2019): Analysis of rural households food security in Western Ethiopia. Food and Nutrition Sciences, 10, 249-265. https://doi.org/10.4236/fns.2019.103019

Sayekti, W., D., Zakaria, W., A., Syafani, T., Y., and Mutolib, A. (2022). Dominant factors on food coping mechanism of poor households in Pringsewu Regency, Indonesia. Mal J Nutr, 28(3), 441-452, https://doi.org/10.31246/mjn-2020-0099

Shumiye, A., Tolossa, D., and Tsehay, S. (2026). Food security determinants and coping strategies among rural households in Ada'a District, Central Ethiopia. Food Science & Nutrition, 14, e71588, 1-25. https://doi.org/10.1002/fsn3.71588

Toma TM, Andargie KT, Alula RA, Kebede BM, Gujo MM, (2023) Household food insecurity and associated factors in South Ari district, Southern Ethiopia: A community-based crosssectional study. PLoS ONE, 18(4), 1-15. https://doi.org/10.1371/journal.pone.0284252

Woleba, G., Tadiwos, T., Bojago, E., and Senapathy, M. (2023). Household food security, determinants and coping strategies among small-scale farmers in Kedida Gamela district, Southern Ethiopia. Journal of Agriculture and Food Research 12(1), 1-15. https://doi.org/10.1016/j.jafr.2023.100597

Yenealem, A., Tolossa, D., and Abate, E. (2023). Determinants of food security among households organized and nonorganized ın self-help groups ın Debre-Markos City and Gozamin District of Northwestern Ethiopia. Journal of Sustainable Development in Africa, 25(1), 130-152.

Yenesew, E., and Masresha, D. (2019). Impact of livelihood diversification on rural household food security: evidence from Goncha-Siso

Enesie Destrict of Amhara Regional State, Ethiopia. International Journal of Agricultural Economics. 4(6), 288-297.

Yenewa, W. (2024). Review on determinants and challenges of household food security in Ethiopia. Middle East Research Journal of Economics and Management, 4(1), 8-14. 10.36348/merjem.2024.v04i01.002

Yohannes, G., Wolka, E., Bati, T., and Yohannes, T. (2023). Household food insecurity and coping strategies among rural households in Kedida Gamela District, Kembata-Tembaro zone, Southern Ethiopia: mixed-methods concurrent triangulation design. BMC Nutrition, 9(4), 1-13. https://doi.org/10.1186/s40795-022-00663-z

PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

DYNAMIC NEXUS BETWEEN FINANCIAL DEEPENING AND INCLUSIVE GROWTH IN NIGERIA: ARDL APPROACH

DOI: 10.17261/Pressacademia.2026.2037
JBEF- V.15-ISS.1-2026(9)-p.100-110

Raymond Rahaj Adegboyega
British American University, Department of Management, School of Business, Florida. USA.
dgbyga@gmail.com , ORCID: 0000-0002-8067-3410

Date Received: March 3, 2026 **Date Accepted:** June 6, 2026

To cite this document
Adegboyega, R.R., (2026). Dynamic nexus between financial deepening and inclusive growth in Nigeria: ARDL approach. Journal of Business, Economics and Finance (JBEF), 15(1), 100-110.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2037

ABSTRACT
Purpose- Within the context of financing-growth nexus, there is contentious argument that financial deepening is the consequence, not a cause of economic growth which tends to increase the demand of financial instruments that leads to the advancements in financial infrastructure. Consequently, to resolve this issue, there is need to examine the impact of financial deepening on inclusive growth in Nigeria.
Methodology- Using ARDL method of analysis on data collected from Central Bank of Nigeria and World Bank database from 1982 to 2024 the study determined the short and long run relationships between financial deepening and inclusive growth in Nigeria.
Findings-The results of the study show that credits to private sector (PSC_GDP), bank lending rate (LR) and rural bank loan to rural deposit ratio (RL_RD) have positive and significant impact on inclusive growth. While financial deepening (FD) has negative but significant impact on inclusive growth. As for the loan deposit ratio (LDR), it has a negative and insignificant impact.
Conclusion- The findings from this study provide new and valid evidence that addressed the controversy between the finance-growth nexus. Also, aligning the results with the theoretical expectations provide the basis for a sound financial system policy that emphasizes strong financial deepening in Nigeria, thereby, making the report a reliable basis for forecasting and policymaking. A limitation of this study is that the available data is restricted to 2023, which presents an opportunity for future research.

Keywords: Cash-based economy, financial deepening, financial intermediation, financial system, inclusive growth
JEL Codes: G00, G20, O40

1. INTRODUCTION

The existence of a well-functioning financial system is highly essential for its critical roles in capital formation and financial intermediation, which facilitate economic growth (Mordi, 2010 & Puatwoe and Piabuo, 2017). This is affirmed by Aslam and Saeed (2023) and Gatsi *et al.* (2020) that financial inclusion, a key component of financial deepening is widely identified as a catalyst for promoting inclusive growth within the BRICS nations.

Acemoglu and Autor (2011) and Aizenman *et al.* (2012) argument is that financial deepening is one of the obstacles hindering the achievement of inclusive growth in developing countries. While, a well-functioning financial structure enhances overall economic efficiency, by creating and transferring financial resources from traditional sectors to the more modern, growth inducing sectors of the economy (Rojan and Zingales, 2003; Akpokerere and Edafiaje, 2016 & Gültekin and Umutlu, 2023). This is corroborated by Sahay *et al.* (2015) that deep and liquid financial system with different types of financial instruments have a tendency to absorb more shocks than a shallow one.

The National Financial Inclusion Strategy Report (CBN, 2012) in Nigeria emphasized that the goal of inclusive growth can only be accomplished through a deepening financial system that makes financial services easily accessible with less stringent conditions attached in accessing loans for investments (Abdul and Adamu, 2016 & Nwolisa and Cyril, 2019). This inclusive growth must be broad-based, creating productive and sustainable economic opportunities for all (Bakker and Messerli, 2017 & Djokoto, 2022).

Based on this concept, the African Development Bank did place inclusive growth and the transition to green growth at the center of its new Ten-Year Strategy (2013-2022) (AfDB, 2012). The strategy underpins the Bank's emphasis on strengthening the robustness, sustainability and inclusiveness of growth on the continent in a time of rapid change in the financial system, with financial deepening as one of the strategies being used.

