



Eastern Africa Journal of Official Statistics

Website: <https://journal.eastfc.ac.tz/index.php/eajos/index>



Drivers of Bilateral Trade Flows between Tanzania and Kenya: An Augmented Gravity Model Approach

Leonard N.M. Amani

Eastern Africa Statistical Training Centre, Box 35103, Dar es Salaam, Tanzania

Email: leombigili@gmail.com

KEYWORDS

Gravity Model,
Bilateral Trade,
VECM, Drivers,
Trade Flows

ABSTRACT

This study examines the determining factors of bilateral trade flows between Tanzania and Kenya using an augmented gravity model. The study relies on secondary time series data obtained from decent international databases, including the World Bank, United Nations Comtrade, and International Monetary Fund, to ensure evenness, reliability, and equivalence. After the estimation, the empirical findings indicate that trade openness has a positive influence on trade flows, as a 1% increase in openness leads to approximately a 16.34% increase in trade, reflecting a strong and substantial effect of liberalisation on trade expansion. Similarly, Kenya's GDP exerts a positive influence, with a 1% increase resulting in about a 13.06% rise in trade flows, highlighting the significant role of Kenya's economic growth in driving bilateral trade. In contrast, Tanzania's GDP shows a negative influence, where a 1% increase leads to a 4.26% decrease in trade flows, suggesting that domestic growth may reduce reliance on bilateral trade. Distance, unexpectedly, has a positive influence, with a 1% increase associated with a 2.46% rise in trade flows, indicating that improved connectivity and trade complementarities may offset distance-related costs. Based on these findings, the study recommends that Tanzania should enhance production in sectors that complement Kenya's demand, particularly in industrial and manufactured goods, to benefit from Kenya's economic growth. It further recommends promoting export-oriented industries, improving product quality, and strengthening trade promotion mechanisms to ensure domestic growth translates into increased trade. Additionally, both countries should sustain trade liberalisation through reduced tariffs, streamlined customs procedures, and harmonised regulations, while continuing to invest in transport infrastructure, logistics, and regional connectivity to improve trade efficiency and maximise bilateral trade potential.

1. Introduction

Regional economic integration has increasingly become a central pillar for enhancing trade flows, economic growth, and structural transformation in developing countries, particularly within the East African Community (EAC). The community with seven countries, including Tanzania, Kenya, Uganda, Rwanda, Burundi, South Sudan, and the Democratic Republic of the Congo established to promote free movement of goods, services, and factors of production. In recent years, the bloc has recorded notable growth in trade volumes, reflecting both improved regional cooperation and expanding productive capacities. For instance, total EAC trade increased substantially, with merchandise trade rising by about 28.4% to USD 38.2 billion in 2025, driven largely by a 40.5% surge in exports, while imports grew at a slower pace (EAC, 2025).

Similarly, the Intra-EAC trade has shown a steady upward trajectory over the past decade, reflecting ongoing efforts toward regional integration, trade facilitation, and the implementation of the Common Market Protocol. Data shows that the value of intra-EAC trade increased from approximately USD 5.8 billion in 2015 to about USD 10.2 billion in 2019, before experiencing a temporary slowdown in 2020 due to disruptions associated with the COVID-19 pandemic (EAC, 2024; World Bank, 2024). Nevertheless, recovery was observed thereafter, with trade rising to around USD 11.4 billion in 2022 and reaching approximately USD 12.1 billion in 2023 (EAC, 2024). Despite this growth, intra-regional trade still accounts for only about 15% of the EAC's total trade, which exceeds USD 80 billion, indicating that a significant share of trade continues to occur with external partners (EAC, 2024; UNCTAD, 2024).

Within the EAC, Kenya and Tanzania account for one of the largest shares of intra-regional trade, consistently ranking among the bloc's leading trading partners. The bilateral trade between Tanzania and Kenya has emerged as one of the most dynamic within the EAC, both in terms of volume and growth. Over the past decade, trade flows between the two countries have expanded significantly, reflecting strong geographical proximity, long-standing historical ties, and increasingly complementary production structures.

The Tanzania and Kenya bilateral trade stands out as the most significant compared to Tanzania's trade with other partner states. In 2024, for instance, Tanzania's exports to Kenya reached approximately USD 320.9 million, while imports from Kenya amounted to about USD 400.4 million, bringing total bilateral trade to over USD 720 million (UN COMTRADE, 2026). In contrast, Tanzania's trade with all other EAC countries remains relatively lower, typically ranging between USD 50 million and USD 250 million annually (EAC, 2024). As a result, Kenya alone accounts for a substantial share often exceeding 25–30% of Tanzania's total intra-EAC trade, clearly indicating

that the Kenya–Tanzania trade corridor dominates both in value and share and plays a central role in shaping regional trade dynamics (UN COMTRADE, 2026; EAC, 2024).

The bilateral trade trend reflects asymmetries in industrial capacity, with Kenya exporting relatively more manufactured goods while Tanzania’s exports are more resource- and agriculture-based. Moreover, recent trends indicate a significant increase in Kenya’s imports from Tanzania, with trade flows expanding substantially in recent years, positioning Tanzania as an increasingly important source of intra-EAC imports for Kenya (World Bank, 2024). This underscores the growing importance of Tanzania in regional supply chains and highlights evolving trade complementarities between the two economies.