Obviously, in both developed and developing countries financial deepening offers a myriad of opportunities in terms of improvement in economic conditions through increased competitive efficiency within financial markets, which thereby expand the depth of financial structure that boost economic growth (Levine and Loayza, 2000 and Balago, 2014).

In Nigeria, these positive outcomes in other countries have not yielded the desired results due to the country highly cash-based economy and some structural challenges. Even with various economic reforms of government there was little confidence in the financial services sector (Okafor and Nwosu, 2018). Despite extensive reforms, Nigeria's financial system remains distorted and current practices contribute to a shallow systems that hinders inclusive growth (Appiah, Li and Frowne, 2020).

From the foregoing, previous studies have focused on the direction of causality between financial deepening and economic growth, which has remained a contentious and unresolved issue. So, the present study departs from these studies by examined the impact of financial deepening on inclusive growth in Nigeria.

The remaining of the paper is structured as follows: section two is literature review, section three presents methodology, section four contain results and findings while section five is conclusion and recommendations.

2. LITERATURE REVIEW

2.1. Theoretical Review

The central theory is based on the assumption of Schumpeter (1934), Goldsmith (1969) and King and Levine (1993) that finance is a vital element of economic growth, which has to be inclusive and encompasses effective participation of the poor. McKinnon (1973) and Shaw (1973) argument is that liberalization of the financial system enhance the rate of economic expansion. In their financial repression, it is assumed that some financial reforms policies were creating unnecessary distortions in the financial market and results in shallow financial market. According to Maxwell (1989) corollary, it is presumed that Mckinnon and Shaw (1973) postulation became the theoretical basis which many developing countries built their policy decisions and financial reforms around. The main motive is to improve capital mobilization and efficiency of financial deepening (intermediation), consequently influencing investment and thereby economic growth with pro-poor oriented (Bencivenga and Smith, 1991; King and Levine, 1993 and Chukwu and Agu, 2009). Schumpeter (1911) assumption is that well-functioning financial intermediaries, particularly banks promote economic development by providing credit to innovative entrepreneurs. These lend credence to the two main diverging theories: the supply leading hypothesis and demand leading hypothesis. Schumpeter (1911) argument for supply-leading hypothesis is that financial development causes economic growth, a view that was elaborated upon by Goldsmith (1969); Calderon and Liu (2003) and Balago (2014), who found evidence for this relationship in developing countries. Whereas, demand side hypothesis is aligned with the Keynesian view of financial deepening, which also in line with Gurley and Shaw (1967), Omotor (2007) and Ndlovu (2013) argument that causality runs from economic growth towards financial deepening. Patrick (1966) emphasized further that demand following approach is related with the demand side of the financial system. That is, economic growth creates additional and new demand for financial services, which causes development in financial system. Apergis, Filippidis and Economidou (2007) neutral hypothesis assertion is that there is no relationship between financial development and economic growth, which is peculiar to developing countries.

2.2. Conceptual Review

The World Bank (1989) and Onyemachi (2012) defined financial deepening as an effort aimed at developing the financial system that encourages an increased in financial assets in the financial markets, leading to the expansion of the real sector of the economy. Hammilton and Godwin (2013); Osinsanwo (2013) and Ngerebo and Lucky (2016) defined financial deepening as a channel of increasing the supply of financial assets in the economy.

Akhator and Marcus (2018); Kiprop (2013); Kolawole *et al*. (2019) and Efanga, et al. (2020) perceived financial deepening as the channel of increasing the provision of financial services with a wider choice of services for the development of the society.

Using relative pro-poor approach, Dollar and Kraay (2002) and Hosono (2022) explained that inclusive growth occurs when the income of poor people increases comparatively faster than the average income of the population. Ravallion and Chen (2003) defined inclusive growth from an absolute pro-poor growth perspective, as the growth that affords the poor individuals benefit unconditionally. Ali and Son (2007) described inclusive growth as the growth that enables equal access to opportunities created by both the rich and poor. UNDP (2017) refers to inclusive growth as equity with growth or shared prosperity from economic growth. Furthermore, Anand *et al.* (2013) explained that inclusive growth occurs when there is an increase in average income through growth or/and an increase in income equality. Mitra and Das (2018) described inclusive growth as broad-based equitable growth, pro-poor growth and financially and environmentally growth. OECD (2015, 2018) and Withers (2018) described inclusive growth as the growth that created equal opportunities for all with reduction in poverty and easy accessibility to social services. Inclusive growth as defined by Barnat et al. (2023) implies an equitable allocation of resources or providing equitable opportunities to all in accessing resources such that it benefits the society at large.

2.3. Empirical Evidences

This study deviated from previous studies by examined the impact of financial deepening on inclusive growth. Kamat and Kamat (2007) results show that the short run effect of financial development causing economic growth. Khan (2008) and Safdar (2014) results revealed that financial deepening impact positively on economic growth in Pakistan. Nzotta and Okereke (2009) study shows that financial deepening index is low in Nigeria and financial system has not sustained an effective intermediation. Pradhan (2010) and Giri and Mohapatra (2012) found that financial deepening promotes growth in India. Iyoboyi (2013) using ARDL technique found a bi-directional causal relationship between financial deepening and economic growth. Wycliffe, *et al.* (2013) found a positive relationship between financial deepening and economic growth in Kenya. Using ARDL method of analysis, Ghildiyal *et al.* (2015) findings revealed a long-run relationship between financial deepening and economic development in India. Agheli and Hadian (2017) findings show that shallowness of financial deepening has no impact on economic growth in the fifteen selected emerging and Middle Eastern countries investigated. Karimo and Ogbonna (2017) results showed that the growth-financial deepening nexus in Nigeria follows the supply-leading hypothesis. Gezer (2018) findings indicate that some countries can be clustered according to supply-leading and demand following approach but bi-directional causality exists for some countries. Igwebuike *et al.* (2019) results show that credit to private sector to GDP ratio has positive effect on economic growth in Nigeria. Nwosu *et al.* (2021) findings show a positive relationship between financial deepening and economic growth in Yemeni. Okafor and Ude (2022) using cointegration approach found that all financial deepening factors have positive impact on economic growth in Nigeria. Al-Shawesh and Kumar (2022) using ARDL method of analysis found that financial deepening impacted economic growth. Also, the results of Eniekezimene and Chiazor (2023) study using ARDL technique revealed that there is a positive impact ofmarket capitalization to GDP on economic growth in Nigeria. Okeke and Akunna (2023) findings confirmed that money supply to GDP and Market capitalization to GDP have positive effect on economic growth in Nigeria. Shan and Liu (2023) results affirmed that financial deepening significantly improves the level of digital economy development in China. Ekane *et al.* (2024) results revealed that financial deepening have negative effect on economic growth in Nigeria. Manasseh, Ngong *et al*. (2024) results show that bidirectional causality exists between financial deepening and economic growth in emerging economies in Africa. Finally, Amakiri and Bobai (2025) findings revealed that there is a positive relationship between financial deepening and economic growth and statistically significant in Nigeria.