It has been noted that regional and bilateral trade are widely recognised as key drivers of countries’ trade expansion and overall economic performance. As such, the effectiveness of these trade activities needs to go hand in hand with the implementation of appropriate measures and policies, which should be grounded in empirical evidence. Knowing that, numerous studies in Tanzania and across other regions have been undertaken to examine the determinants of such trade dynamics. Based on the empirical examination, several factors have been identified as the main drivers of the trade, mainly through the gravity framework.

For example, Makilagi and Oh (2023) found that GDP significantly promotes Tanzania’s trade flows, while distance reduces trade, indicating that the country trades more intensively with larger and geographically closer economies. The study further emphasises that Tanzania’s trade potential is stronger within nearby regional markets, pointing to the role of regional integration frameworks such as the EAC and AfCFTA in enhancing trade outcomes. In addition, Taylor and Luvanda (2022) show that exchange rates and long-run economic relationships are key determinants of Tanzania’s bilateral trade with China, suggesting that macroeconomic stability and sustained economic linkages are crucial in shaping trade patterns. Their findings further indicate that factors such as foreign direct investment (FDI) and trade policies play an important role in influencing trade dynamics.

Across Africa, empirical studies reveal a consistent set of factors, alongside important institutional and policy influences. Evidence shows that larger economies, higher population levels, shared borders, and a common language significantly enhance trade, while distance and high transport costs continue to constrain trade expansion (Gülseven et al., 2023). Regional integration efforts, including COMESA, EAC, and SADC, have generally contributed to increased trade flows, although their effectiveness varies depending on institutional strength and implementation capacity (Olayungbo & Iqbal, 2021). Furthermore, governance-related factors have been found to play a critical role, with democratic governance improving trade performance, particularly in manufacturing sectors

(Boungou et al., 2023), while weak governance and corruption can hinder trade within regions such as the EAC (Oparanya et al., 2019).

Additional evidence underlines the role of trade policies and broader structural factors in shaping trade outcomes across the continent. Trade liberalisation measures, including tariff reductions and regional trade agreements, have been shown to significantly boost intra-regional trade in East Africa, although trade imbalances remain a persistent challenge for countries such as Tanzania and Kenya (UNCTAD, 2023). Similarly, participation in regional and intercontinental trade agreements has been found to increase bilateral trade intensity in countries such as South Africa (Mchani & Phiri, 2025). Beyond economic and policy factors, socio-cultural and historical linkages, including shared language, colonial ties, and cultural proximity, also facilitate trade, while improvements in economic freedom and institutional reforms further strengthen trade relationships (Gonel, 2017).

Despite notable progress in the literature, important gaps remain. The review reveals that much of the existing research focuses on aggregate African trade patterns or relationships with external partners, with limited attention to detailed bilateral dynamics within the EAC, particularly between Tanzania and Kenya. Hence, there is a clear lack of recent empirical analysis on bilateral trade between the two countries. Available evidence indicates that the most prominent study examining the drivers of Tanzania and Kenya trade was conducted over a decade ago by Mahona and Mjema (2014). This gap is especially critical given that trade between the two countries has grown rapidly, signalling strengthening economic linkages, yet the underlying drivers of this expansion remain poorly understood. The absence of up-to-date empirical evidence on the factors shaping this bilateral relationship emphasises the need for a more focused and systematic investigation.

Against this backdrop, the study employs an augmented gravity model to achieve two specific objectives: first, to examine the effects of economic sizes and geographical distance on bilateral trade flows between Tanzania and Kenya; and second, to assess the influence of trade liberalisation on the trade relationship between the two countries. Accordingly, the study seeks to answer the research question: What are the effects of economic size, geographical distance, and trade liberalisation on bilateral trade flows between Tanzania and Kenya? The ultimate goal is to generate updated and policy-relevant insights that can inform trade policy formulation and regional integration strategies.

2. Methodology

2.1 Data Type and Source

This study relies on secondary time series data obtained from trustworthy international databases to ensure consistency, reliability, and comparability. Data on Gross Domestic Product (GDP) for both Tanzania and Kenya were sourced from the World Bank, while the Real Effective Exchange Rate

(REER), the proxy of distance, was obtained from the International Monetary Fund. Bilateral trade flow data between Tanzania and Kenya were drawn from the United Nations Comtrade database. In addition, trade openness was computed by the researcher using Tanzania's GDP alongside export and import data. The use of these secondary data sources enhances the credibility of the analysis, as they are widely recognised for their methodological rigour and comprehensive coverage, making them suitable for econometric analysis of international trade relationships.

2.2 Pre-estimation Tests

Pre-estimation tests are essential in time series analysis to ensure the validity and reliability of econometric results. In this study, two tests: unit root and cointegration tests are conducted before model estimation. Together, the results of these tests guide the appropriate model specification.

2.2.1 Unit Root Test

The unit root test is a fundamental step in time series analysis, as it determines whether the statistical properties of a variable, such as its mean and variance, remain constant over time. Ensuring stationarity is crucial because the use of non-stationary data in regression analysis can lead to misleading and spurious results, thereby compromising the validity of the findings. There are several techniques for testing stationarity, including the Dickey–Fuller (DF) test, Augmented Dickey–Fuller (ADF) test, the Phillips–Perron (PP) test, and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. In this study, the Phillips–Perron (PP) test is employed over the others. The PP test uses non-parametric corrections to account for serial correlation and heteroskedasticity in the error terms. This makes it particularly advantageous in empirical applications, as it reduces the risk of model misspecification and provides more consistent and efficient estimates.

2.2.2 Cointegration Test

It is crucial to verify whether variables share a long-run equilibrium relationship to avoid deceptive results. Cointegration tests serve this purpose by identifying stable long-term associations among variables. Cointegration can be tested using a number of methods, including the Engle–Granger two-step method, the Johansen test, the Autoregressive Distributed Lag (ARDL) bounds testing approach, the Phillips–Ouliaris test, the Banerjee–Dolan method, and the Stock–Watson dynamic ordinary least squares (DOLS) approach.

In the present study, the Johansen cointegration test is employed due to its ability to handle multiple variables and identify more than one cointegrating relationship simultaneously. The Johansen approach uses a maximum likelihood framework, providing greater efficiency and robustness when analysing complex systems. These advantages make the Johansen test particularly suitable for capturing the long-run dynamics of the variables in this study.

2.3 Lags Selection

Selecting the appropriate number of lags in time series analysis is a critical issue because it ensures that the model adequately captures the dynamics of the variables without overfitting or losing efficiency. This is particularly important because, in time series data, past periods often influence the current period's outcomes. Thus, failing to account for these effects can lead to inaccurate results. It is important to note that incorrect lag length can result in biased parameter estimates, poor forecasting performance, or mis-specified error structures.

In selecting the optimal lag, several techniques are used, including the Akaike Information Criterion (AIC), Schwarz Bayesian Information Criterion (BIC), Hannan–Quinn Information Criterion (HQIC), Final Prediction Error (FPE), and the Likelihood Ratio (LR) test. In this study, all these techniques are applied to determine the optimal lag length. Using multiple criteria provides robustness to the selection process, as each method balances model fit and parsimony differently. By considering all methods, the study ensures that the chosen lag structure reliably captures the short-run dynamics while maintaining model efficiency and accuracy.

2.4 The Gravity Model Framework

The gravity model of trade in international economics predicts that bilateral trade flows are based on the economic sizes (often using GDP measurements) and the distance between two units. The basic theoretical model (standard model) for trade between two countries (i and j) takes the form of:

$$X_{ij} = G \cdot \frac{GDP_i^\beta * GDP_j^\beta}{D_{ij}^\beta} \dots \dots \dots (1)$$

where X is the bilateral trade between countries i and j, GDP is the gross domestic product of countries i and j, respectively, D is the distance between the trading countries, and G is a constant. The coefficients β represent the elasticities. In other words, how sensitive is the volume of bilateral trade flow to the GDP of the home country, to the GDP of the partner country, or their bilateral distance. Taking logarithms, we can convert equation (1) into a linear form for econometric analysis, where constant G becomes α as shown in equation (2), and the equation becomes:

$$\ln X_{ij} = \alpha + \beta_1 \ln GDP_i + \beta_2 \ln GDP_j - \beta_3 \ln D_{ij} + \varepsilon \dots \dots \dots (2)$$

The mutual attractive force in the gravity models in trade is represented by trade flow, and the model suggests an analogical relationship between this force, GDP, which stands for the size (mass) of both partner countries, and their distance. The trade appears to depend positively on the GDPs of the trade partners and negatively on the shipping distance separating them (Bergeijk and Brakman, 2010).

This trade model provides a strong analytical foundation for this study, as it has been widely used to identify the key determinants of bilateral trade flows. In particular, core variables emphasised by the model, which are economic sizes (GDP) and geographical distance, are incorporated into this analysis due to their established influence on trade patterns. In addition, the model is extended by including trade liberalisation as an augmented variable, allowing for a more comprehensive examination of policy effects alongside traditional determinants. So, the model came to be known as the augmented gravity model.

The use of the augmented gravity model is justified by the limitations of the basic gravity model in explaining modern international trade patterns. While the basic gravity model mainly focuses on economic size and geographical distance as determinants of trade flows, it overlooks other important factors such as trade liberalisation, exchange rates, foreign direct investment, institutional quality, infrastructure, and trade policies that significantly influence trade performance. Therefore, researchers augment the gravity model to incorporate these additional variables, enabling a more comprehensive, realistic, and empirically robust analysis of the determinants of bilateral trade flows.

A gravity model augmented by trade liberalisation can be expressed as:

$$X_{ij} = \left(G \cdot \frac{GDP_i^{\beta_1} * GDP_j^{\beta_2} * \beta_4 TO_{ij}}{D_{ij}^{\beta_3}} \right) \dots \dots \dots (3)$$

In logarithmic form, which is commonly used for estimation, the model becomes:

$$\ln X_{ij} = \alpha + \beta_1 \ln GDP_i + \beta_2 \ln GDP_j - \beta_3 \ln D_{ij} + \beta_4 \ln TO_{ij} + \varepsilon \dots \dots \dots (4)$$

In addition to the definitions of all other variables, TO stands for trade Openness.

2.5 Theoretical Framework of the Vector Error Correction Model (VECM)

The VECM is widely employed in econometric analysis and is typically estimated using the maximum likelihood method under the assumption of normality (Maekawa & Setiawan, 2014). The model is derived from the standard Vector Autoregression (VAR) framework. However, it was observed that the VAR model alone may fail to adequately capture both the short-run dynamics and the long-run equilibrium relationships among variables, revealing limitations in time series analysis. To address this limitation, the concept of error correction was introduced through the Engle–Granger two-step method, which was later extended to the multivariate context by Johansen (Engle & Granger, 1987).