This study deviated from the previous studies by examined the impact of financial deepening on inclusive growth, using these variables: BLR (bank lending rate), LDR (loan deposit ratio) and RL_RD (rural loan to rural deposit ratio) which were not examined by any of the previous studies. This study filled these gaps that formed part of our contributions to knowledge.

3. METHODOLOGY

The study made use of the data collected from Central Bank of Nigeria statistical bulletin and World Bank database between 1982 and 2024 using ARDL method of analysis. The variables of the study include: inclusive growth proxy by GDP growth rate (GDPGR), while independent variables are private sector credit to GDP ratio (PSC_GDP), financial deepening ratio (FD), prime lending rate (LR), loan deposit ratio (LDR) and rural loan to rural deposit ratio (RL_RD).

3.1. Model Specification

Based on our theoretical review in this study, we follow Schumpeterian finance endogenous growth model developed by Aghion *et al.* (2005) & Acemoglu *et al.* (2006) which is built around the aggregate production function given as:

$$x = \gamma * \delta * q \tag{1}$$

In equation (1) above, technological progress (x) is defined as a function of research and development (R&D) (q), while the two parameters define the probability that each unit spent on R&D yields a successful innovation (γ) and the extent to which each innovation raises the productivity parameter (δ), respectively. The economic determinants of the R&D are assumed to be taken as exogenous by the entrepreneur. Thus, these may include; the discounted value of expected returns, the real interest rate, capital per efficiency unit, and institution features of the economy.

$$q = q\ \{\gamma, \delta, r, comp, ppr, \varepsilon\} \tag{2}$$

From the equation (2) above; the R&D intensity (q) is assumed to be positively related to the discounted value of expected return as measured by γ and δ, negatively related to real interest rate (r), and positively related to capital per efficiency unit (k), while product market competition (comp.) and property right (*ppr*) are examples of institutional features within the economy. $\mathcal{E}$ depicts all other institutional features of the economy not cited in the equation. From equations 1 and 2, the "Schumpeter finance-growth relationship" can be derived as:

$$x = x\{k\} \tag{3}$$

This states that since the rate of technology (x) depends on q, which in turn, depends on k, x is a function of k, the capital efficiency per unit. A positive relationship also exists between the two variables. Thus, an increase in the saving rate in the economy will increase the capital efficiency per unit, which in turn stimulates more R&D activities via innovation. This will bring about growth in the economy. Thus, in a steady state, *x* is similar to economic growth, *Yt* below.

3.2 Analytical Model of the study

Following a detailed review of previous studies and improving upon the theoretical postulate described in equation (3) above and also in line with the endogenous growth model of Bencivenga and Smith (1991) which assumes that financial deepening (intermediaries) are the channels of capital formation which promotes growth, therefore economic growth (*Yt*) is expressed as a function of financial intermediation, *FIt*, and a set of control variables, Z.. The adopted production function model in equation (3) above can be rewritten and specified in line with the major variables of the study as follows:

$$Yt= f \{FIt, Zt\} \tag{4}$$

Following the empirical specifications in Nwosu *et al.* (2021) and Al-Shawesh and Kumar (2022), the equation (4) above is expanded to accommodate other indicators of financial intermediation (*FIt*), as well as control variables (*Zt*) which are determinants of traditional growth. Thus, in line with our study, the model is stated as follows:

$$GDPGR = PSC\_GDP, FD, LDR, LR, RL\_RD \tag{5}$$

Therefore, following the adopted modified models of Nwosu *et al*. (2021) and Al-Shawesh and Kumar (2022) methods of analysis that used a time subscript (t) and first difference operator (Δ), we therefore model the relationship between financial intermediation and economic growth as follows:

$$ln\Delta GDPGR_t = f(ln\Delta PSC\_GDP_t, ln\Delta FD_t, ln\Delta LDR_t \, ln\Delta LR_t \, ln\Delta RL\_RD_t) \tag{6}$$

In order to empirically test the long-run relationship between financial intermediation and economic growth the transformation of equation (6) into a linear equation then becomes:

$$ln\Delta aGDPGR_t = \alpha + \psi \, ln\Delta PSC\_GDP_t, + \gamma \, ln\Delta FD_t, + \varphi \, ln\Delta LDR_t + \phi ln\Delta LR_t + \text{Ɑ} \, ln\Delta RL\_RD_t \tag{7}$$

where, *ln* is the natural logarithm of the variables, and the estimates of ψ, γ, φ, ϕ and Ɑ represent elasticities. The error term ε_t is assumed to be white noise normally and identically distributed. The reasons for using ARDL technique are the following: it has advantage of not requiring a specific identification of the order of the underlying data because it allows a mixture of I(1) and I(0) variables as regressors, that is, the order of integration of appropriate variables may not necessarily be the same. Also, it circumvents the low power of unit root tests and the resulting degree of uncertainty regarding the order of integration of the underlying variables. Additionally, it is also suitable for small or finite sample size (Pesaran *et al*., 2001).