Subsequently, Johansen proposed the use of the VECM to estimate and test for co-integration among multiple time series. He emphasised that the model is superior because it simultaneously captures short-run dynamics and long-run equilibrium relationships by incorporating an error correction term that adjusts deviations from equilibrium (Johansen, 1988, 1991). In essence, the VECM can be regarded as a special case of the VAR model designed specifically for non-stationary time series that share a common stochastic trend. It is often described as a constrained VAR model that includes a co-integration term, thereby preserving long-run equilibrium relationships among variables while allowing for short-run fluctuations (Enders, 2014). Similarly, Kazanas (2017) describes the VECM as a constrained VAR model expressed in first differences.

The VECM specification, therefore, restricts the long-run behaviour of endogenous variables to converge toward their long-run equilibrium relationships while simultaneously accommodating short-run adjustments (Engle & Granger, 1987). This is achieved by incorporating an error correction mechanism (ECM) into the model, which has proven highly useful for modelling non-stationary time series (Kazanas, 2017). As a result, the VECM explicitly integrates co-integration relationships, making it particularly valuable for policy analysis and forecasting (Engle & Granger, 1987). The VECM representation can be specified from a VAR(p) model as follows:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t \dots \dots \dots (5)$$

Y_t is an $(n \times 1)$ vector of endogenous variables, A_i is a coefficient matrix, and ε_t is a white noise error term.

Rearranging the VAR to express it in first differences:

$$Y_t - Y_{t-1} = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} - Y_{t-1} + \varepsilon_t \dots \dots \dots (6)$$

Define the first difference:

$$\Delta Y_t = Y_t - Y_{t-1} \dots \dots \dots (7)$$

Rearrange the lag structure:

$$\Delta Y_t = (A_1 + A_2 + \dots + A_p - I)Y_{t-1} - (A_2 + \dots + A_p)\Delta Y_{t-1} - \dots - A_p \Delta Y_{t-p+1} + \varepsilon_t \dots \dots \dots (8)$$

Define the long-run and short-run matrices:

$$\pi = (A_1 + A_2 + \dots + A_p) - I \dots \dots \dots (9)$$

$$\tau_i = -(A_{i+1} + \dots + A_p), i = 1, 2, \dots, p - 1 \dots \dots \dots (10)$$

Obtain the VECM representation:

Substituting the above definitions gives the VECM

$$\Delta Y_t = \pi Y_{t-1} + \sum_{i=1}^{p-1} \tau_i \Delta Y_{t-i} + \varepsilon_t \dots\dots\dots (11)$$

Introduce the cointegration decomposition

If the variables are cointegrated, the mark π can be decomposed:

$$\pi = \alpha \beta' \dots\dots\dots (12)$$

Where:

- β' represent the cointegrating vector (long-run relationships)
- α represents the speed of adjustment parameter toward equilibrium

Final VECM form:

The VECM therefore becomes,

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \sum_{i=1}^{p-1} \tau_i \Delta Y_{t-i} + \varepsilon_t \dots\dots\dots (13)$$

Here:

$\beta' Y_{t-1}$ is the error correction term (ECT) representing long-run equilibrium deviation.

α measures how quickly variables adjust back to equilibrium.

τ_i capture short-run dynamics.

The application of the model requires that all variables included in the system be integrated of the same order, preferably integrated of order one, I(1), meaning they are non-stationary at levels but become stationary after first differencing. A necessary condition for the use of VECM is the existence of at least one cointegrating relationship among the variables, which indicates the presence of a stable long-run equilibrium relationship.

Then equation (17) was transformed into a VECM equation. Since the VECM incorporate short-run dynamics (differences) and long-run equilibrium (error correction term), the equation became:

$$\begin{aligned} \Delta TF_t = & \beta_0 + \lambda ECT_{t-1} \\ & + \sum_{i=1}^{p-1} \beta_{1i} \Delta \ln TF_{t-1} \\ & + \sum_{i=1}^{p-1} \beta_{2i} \Delta \ln TGDP_{t-1} \\ & + \sum_{i=1}^{p-1} \beta_{3i} \Delta \ln KNGDP_{t-1} - \sum_{i=1}^{p-1} \beta_{4i} \Delta \ln DIST_{t-1} + \sum_{i=1}^{p-1} \beta_{5i} \Delta \ln TO_{t-1} + \varepsilon_t \dots \dots (18) \end{aligned}$$

The definition of variables remains the same as in equation (14). But β_0 represents a constant, β_1 to β_5 refer to the coefficients of the variables (the actual effect that each variable is expected to have on trade flow), ECT is the error correction term, and ε_t represents the error term. The expected signs for all determinants' coefficients are positive, except distance, which is expected to have a negative coefficient, as stated by the gravity model.

The model specification demonstrates that the augmented gravity model establishes the theoretical long-run relationship by expressing trade flows as a function of economic size, geographical distance, and trade liberalisation in a log-linear form. Since the variables are non-stationary but cointegrated, this long-run gravity relationship is incorporated into the VECM as the cointegrating equation. Consequently, the VECM extends the augmented gravity model by not only capturing the long-run equilibrium relationship among the variables, but also explaining the short-run dynamics through the error correction mechanism, which measures the speed at which deviations from the long-run equilibrium are corrected over time.

2.7 Model Diagnosis Tests

Model diagnostics are essential in econometric analysis to ensure that the estimated results are reliable, unbiased, and suitable for inference. This study conducted a series of diagnostic tests, including checks for autocorrelation, normality, seasonality, and heteroskedasticity, to validate the underlying assumptions of the time series model. By performing these tests, the study ensures that any violations are identified and addressed, thereby enhancing the accuracy, robustness, and credibility of the empirical findings. The descriptions of each test are in the sub-sections below.

2.7.1 Autocorrelation Test

Autocorrelation testing plays a crucial role in time series analysis as it ensures that the error terms are independent over time, a key requirement for efficient estimation and reliable statistical inference. When this assumption is violated, standard errors may become biased, leading to misleading conclusions regarding the significance of variables. In this study, the Breusch–Godfrey test was employed due to its superiority over alternative methods. The Breusch–Godfrey test is particularly advantageous because it can identify higher-order serial correlation and remains valid in more complex model specifications. Moreover, its flexibility makes it well-suited for econometric frameworks such as VECM, thereby providing a robust and reliable approach for ensuring model adequacy.

2.7.2 Heteroscedasticity test

Testing for heteroskedasticity is essential because it helps verify whether the variance of the error terms in a regression model remains constant across observations, an assumption underpinning the Ordinary Least Squares framework. When heteroskedasticity is present, the estimated coefficients remain unbiased but become inefficient, and more importantly, the standard errors are distorted, leading to unreliable hypothesis testing and misleading statistical inferences. This study employs the Breusch–Pagan test because it is superior compared to other techniques. Principally, its simplicity, direct linkage to the regression model, and ability to explicitly test whether the variance of the residuals is systematically related to the explanatory variables, thereby providing clear and reliable results for empirical analysis.

2.7.3 Normality test

Ensuring that the residuals follow a normal distribution is crucial for obtaining reliable statistical inference in econometric analysis. When this condition is not satisfied, hypothesis tests such as t-tests and F-tests may yield misleading results, potentially leading to incorrect conclusions about the significance of variables. The study was therefore required to conduct a normality test to validate this assumption. In doing so, the Jarque–Bera test was employed because of its effectiveness in jointly evaluating skewness and kurtosis simply and efficiently. Its clear decision rule and strong applicability in large samples make it a reliable and widely accepted tool for testing normality in empirical analysis.

3. Results Presentation and Discussion

3.1 Results Presentation

3.1.1 Descriptive Statistics

Results from Table 1 reveal that Kenya's GDP exhibits the highest variance (0.76474), followed by Tanzania's GDP (0.62459). This indicates that both Kenya's and Tanzania's GDP experienced relatively large fluctuations over the period under study. In practical terms, these variables are more volatile and show greater deviations from their mean values, suggesting that economic activity in both countries has not been entirely stable. Such high variability can strongly influence trade dynamics, as GDP fluctuations often affect production capacity, income levels, and demand.

In contrast, Trade Flow (0.14961) and Trade Openness (0.06331) show moderate variance, implying that they are relatively more stable compared to GDP. Although they do vary over time, their fluctuations are less pronounced, suggesting a certain level of consistency in trade-related behaviour. This stability is often desirable in econometric modelling, as it reduces the risk of erratic relationships.

Finally, Distance (proxied by real effective exchange rate) has the lowest variance (0.022724), indicating that it is highly stable with minimal dispersion. This is expected because distance is typically a fixed or slowly changing variable in gravity models. Its very low variability confirms that it does not change significantly over time and mainly serves as a structural factor rather than a dynamic one. Overall, the variance analysis suggests a clear distinction: macroeconomic variables (GDPs) are highly volatile, while trade-related variables are moderately stable, and geographical factors (distance) are nearly constant.

Regarding skewness and kurtosis, the results show that Trade Flow (-0.3058), Tanzania GDP (-0.3766), and Kenya GDP (-0.1278) are slightly negatively skewed, indicating mild left-tailed distributions, while Distance (0.3115) and Trade Openness (0.1721) are positively skewed, suggesting slight right-tailed behaviour; however, all skewness values are close to zero, implying near symmetry. In terms of kurtosis, Trade Flow (1.9746), Tanzania GDP (2.0519), Kenya GDP (1.4681), and Trade Openness (2.0159) have values below 3, indicating platykurtic distributions (flatter with fewer extreme values), whereas Distance (3.7044) exceeds 3, suggesting a leptokurtic distribution (more peaked with heavier tails). Overall, these figures indicate that most variables are approximately normally distributed with only minor deviations.

Table 1: Descriptive statistics results

Variables	Mean	Std. dev.	Variance	Min	Max	Skew	Kurtosis
lnTF	6.28030	0.38679	0.14961	5.55296	6.85646	-0.3058	1.97461
lnTGDP	3.31566	0.79031	0.62459	1.66013	4.37021	-0.3766	2.05191
lnKNGDP	3.61027	0.87450	0.76474	2.20277	4.82431	-0.1278	1.46813
lnDIST	4.66668	0.15074	0.02272	4.360053	5.0123	0.31154	3.70436
lnTO	3.64878	0.25161	0.06331	3.23869	4.17105	0.17209	2.01585

Source: WB, IMF, UN Comtrade

3.1.2 Unit Root

The results of the Phillips–Perron Unit Root Test in Table 1 indicate that all variables are non-stationary at the level but become stationary after first differencing. At level, the p-values for Tanzania GDP (0.0591), Kenya GDP (0.8031), Distance (0.0793), and Trade Openness (0.1083) are all above the 5% significance level. This means the null hypothesis of a unit root cannot be rejected, and the variables are non-stationary. Thus, variables were differenced to see the possibility of becoming stationary. Interestingly, after first differencing, all variables become statistically significant with p-values below 0.05. Specifically, Tanzania GDP (0.0005), Kenya GDP (0.0009), Distance (0.0004), and Trade Openness (0.0161), indicating that they are stationary at the first difference. Therefore, all variables are integrated of order one, I(1), which justifies proceeding with cointegration analysis to examine if long-run relationships exist among them.