In order to conduct the bounds test, equation (7) is converted into an unrestricted error correction model (UECM) form:

$$\begin{aligned} ln\, \Delta GDPGR_t = {} & \alpha + \textstyle\sum_{k=1}^{n} \delta_1 ln\Delta GDPGR_{t-k} + \sum_{k=0}^{n} \delta_2 ln\Delta PSC\_GDP_{t-k} \\ & + \textstyle\sum_{k=0}^{n} \delta_3 ln\Delta FD_{t-k} + \sum_{k=0}^{n} \delta_4 ln\Delta LDR_{t-k} + \sum_{k=0}^{n} \delta_5 ln\Delta LR_{t-k} \\ & + \textstyle\sum_{k=0}^{n} \delta_6 ln\, \Delta RL\_RD_{t-k} + \psi \, ln\, PSC\_GDP_{t-1} + \gamma \, ln\, FD_{t-1} \\ & + \varphi \, ln\, COBS\_M2_{t-1} + \phi lnLR_{\,t-1} + \text{Ɑ} \, ln\, SR_{t-1} + \varepsilon_t \end{aligned} \tag{8}$$

where, **α** is the drift component, **Δ** represents the first difference operator, and ε_t are white noise errors. In this setup, the short-run effects are inferred by the sign and significance of the estimates of $\boldsymbol{\delta_1}$, $\boldsymbol{\delta_2}$, $\boldsymbol{\delta_3}$, $\boldsymbol{\delta_4}$ and $\boldsymbol{\delta_5}$ while the long-run effects are inferred by the sign and significance of the estimates of $\boldsymbol{\psi}$, $\boldsymbol{\gamma}$, φ, ϕ and Ɑ**.** Equation (8) indicates that economic growth tends to be influenced and explained by its past values. The structural lags are established by using minimum Akaike's information criteria (AIC). The Wald test (F-statistic) was also computed to differentiate the long-run relationship between the concerned variables.

Since all the variables in the model appear to be trended, a second ARDL-UECM including a trend term **(ξt)** is presented in the form:

$$\begin{aligned} ln\, \Delta GDPGR_t = {} & \alpha + \xi_t + \textstyle\sum_{k=1}^{n} \delta_1 ln\Delta GDPGR_{t-k} + \sum_{k=0}^{n} \delta_2 ln\Delta PSC\_GDP_{t-k} \\ & + \textstyle\sum_{k=0}^{n} \delta_3 ln\Delta FD_{t-k} + \sum_{k=0}^{n} \delta_4 ln\Delta \, LDR_{t-k} + \sum_{k=0}^{n} \delta_5 ln\Delta LR_{\,t-k} \\ & + \textstyle\sum_{k=0}^{n} \delta_6 ln\, \Delta RD\_RL_{t-k} + \psi \, ln\, PSC\_GDP_{t-1} + \gamma \, ln\, FD_{t-1} \\ & + \varphi \, ln\, LDR_{t-1} \, \phi lnLR_{\,t-1} + \text{Ɑ} \, ln\, RL\_RD_{t-1} + \xi_t \end{aligned} \tag{9}$$

In this case, the null hypothesis of no cointegration, that is, no long run relationship ($H_0 = \psi = \boldsymbol{\gamma} = \varphi = \phi =$ Ɑ $= 0$) is tested against the alternative of long run relationship (H_1: $\psi \neq \boldsymbol{\gamma} \neq \varphi \neq \phi \neq$ Ɑ $\neq 0$) using the familiar F-test with critical values tabulated by Pesaran *et al*, (2001). Accordingly, it is hypothesized that the estimates of ψ, $\boldsymbol{\gamma}$, φ, ϕ and Ɑ are positive and statistically significant, thus confirming the diversification-led growth hypothesis.

4. FINDINGS AND DISCUSSION

4.1. Descriptive Statistics

In the table 1 below, the values of mean and median which are within the maximum and minimum values of the series reflected high level of consistency of all the series. There is a positive trend in all the variables because they have positive mean values. The extra ordinary highest maximum value and high standard deviation of RL_RD call for concern, which implies that at a particular period of the study more rural loans were granted than rural deposits. It is a manifestation of shallow financial deepening of the rural areas in Nigeria, which has policy implication. Also, the increase in rural loans that did not reflect in the rural development shows that these loans were granted to business activities in the urban areas. The financial deepening is very shallow in Nigeria taking into consideration the least value of the maximum value of private sector credit to GDP (CPS_GDP) comparable to other variables' maximum values. The high standard deviation of lending rate (LR) is also high and it is an evident of high lending rate during the period of the study. The positive Kurtosis indicates too few cases at the tail of the distribution. Also, all variables had their entire kurtosis coefficient >0 which shows that they are leptokurtic. The Jarque-Bera confirms normal distributions across all datasets except GDP_GR and RL_RD. The Skewness coefficient indicates normal curves for all the variables with the values ranging between -3 and +3.

Table 1: Descriptive Statistics Results

	GDPGR	CPS_GDP	LR	FD	LDR	RL_RD
Mean	3.229442	9.548837	48.37835	15.90116	67.25387	97804.06
Median	3.300000	8.200000	46.80000	13.02000	66.90000	56.30000
Maximum	33.70000	19.60000	70.40000	27.56000	96.81702	3699600.
Minimum	-13.10000	4.900000	29.10000	8.460000	37.55947	21.30000
Std. Dev.	7.020395	3.687648	10.44313	5.684939	13.22592	563740.1
Skewness	1.398879	0.974035	0.249344	0.492069	-0.162645	6.278607
Kurtosis	10.17875	3.368656	2.434974	1.636003	2.685168	40.62001
Jarque-Bera	106.3566	7.042833	1.017566	5.068652	0.367172	2818.199
Probability	0.000000	0.029558	0.601227	0.079315	0.832280	0.000000
Observations	43	43	43	43	43	43

4.2. Correlation Matrix Tests

The results in table 2 below indicate that there is a weak negative correlation of LDR with GDPGR, which implies that loan deposit ratio affects inclusive growth in Nigeria negatively. This also obvious in the negative and weak correlation of LDR with FD, which indicates that the level of financial development is relatively low in Nigeria. In addition, FD has a positive but very weak correlation with GDPGR, which indicates that financial deepening has positive effect on inclusive growth in Nigeria. Furthermore, the negative and very weak correlation of CPS with LR shows that high lending rate discourages credit to private sector which supposed to improve financial development that can drive inclusive growth in Nigeria.

Table 2: Correlation Matrix Test Results

	GDPGR	CPS_GDP	BLR	FD	LDR	RL_RD
GDPGR	1.000000					
CPS_GP	0.141853	1.000000				
LR	0.050966	-0.037907	1.000000			
FD	0.028402	0.831942	0.210290	1.000000		
LDR	-0.120751	-0.009307	-0.108991	-0.282810	1.000000	
RL_RD	0.056175	0.098396	-0.048977	0.219306	-0.373161	1.000000

4.3. Unit Root Tests

The results in table 3 show that the variables under the study were integrated at either I(0) or I(1). Thus, as the order of integration varies, the study made use of ARDL approach to detecting long and short-run relationships.