Table 2: Unit root test results

Variable	Level I(0) Test Statistic (P-Value)		First Difference I(1) Test Statistic (P-Value)		Order Of Integration
	Coeff.	P-value	Coeff.	P-value	
lnTGDP	-2.794	0.0591	-4.263	0.0005	I(1)
lnKNGDP	-0.853	0.8031	-4.134	0.0009	I(1)
lnDIST	-2.670	0.0793	-4.352	0.0004	I(1)
lnTO	-2.53	0.1083	-3.274	0.0161	I(1)

Source: Author's computation

3.1.3 Cointegration Results

The results of the Johansen Cointegration Trace Test indicate the existence of long-run relationships among the variables. The trace statistics for $r = 0$ (153.1075), $r = 1$ (73.7746), $r = 2$ (42.1409), and $r = 3$ (18.3697) are all greater than their respective critical values at the 5% level, leading to rejection of the null hypotheses at those ranks. However, at $r = 4$, the trace statistic 2.3266 is less than the critical value 3.76, meaning the null hypothesis cannot be rejected. This implies that there are four cointegrating equations, confirming that the variables share long-run equilibrium relationships. All are clearly presented in Table 3.

Table 3: Cointegration test results

Max. Rank (r)	Params	LL	Eigenvalue	Trace Statistics	Critical Value at 5%
0	30	204.31279		153.1075	68.52
1	39	243.97924	0.94118	73.7746	47.21
2	46	259.79604	0.67689	42.1409	29.68
3	51	271.68166	0.57215	18.3697	15.41
4	54	279.70324	0.43615	2.3266*	03.76
5	55	280.86652	0.07973		

Source: Author's computation

3.1.4 Lags Selection

To determine the optimal lag structure, several selection techniques were applied, including the Likelihood Ratio Test (LR), Final Prediction Error (FPE), Akaike Information Criterion (AIC), Hannan–Quinn Information Criterion (HQIC), and Schwarz Bayesian Information Criterion (SBIC). The results as presented in Table 4 indicate that the optimal lag length is four, as evidenced by the significant LR statistic (168.61) and the lowest values of FPE ($1.0e-17$), AIC (-28.054), HQIC (-26.5909), and SBIC (-22.9732) at lag 4, suggesting that a four-lag specification is the most appropriate for the estimation of the present model.

Table 4: Lag selection results

LAG	LL	LR	DF	P	FPE	AIC	HQIC	SBIC
0	80.1282				$2.1e-09$	-5.77909	-5.70942	-5.53715
1	267.431	374.61	25	0.000	$8.4e-15$	-18.2639	-17.8459	-16.8123
2	315.000	095.19	25	0.000	$1.9e-15$	-20.0023	-19.2359	-17.3409
3	385.398	140.74	25	0.000	$1.2e-16$	-23.4922	-22.3775	-19.6211
4	469.702	168.61*	25	0.000	$1.0e-17^*$	-28.054*	-26.5909*	-22.9732*

Source: Author's computation

3.1.5 Vector Error Correction Model (VECM) Estimation Results

Table 5 reveals that the error correction term (ce2-L1) is negative and statistically significant with a coefficient of -0.6045 ($p = 0.013$), confirming the existence of a stable long-run equilibrium relationship among the variables. The magnitude implies that approximately 60.45% of the previous period's disequilibrium is corrected within one period, indicating a relatively fast adjustment toward the long-run equilibrium. In addition, the model shows a reasonably good explanatory power with an R-squared of 0.5829, meaning that about 58.29% of the variation in trade flows is explained by the variables included in the model. The Root Mean Square Error (RMSE) of 0.1695 suggests that the model's prediction errors are relatively small. Furthermore, the Determinant of the covariance matrix [$\text{Det}(\text{Sigma_ml}) = 2.31e-18$] is positive, indicating that the estimated variance–covariance matrix is valid and the model estimation is stable.

The short-run results show that only trade openness with a two-period lag (lnTO-L2D) significantly influences trade flows. The coefficient of 16.3416 with a p-value of 0.064 indicates significance at the 10% level. This positive coefficient implies that an increase in trade openness leads to a rise in trade flows in the short run after two periods. Specifically, a 1% increase in trade openness is associated with about a 16.34% increase in trade flows, suggesting that liberalisation measures and reduced trade barriers gradually stimulate bilateral trade between the two countries over time.

The long-run results reveal several variables that significantly influence trade flows. Distance (lnDIST) has a positive and statistically significant coefficient of 2.4603 ($p = 0.000$), indicating that a 1% increase in distance is associated with a 2.46% increase in trade flows in the long run. Similarly, Kenya's GDP (lnKNGDP) shows a positive and highly significant coefficient of 13.0619 ($p = 0.000$), implying that a 1% increase in Kenya's economic growth leads to about a 13.06% increase in trade flows, reflecting the strong role of Kenya's economic expansion in stimulating bilateral trade. In contrast, Tanzania's GDP (lnTGDP) has a negative and statistically significant coefficient of -4.2588 ($p = 0.000$), suggesting that a 1% increase in Tanzania's GDP leads to about a 4.26% decrease in trade flows, which may indicate increased domestic production reducing reliance on imports.

Table 5: VECM estimation results

	Variables	Coefficient	Std. Error	Z-Stat.	P-Value
SHORT RUN	ce2- L1	-0.6045	0.4081752	-1.48	0.013
	lnTF-LD	0.2893	0.289646	1.00	0.318
	lnTF-L2D	-0.1243	0.31242	-0.40	0.691
	lnTO-LD	0.19302	5.778734	0.03	0.973
	lnTO-L2D	16.3416	8.812729	1.85	0.064
	lnTGDP-LD	0.65663	0.631273	1.04	0.298
	lnTGDP-L2D	-0.15309	0.489943	-0.31	0.755
	lnKNGDP-LD	0.12931	0.565626	0.23	0.819
	lnKNGDP-L2D	0.0637	0.605071	0.11	0.916
	lnDIST-LD	-0.2486	0.482063	-0.52	0.606
	lnDIST-L2D	0.4515	0.466720	0.97	0.333
Conc.	0.0232	0.261508	0.09	0.929	
LONG RUN	lnTFLOW	0.0000	(omitted)		
	lnTO	1.0000			
	lnDIST	2.4603	.2513461	9.79	0.000
	lnTGDP	-4.2588	.4891988	-8.71	0.000
	lnKNGDP	13.0619	1.659168	7.87	0.000
	Conc.	-48.1313			
R-SQ				0.5829	
RMSE				0.1695	
DET (SIGMA_ML)				2.31e-18	

Source: WB, IMF, UN Comtrade

3.1.6 Model Diagnosis Tests

Table 6 presents the three diagnostic test results conducted to examine the adequacy and reliability of the estimated model. They include autocorrelation, heteroskedasticity, and normality tests. The Breusch–Godfrey test for autocorrelation produced a chi-square statistic of 4.434 with a p-value of 0.2182. Since the probability value exceeds the 5% significance level, the null hypothesis of no autocorrelation cannot be rejected, implying that the residuals are free from serial correlation. Similarly, the Breusch–Pagan test was conducted to examine the presence of heteroskedasticity in the model. The results show a chi-square statistic of 1.52 with a p-value of 0.1086, which is also greater than the 5% significance level. Therefore, the null hypothesis of homoskedasticity cannot be rejected, indicating that the variance of the residuals remains constant and the model does not suffer from heteroskedasticity.

Furthermore, the Jarque–Bera test was employed to assess whether the residuals follow a normal distribution. The test yielded a Chi-square of 0.2876 with a p-value of 0.866, suggesting that the null hypothesis of normally distributed residuals cannot be rejected. This confirms that the residuals satisfy the normality assumption. Taken together, these diagnostic results indicate that the model satisfies the key classical regression assumptions, confirming that the estimated model is statistically adequate and well specified.

Table 6: Model diagnosis test results

Test Name	Chi-Square (Chi2)	P-Value
Autocorrelation Test	4.4340	0.2182
Heteroscedasticity Test	1.5200	0.1086
Normality Test	0.2876	0.8660

Source: Author’s computation

3.2 Results Discussion

The results show that trade openness significantly influences bilateral trade flows, implying that liberalisation and reduced trade barriers increase trade between the two countries. This is consistent with Taylor and Luvanda (2022), who emphasise the role of macroeconomic and trade-related factors in shaping bilateral trade. Similarly, evidence across Africa shows that trade liberalisation and regional agreements enhance trade, though their effectiveness depends on institutional capacity. Although the Gravity Model of Trade does not explicitly include policy variables, this result aligns with its core idea that lower trade costs increase trade. The large elasticity may reflect strong responsiveness to policy changes, especially where initial trade barriers were high.

Building on this, the findings on economic size support and contradict both the theory and the empirical results, depending on the country. The positive effect of Kenya’s GDP aligns with the

Gravity Model of Trade, which predicts that larger economies trade more, and is supported by Makilagi and Oh (2023). In contrast, the negative effect of Tanzania's GDP contradicts both theory and empirical evidence, which typically expect growth to promote trade. This may reflect import substitution, where increased domestic production reduces imports, or a stronger domestic market orientation. Moreover, structural differences, particularly Kenya's more diversified and export-oriented economy, may explain why Kenya's growth boosts trade while Tanzania's reduces it.

Again, the positive effect of distance contradicts both theory and empirical evidence. The Gravity Model of Trade predicts a negative relationship due to transport costs, as supported by Gülseven et al. (2023). However, the positive result may reflect factors not captured in the model, such as improved infrastructure, established trade corridors, or regional integration within the EAC. In line with this, evidence from Olayungbo and Iqbal (2021) suggests that integration can reduce trade costs and offset distance effects. Overall, these differences likely arise from structural factors, regional dynamics, and possible model limitations.