Table 3: Unit Root Tests Results

Variables	ADF Test Statistic	Critical Value of ADF	Order of Integration	Remarks
GDPGR	-5.262351*	-3596616	I(0)	Level Stationary
PSC_GDP	-5.275397*	-3.610453	I(1)	Difference Stationary
LR	-3.022368**	-3.596616	I(0)	Level Stationary
FD	-5.634911*	-3.600987	I(1)	Difference Stationary
LDR	-4.917382*	-3.600987	1(0)	Level Stationary
RL_RD	-5.965745*	-3.596616	1(0)	Level Stationary

4.4. Bound Tests

The results in table 4 contain the Bound F-test for Co-integration along with the asymptotic critical values. The results show that F-statistics is greater than the lower critical bound value at 5% significance level and there is existence of cointegration among the variables. Therefore, there is a long run relationship among the variables in the presence of structural breaks stemming in the series for period 1981 to 2023 in Nigeria. This is also confirmed by the high COINTEQ coefficient in the error correction regression, which is highly significant (Table 6)

Table 4: ARDL Bounds Test Results/ Bounds Critical Values

Test Statistic	Value	K
F-statistic	4.355300	5

	Sample Size	10%	5%	1%
I(0)	35	2.331	2.804	3.900
I(1)	35	3.417	4.013	5.419
I(0)	40	2.306	2.734	3.657
I(1)	40	3.353	3.390	5.256
I(0)	Asymptotic	2.080	2.390	3.060
I(1)	Asymptotic	3.000	3.380	4.150

* I(0) and I(1) are respectively the stationary and non-stationary bounds

4.5. Long Run Estimate

The results in table 5 below show that financial deepening (FD) has negative impact on inclusive growth and significant. While loan deposit ratio (LDR) is insignificant and has negative impact on inclusive growth in Nigeria. All other variables are significant with positive impact on inclusive growth. This shows that the deposits mobilized were not properly channeled into productive investment (loans) in the economy. The current year value of private sector credit to GDP (CPS_GDP) and past one year value of lending rate (LR) contribute to inclusive growth in Nigeria. These findings occur due to moderate lending rate which stimulates borrowings. This past one year of LR tends to constitute a springboard for further deepening of financial system through policy consistence. The result supports the findings of Igwebuike *et al.* (2019) and Okafor and Ude (2022). The rural loan to rural deposit ratio (RL_RD) has a very low impact on inclusive growth but highly significant. This indicates low degree of rural financial system efficiency and shallow rural financial deepening on inclusive growth in Nigeria. This occurred due to low level of deposits mobilization in the rural areas, and assumption that rural people don't save. This has policy implication and formed part of our contribution to knowledge. Also, the past one year value of financial deepening (FD) is negative but significant. Thus, a unit increase in financial deepening reduces inclusive growth by 1.05. This finding corroborates the result of Agheli and Hadian (2017). This is also reflected in the past one year of inclusive growth which is negative and not in line with theoretical expectation. This decrease in growth discourages financial deepening and does not stimulate financial activities.

Table 5: Long Run Results

Variable	Coefficient	Std. Error	t-Statistic	Prob.*
GDPGR(-1)	-0.868430***	0.210436	-4.126814	0.0003
CPS_GDP	1.452809*	0.722041	2.012086	0.0547
LR(-1)	0.444366*	0.223069	1.992057	0.0570
FD(-1)	-1.052873**	0.477408	-2.205396	0.0365
LDR	-0.061598	0.057334	-1.074373	0.2925
RL_RD(-1)	5.88E-06***	1.98E-06	2.972633	0.0063

***1%significant level, **5%significancelevel,*10%significancelevel

4.6. Short Run Estimate

The results in table 6 below contain the short-term dynamics of the estimated parameter of the error correction term. All the variables under consideration are significant. Bank lending rate (LR) current year value is positive and consistent with long-run results. Also, the current year value of rural loan to rural deposit ratio (RL_RD) is very low but positive, which is consistent with long-run results. In addition, the current year value of financial deepening (FD) is negative but significant. Similar results were obtained by Agheli and Hadian (2017). The coefficient of the error term is negative and significant at 1 percent level.

The coefficient of 0.87 indicates that the concurrent speed of adjustment of financial deepening to long-run equilibrium after temporary disequilibrium and disconcertion is 87 percent.

Table 6: Short Run Results

Variable	Coefficient	Std. Error	t-Statistic	Prob.*
D(BLR)	0.285537***	0.088061	3.242474	0.0028
D(BLR(-1))	-0.131743	0.176579	-0.746087	0.4611
D(BLR(-2))	-0.290366***	0.066233	-4.383985	0.0001
D(BLR(-3))	-0.352902**	0.161663	-2.182943	0.0365
D(FD)	-2.165664***	0.650353	-3.329984	0.0022
D(RL_RD)	2.81E-06***	6.98E-07	4.025776	0.0003
COINTEQ	-0.868430***	0.202968	-4.278660	0.0002

***1%significant level, **5%significancelevel,*10%significancelevelSource:

4.7. Post Estimation Diagnostic Tests

Deducing from the results in table 7 below, the null hypothesis of no coexistent serial correlation (with F-Statistic = 1.46 (0.25)) cannot be rejected, as the p-value of the test statistic is greater than 0.05. So, the model is free from auto-correlation and homoscedastic. The model is not normally distributed, which is the same with normality test under descriptive statistics and requires for caution. In addition, the Ramsey RESET specification test also showed that the model does not suffer from the problem of omitted variables and linearity assumption at 5% level of significance. Thus, there is no reasonable evidence to invalidate the model, considering the fact that the estimates are robust in the absence of serial correlation and homoscedastic. Therefore, the model can be used for structural and policy implication analysis.