4. Conclusion and Recommendation

4.1 Conclusion

In conclusion, the study highlights that trade openness, Kenya's economic size and distance are key drivers of bilateral trade between Tanzania and Kenya, confirming the gravity model's prediction that larger economies and lower trade barriers promote trade. The findings also reveal that Tanzania's GDP has a negative influence on the trade flow. The negative effect of Tanzania's GDP and the positive effect of distance showed unexpected patterns, suggesting that structural differences, domestic market orientation, and regional integration dynamics play important roles in shaping trade flows. These results underline that while economic size, trade policies, and regional agreements facilitate bilateral trade, country-specific characteristics and evolving institutional frameworks can lead to deviations from standard theoretical expectations.

4.2 Recommendation

The study recommends several targeted policy measures to enhance bilateral trade between Tanzania and Kenya based on the key variables of the study. Regarding Kenya's GDP, which positively affects bilateral trade, policymakers in Tanzania should focus on enhancing sectors that complement Kenya's demand, particularly industrial and manufactured goods. Encouraging trade in these sectors will allow Tanzania to benefit from Kenya's economic growth while deepening regional economic linkages.

For Tanzania's GDP, which surprisingly shows a negative effect on trade, strategies should aim to convert domestic growth into increased regional trade. This can be achieved by supporting export-

oriented industries, improving product quality, and strengthening marketing and trade promotion mechanisms, ensuring that rising domestic production does not reduce participation in bilateral trade.

In terms of trade openness, both countries should continue to liberalise trade through reduced tariffs, streamlined customs procedures, and harmonised regulations. Such measures will enhance the responsiveness of trade flows to policy changes and sustain the positive impact of liberalisation on bilateral trade.

Finally, although distance showed a positive effect in the results, improving transport infrastructure, logistics, and connectivity between the two countries will further reduce trade costs, making trade more efficient and accessible across all regions. Investments in roads, ports, and border facilities, combined with regional integration efforts, will help maximise trade potential despite geographic constraints.

References

Bank of Tanzania. (2024). Monthly economic review: December 2024.

Bergeijk, P. A. G. van, & Brakman, S. (Eds.). (2010). *The gravity model in international trade: Advances and applications*. Cambridge University Press

Boungou, W., Osei-Tutu, F., & Zongo, A. (2023). Democracy and intra-Africa trade. *Empirical Economics*.

East African Community (EAC). (2024). EAC trade and investment report.

East African Community (EAC). (2025). Quarterly statistics bulletin.

Enders, W. (2014). *Applied Econometric Time Series* (4th ed.). Wiley.

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

Gonel, F. (2017). Turkey-Africa trade: A gravity model estimation of determinants.

Gulseven, O., Salam, S. A., & Alhadi, A. A. (2023). Can WTO membership boost intra-African trade? *Scientific African*, 21, e01728.

International Monetary Fund. (2026). *Real effective exchange rate (index, CPI-based)*. Retrieved from Federal Reserve Bank of St. Louis, FRED:

<https://fred.stlouisfed.org/series/TZAEREERIX>

Johansen, S. (1988). Statistical analysis of co-integration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.

- Johansen, S. (1991). Estimation and hypothesis testing of co-integration vectors in Gaussian vector autoregressive models. *Econometric*, 59(6), 1551-1580.
- Kazanas, T. (2017). A vector error correction forecasting model of the Greek economy. *Hellenic Fiscal Council*, 2, 2-6.
- Maekawa, K., & Setiawan, K. (2014). Estimation of vector error correction model with GARCH errors: Monte Carlo simulation and application. In Paper Presented at the International Conference on Economic Modelling–ECOMOD 2014.
- Mahona, B. K., & Mjema, G. D. (2014). Determinants of Tanzania and Kenya trade in the East African community: A gravity model approach. *Journal of Economics and Sustainable Development*, 5(4).
- Makilagi, B. N., & Oh, J. (2023). What determines Tanzania's trade? A gravity approach. *International Area Studies Review*, 26(4), 438–448.
<https://doi.org/10.1177/22338659231194946>
- Mchani, S., & Phiri, A. (2025). Can Increased Intra-Continental Trade Partnerships Diversify Export Baskets in Africa? *Managing Global Transitions*, 23(1).
- Olayungbo, D. O., & Iqbal, B. A. (2021). An empirical analysis of African trade blocs effects on the global economy: New evidence from the gravity model. *Future Business Journal*, 7(45).
- Oparanya, W., Mdadila, K. P., & Rutasitara, L. K. (2019). Determinants of bilateral trade in the East African Community: Application of the gravity model. *International Journal of Economics and Finance*.
- Taylor, R. S., & Luvanda, E. G. (2022). Analysis of Sino-Tanzania trade relations: The gravity model approach. *The African Review*, 49(3).
- United Nations COMTRADE. (2026). International trade statistics database.
- United Nations Conference on Trade and Development (UNCTAD). (2023). *Economic development in Africa report 2023: Trade and regional integration for sustainable development*. United Nations.
- United Nations Statistics Division. (2026). *UN Comtrade international trade statistics database (bilateral trade data for Tanzania and Kenya, 1995–2024)*. Retrieved March 13, 2026, from Trading Economics: <https://tradingeconomics.com>
- World Bank. (2024). *World Integrated Trade Solution (WITS): Kenya trade profile*.

World Bank. (2024). World Integrated Trade Solution (WITS): Tanzania trade summary.

World Bank. (2026). *World development indicators: GDP (current US\$) for Tanzania and Kenya.*

Retrieved from [World Bank Data Portal](#)