Table 7: Serial Correlation LM, Homoscedasticity Jarque-Bera and Ramsey Tests Results

Test	F-Statistic	Prob. Value
Breusch-Godfrey Serial Correlation	0.935699	0.4062
Heteroskedasticity Test Breusch-Pagan-Godfrey	1.312194	0.2702
Jarque-Bera	15.53325	0.0004
Ramsey Stability Test	5.475455	0.0276

4.8. Granger Causality Tests

The results in table 8 below show that only LDR granger cause RL_RD. This indicates that total loan to deposit ratio is the major contributor to the rural loan deposit ratio in Nigeria. It shows that majority of the rural loans granted were funded by mobilized deposits outside rural areas. This is also confirmed by descriptive statistics that at a particular period of the study more rural loans were granted than rural deposits. There is no causality between FD and GDPGR. This requires caution because of its policy implications. This finding is corroborated by Agheli and Hadian (2017). Also, it is in line with Apergis and Levine (2007) neutral hypothesis, which asserts that there is no relationship between financial development and economic growth. It corroborates the evidence from the literature that financial deepening is very shallow in Nigeria (Ozekhome, 2020).

Table 8: Granger Causality Tests Results

	Obs.	F-Statistic	Prob.
FD does not Granger Cause GDPGR	41	0.24838	0.7814
GDPGR does not Granger Cause FD		0.17769	0.8379
LDR does not Granger Cause RL_RD	41	0.56112	0.0912

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

The study critically examined the linkage between financial deepening and inclusive growth from the context of financing-growth nexus. The findings showed that private sector credits (PSC_GDP), bank lending rate (LR) and rural bank loan to rural deposit ratio (RL_RD) have positive and significant impact on inclusive growth. While financial deepening (FD) has negative but significant impact on inclusive growth. As for the loan deposit ratio (LDR), it has negative impact and insignificant. It can de deduced from the results that the real sector of the economic have been neglected and necessary to mobilize financial resources to finance the growth of the sector. Conclusively, it is evident that the findings from this study addressed the controversy between the finance-growth nexus as the relationship appears to produce new evidence and more valid results.

Precisely, the existence of neutral causality between financial deepening and inclusive growth requires for concern. This requires offering insights into the policies that can maximize deposits channelization and accumulation for the growth of private sector that can drive inclusive growth.

5.2. Recommendations

Extracting from the study discussions and outcomes, the following recommendations that are inclusive growth driven are made as follows: There is a need for the proper implementation of policies and strategies geared towards financial deepening in Nigeria, proper enhancement of credit delivery to the private sector, and augmentation of deposit mobilization and accumulation both in the urban and rural areas.

REFERENCES

Abdul, Z.H., & Adamu, A. (2016). Effect of financial inclusion on Human Development in Nigeria, In Aruwa S.A.S and Gbande, C.A (eds), Managing national economy in transition: Conference proceedings of 6th national conference of Faculty of Administration (pp. 62 – 75). Nasarawa State University, Keffi, Nigeria.

Acemoglu, D.; Aghion, P., & Zilibotti, F. (2006). Distance to Frontier, Selection and Economic Growth, Journal of the European Economic Association, 4(2), 34 – 74. https://doi.org/10.1162/jeea.2006.4.1.37

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In Handbook of Labor Economics (Vol. 4, pp., 1043 – 1171). Amsterdam: Elsevier-North. https://doi.org/10.1016/S0169-7218(11)02410-5

African Development Report (2012). African development report 2012:Towards Green Growth in Africa (Overview). https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/African_Development_Report_2012.pdf

Agheli, L., & Hadian, G. (2017). Financial Development, Commercial Development, and Economic Growth in the Selected Emerging and the Middle Eastern Countries. International Journal of Economics and Financial Issues, 7(3), 362-370. https://www.econjournals.com/index.php/ijefi/article/view/4829

Aghion P., Howitt, P., & Mayer-Foulke, D. (2005). The Effect of Financial Development on Convergence. Theory and Evidence, 120(3), 173 – 222. https://doi:10.1162/qjec.2005.120.1.173

Aizenman, J., Lee, M., & Park, D. (2012). The Relationship between Structural Change and Inequality: A Conceptual Overview with Special Reference to Developing Asia (ADBI Working PaperSeries No. 396). Asian Development Bank Institute. https://www.adb.org/publications/relationship-between-structural-change-and-inequality-conceptual-overview-specia

Akhator, A. P., & Marcus, G. O. (2018). Financial deepening and economic development of Nigeria: an empirical investigation. Veritas International Journal of Entrepreneurship Development, 1(1), 125-139.

Akpokerere, O.E. & Edafiaje, L.A. (2018). Effect of Financial Deepening on the Nigerian Economic Growth and Development. Journal of Social Development, 7(1), 1-16.

Ali, I., & Son, H. H. (2007). Measuring Inclusive Growth. Asian Development Review, 24(1), 11-31.

Al-Shawesh, M.H.A. & Kumar (2022). The Impact of Financial Deepening on Economic Growth: Empirical Evidence from Yemen (1994-2018). Journal of Economics, Management and Trade, 28(8), 70-77. https://doi.org/10.9734/jemt/2022/v28i830431

Amakiri, A.J., & Bobai, J.N. (2025). Financial Deepening and Economic Growth in Nigeria. IIARD International Journal of Economics And Business Management, 11(12). 308-335. https://doi:10.56201/ijebm.vol.11.no12.2025.pg308.335

Anand, R., Mishra, S., & Peiris, S. J. (2013). Inclusive Growth: Measurement and Determinants (IMF Working Paper No. 13/135). International Monetary Fund

Apergis, Ioannis Filippidis, I., & Economidou, C. (2007). Financial Deepening and Economic Growth Linkages: A Panel Data Analysis. Review of World Economics, 143(1), 179 – 198. https://doi.org/10.1007/s10290-007-0102-3

Appiah, M., Li, F., & Frowne, D. I. (2020). Financial Development, Institutional Quality and Economic Growth: Evidence from ECOWAS Countries, Organizations and Markets in Emerging Economies, 11(1), 6-7. https://doi.org/10.15388/omee.2020.11.20

Aslam, N., & Saeed, R. (2023). How financial development and regional integration promotes inclusive growth? Evidence from BRI countries? Journal of Economics Finance and Management Studies, 06(10). https://doi.org/10.47191/jefms/v6-i10-27

Bakker, M., & Messerli, H.R. (2017). Inclusive growth versus pro-poor growth: Implications for tourism development. Tourism and Hospitality Research, 17(4), 384-391. https://doi.org/10.1177/1467358416638919

Balago, G. S. (2014). Financial sector development and economic growth in Nigeria, An empirical investigation, International Journal of Finance and Accounting, 3 (4), 253 - 265. https://doi:10.5923/j.ijfa.20140304.05

Barnat, N., MacFeely, S., Cantú, F., Ryabtsev, N., Strobino, J. M., Filippova, M., & Karlan, A. (2023). Compiling an inclusive growth index: Methodological challenges, considerations and conclusions. Statistical Journal of the IAOS, 39(3), 1–21. https://doi.org/10.3233/SJI-230025

Beck, T., Levine, R., & Loayza N. (2000). Finance and Sources of Growth, Journal of Financial Economics, 58(3), 261 - 310. https://doi.org/10.1016/S0304-405X(00)00072-6

Bencivenga, V. R., & Smith, B. D. (1991). Financial Intermediation and Endogenous Growth, Review of Economic Studies, 58 (2), 195 – 209.

Calderon, C., & Liu, L. (2003). The direction of causality between financial development and economic growth, Journal of Development Economics, 72(3), 321-334. http://dx.doi.org/10.1016/S0304-3878(03)00079-8

Central Bank of Nigeria. (2012). National Financial Inclusion Strategy: Summary Report, Abuja, Nigeria.

Chukwu, J. O., & Agu, C. C. (2009). Multivariate causality between financial depth and economic growth in Nigeria. African Review of Money, Finance and Banking, 7-22. Research Center on International Cooperation of the University of Bergamo.

Djokoto, J.G. (2022). The investment development path and human development: Is there a nexus? Research in Globalization, 4, Article 100079. https://doi.org/10.1016/j.resglo.2021.100079

Dollar, D., & Kraay, A. (2002). Growth is Good for the Poor. Journal of Economic Growth, 7(3), 195 - 225. https://doi.org/10.1023/A:1020139631000

Efanga, U., Ogochukwu, C., & Ugwuanyi, O. (2020). The impact of financial deepening on the economy of Nigeria (1981-2018). Asian Journal of Economics, Business and Accounting, 15(1), 12-21. https://doi.org/10.9734/ajeba/2020/v15i130204

Ekane, R.O., Aguwamba, M.S., & Ughulu, S.E. (2024). Financial Deepening and Economic Growth in Nigeria. African Banking and Finance Review Journal (ABFRJ), 17(17), 100 – 109. https://www.abfrjournal.com/index.php/abfr/article/view/253.

Eniekezimene, F.A., & Chiazor, D. (2023). Financial Deepening and Economic growth in Nigeria. Numu Journal of Economic and Development Issues (YJEDI), 6(2), 103 – 119. ISSN: 2630-726X

Gatsi, J., Idun, A. A. A., & Mensah, F. (2020). Monetary policy and financial development in Africa: Do governance mechanisms matter? Journal of Economic and Financial Sciences, 13(1). https://doi.org/10.4102/jef.v13i1.419

Gezer, M.A. (2018). The relationship between financial deepening and economic growth: Bootstrap causality approach for the selected upper middle income countries, Theoretical and Applied Economics, 25(1), 95 – 112.

Ghildiyal, V., Pokhriyal, A.K., & Mohan, A. (2015). Impact of Financial Deepening on Economic Growth in India Perspective: ARDL Bound Testing Approach to Cointegration, Asian Development Policy Review, 3(3), 49 – 60.

Giri, A. K., & Mohapatra, G. (2012). Financial Development and Economic Growth: Evidence from Indian Economy. International Journal of Applied Research and Studies, 1(2), 1 - 17. http://www.ectap.ro/articol.php?id=1322&rid=130

Goldsmith, R. W. (1969). Financial Structure and Development. New Haven, C.T.: Yale University Press.

Gültekin, M., & Umutlu, M. (2023). Economic Growth and Financial Development: Evidence from Panel Cointegration Tests in Emerging Countries, International Journal of Contemporary Economics and Administrative Sciences, 8(1), 101-126. https://api.semanticscholar.org/CorpusID:259806223

Gurley, G. J., & Shaw, E. S. (1967). Financial Structure and Economic Development, Economic Development and Cultural Change, 15(3), 257 - 268. https://doi.org/10.1086/450226

Hammilton, O.I., & Godwin (2013). Does Financial Deepening Follow Supply Leading or Demand Following Hypothesis? A Critical Look at the Nigerian Evidence. Asian Journal of Science and Technology, 5 (1),10-15. http://www.journalajst.com

Hosono, A. (2022). Transforming economies for jobs and inclusive growth. In SDGs, Transformation, and quality growth. Insights from International Cooperation (pp.103-120), Springer. https://doi:10.1007/978-981-16-9748-7_5

Igwebuike, E. C., Udeh S. N., & Okonkwo, O. M. (2019). Effects of Financial Deepening on Economic Growth of Nigeria (1981-2016), International Academy Journal of Business Administration Annals, 7(1), 54 -67. ISSN: 2382-9175, April.

Iyoboyi, M. (2013). Bank and non-bank financial deepening and economic growth: The Nigerian experience (1981–2010). Review of Banking, Finance and Monetary Economics, 42(3), 247-272. https://doi.org/10.1111/j.1468-0300.2013.12008.x

Kamat, M., & Kamat, M. (2007). Does Financial Growth lead Economic Performance in India? Causality-Cointegration using Unrestricted Vector Error Correction Models (MPRA Paper No. 6154). Munich Personal RePEc Archive. uni-muenchen.de

Karimo, T. M., & Ogbonna, O. E. (2017). Financial Deepening and Economic Growth Nexus in Nigeria: Supply-Leading or Demand-Following? Economies, 5(4), 1 – 18. https://doi:10.3390/economies5010004

Kiprop, J. G. (2013). The relationship between financial deepening and the growth of small and medium enterprises in Kenya. Master's Thesis, University of Nairobi, Kenya.

Kolawole, K. D., Ijaiya, M. A., Sanni, M., & Aina, T. (2019). Impact of Financial Deepening on Economic Growth in Nigeria. Fountain University Osogbo Journal of Management (FUOJM). 4(2), 56-72.

Lewis, W. A. (1995). The theory of economic growth. London: Allen and Unwin.

Maxwell, F. J. (1989). Financial development: Theories and recent experience. Oxford Review of Economic Policy, 5(4), 13-28. https://ideas.repec.org/s/oup/oxford.html

Manasseh, C.O., Ngong, C.A., Logan, C.S., Okanya, O.C., & Olelewe, C.A. (2024). Financial deepening and economic growth nexus in emerging economies in Africa: does supply-leading or demand-following hold? International Review of Applied Economics, 38(3), 1-6. DOI: 10.1080/02692171.2023.2296506 https://www.researchgate.net/publication/377203100

Mckinnon, R. I. (1973). Money and Capital in Economic Development, Washington, D. C.: Brooking Institution.

Mitra, A., & Das, D. (2018). Inclusive growth: Economics as if people mattered. Global Business Review, 19(3), 756–770. https://doi:10.1177/0972150917713840

Mordi, N.O.C. (2010). The link between the financial (banking) sector and the real economy. Central Bank of Nigeria, Economic and Financial Review, 48(4), 7 – 30. https://dc.cbn.gov.ng/efr/vol48/iss4/11/

Ndlovu, G. (2013). Financial sector development and economic growth: Evidence from Zimbabwe, International Journal of Economics Finance, 3(5), 435 – 446. https://www.econjournals.com/index.php/ijefi/article/view/401

Ngerebo-a, T.A., & Lucky, A. L., (2016). Interest rates and deposit money bank's profitability: Evidence from Nigeria (1980 – 2014), International Journal of Empirical Finance, 5 (1), 22 – 35. https://ideas.repec.org/s/rss/jnljef.html

Nwaolisa, E., & Cyril, U. (2019). Assessment of financial deepening on the growth of Nigerian economy: 1990-2016. Journal of Contemporary Research in Business Economics and Finance, 1(2), 16-29. https://doi.org/10.33094/26410265.2019.12.16.29

Nwosu, C. P., Itodo A. I., & Ogbonnaya-Orji, N. (2021). Financial Deepening, Financial System Fragility and Economic Growth in Nigeria, Central Bank of Nigeria Economic and Financial Review, 59(1), 10 – 27. https://doi:10.48028/iiprds/ijsrssms.v8.i2.03

Nzotta, S.M., & Okereke, E.J. (2009). Financial deepening and economic development of Nigeria: An empirical investigation, African Journal of Accounting, Economics, Finance and Banking Research, 5(5), 52 - 66. https://ssrn.com/abstract=1534212

OECD (2015). All on board: Making inclusive growth happen. OECD Publishing.

OECD (2018). Opportunities for All. A Framework for Policy Action on Inclusive Growth; OECD: Paris, France.

Okafor H., & Nwosu, C. P. (2018). Dynamics of Nigeria banking system: Evolution, crises and reforms. Unpublished Manuscript. Chapter 2.

Okafor S. O., & Ude C. L. (2022). Financial Deepening and Economic Growth in Nigeria. International Journal of Management Research and Economics, 2(2), 81-85. https://doi.org/10.51483/IJMRE.2.2.2022.81-85

Okeke, I.J., & Akuna, R.C. (2023). Effect of Financial Deepening on Economic Growth in Nigeria. ANAN Journal of Contemporary Issues, 4(2), 51-67, August. https://journals.unizik.edu.ng/jocia/article/view/2520

Omotor, D. (2007). Financial development and economic growth: Empirical evidence from Nigeria, Niger. Journal of Economics Social Studies, 49(2), 209 – 234.

Onyemachi, C. (2012). An empirical analysis of financial deepening and economic growth (1986-2010), A Thesis submitted to the Postgraduate School in Ahmadu Bello University, Zaria, Nigeria in partial fulfilment for the award of Master of Science in Economics (M.SC). https://kubanni.abu.edu.ng/handle/123456789/1704

Osinsanwo, B.G. (2013). The macroeconomic effect of financial development on economic growth in Nigeria: A long run analysis, 1970-2011, Journal of African Macroeconomic Review, 4(2), 227 - 245.

Pradhan, R.P. (2010). Financial Deepening, Foreign Direct Investment and Economic Growth: Are They Cointegrated? International Journal of Financial Research, 1(1), 37 - 43. https://ideas.repec.org/s/jfr/ijfr11.html

Puatwoe, T.J., & Piabuo, M.S. (2017). Financial sector development and economic growth: Evidence from Cameroon. Financial Innovation. 3(1), 1-18. https://doi:10.1186/s40854-017-0073-x

Quartey, P., & Prah, F. (2008). Financial development and economic growth in Ghana: Is there a causal link? African Finance Journal, 10(1), 28 – 54. https://ideas.repec.org/s/afj/journl.html

Rajan R.G., & Zingales I. (20O3). Saving Capitalism from Capitalist: Unleashing the power of financial markets to create wealth and spread opportunity. Princeton University Press.

Ravallion, M., & S. Chen (2003). Measuring Pro-Poor Growth. Economics Letters, 78(1), 93–99. https://ideas.repec.org/s/eee/ecolet.html

Safdar, L. (2014). Financial Deepening and Economic Growth in Pakistan: An Application of Cointegration and VECM Approach, Interdisciplinary Journal of Contemporary Research in Business, Institute of Interdisciplinary Business Research 5(12), 368 – 384, April. https://journal-archieves36.webs.com/368-384apr14.pdf

Sahay, R., Cihak, M., N'Diaye, P., & Barajas, A. (2015). Rethinking financial deepening: Stability and growth in emerging markets. Journal of Institutional Economics, 17 (33), 73-107. https://ideas.repec.org/s/imf/imfsdn.html

Schumpeter J.A. (1934). The Theory of Economic Development, Cambridge, MA, Harvard University Press.

Shan, S., & Liu, C. (2023). The Impact of Financial Deepening on Digital Economy Development: An Empirical Analysis from China. Sustainability, 15, p. 11358. https://doi.org/10.3390/su151411358

Shaw, E. S. (1973). Financial Deepening in Economic Development, Cambridge, M.A.: Harvard University Press.

United Nations Development Programme. (2017). UNDP's strategy for inclusive and sustainable growth. United Nations Development Programme. Available online: https://www.undp.org/publications/undps-strategy-inclusive-and-sustainable-growth

Withers, G. (2018). Inclusive growth. Routledge EBooks, pp. 235–252. https://doi.org/10.43 24/9781351245944-14

World Bank (1989). World Development Report.

Wycliffe, M., Jonathan, M., Tabby, G., Ruth, N., & Kitheka, S. (2013). Financial Deepening and Economic Competitiveness In Kenya: The Strides To Being An Economic Power House, International Journal of Management & Information Technology. 6(2), 817-825. https://doi:10.24297/ijmit.v6i2.737