EFFECT OF ETHNIC FRACTIONALISATION ON INFORMALITY IN SUB-SAHARAN AFRICA

ABDULKAREEM, Abdulhafis Olagunju

Department of Economics
Kwara State University, Malete
a.oabdulkareem@yahoo.com
08034614045

Abstract

This study examined the effect of ethnic fractionalisation on informality in Sub-Saharan Africa, utilising two measures of informality (viz: Dynamic General Equilibrium (DGE) and Multiple Indicator Multiple Cause (MIMIC)) and employing a theoretical proposition on ethnic fractionalisation and economic dualism theory as the theoretical framework. Conducting a unit root test and then using the Kao panel test for the cointegration of data from 1993 to 2021, the paper aimed to establish a long-run relationship among the variables. Panel ARDL estimation technique was then utilised to analyse the data. Findings revealed a positive effect of ethnic fractionalisation on the informal sector, in line with the proposition advanced in the paper. Also, the control variables, comprising economic development, unemployment, educational attainment, and World governance combined indicators, were all found to have the expected effects on informality. The study concluded with recommendations for policymakers to formulate and implement policies aimed at reducing the rather undesirable effect of ethnic fractionalisation on the size of the informal economy and to also check the extent of informality effectively through other targeted interventions.

Keywords: Ethnic Fractionalisation, Informality, and Sub-Sahara Africa,

JEL Codes: M14, E26, O17

1. Introduction

In the past informality was considered as a symptom of underdevelopment. However, recently, the informal sector has been duly recognised to exist in both developed and developing countries (Loayza, 2018). In Africa, the informal sector forms a substantial part of the economy, providing jobs and a source of livelihood for millions of its population, especially women and youth (Schneider & Enste, 2000). This sector plays an important role in the social and economic space of Africa, as it reduces unemployment and contributes to the growth of GDP. However, the lack of registration and regulation that characterises this sector, poses several challenges such as reduced tax revenues for

the government, limited worker protections, and impediments to economic development (Loayza, 2018). Therefore, policymakers should seek to promote formalisation policies which require identifying those factors that determine the size of the informal sector.

Given the importance of formulating such a policy, some theoretical and empirical studies have addressed the issue. For instance, some theories have posited that economic development, stringent government regulation, globalisation, and institutions are important determinants of informality (Rostow, 1960; Boeke, 1953; De Soto, 1989). At the empirical level, several factors have also been identified as determining informality. These include economic development, government size, income inequality, inflation rate, and ease of starting a business (Berdiev et al., 2018; Esaku, 2021; Bulut, 2021; Nageri & Gunu, 2020; Schneider & Enste, 2000). It is however observed in the theoretical and empirical literature that it is virtually these orthodox or conventional explanatory variables that were being "re-cycled" from time to time and from study to study, with very few additions of new explanatory variables, thereby stalling the process being made in enlarging the number of factors identified as the determinants of informality.

What is therefore called for in this regard is to further expand the frontier of knowledge on the list of informality determinants by "thinking outside the circle" and aggressively exploring and searching for new or "virgin" informality determinants, with plausible propositions being put forward to rationalise why they are considered. In this regard, one such factor grossly overlooked in the existing literature is ethnic fractionalisation. This is a serious gap because, as discussed later in this paper, it can affect the extent of informality size in an economy. Accordingly, this issue is taken up in this paper which puts forward a proposition on the effect of ethnic fractionalisation informality and subjects the proposition to an empirical test.

The subsequent sections of the paper are structured as follows: the second section provides a comprehensive review of the pertinent literature, while the third section focuses on the methodology. The fourth section of the paper presents and discusses the results obtained from data analysis while the fifth section is on the conclusion and recommendations of the paper.

2.0 Literature Review

2.1 Theoretical Review

This section gives a review of informality theories. These theories include the modernisation theory, the theory of economic dualism, the structural theory, and the neo-liberal theory. The review of these theories is done below.

The modernisation theory, proposed by Rostow (1959) posited that an economy goes through five stages of economic growth in a linear progression: traditional society, preconditions for take-off, take-off, drive to maturity, and age of high mass consumption. Each stage represents a distinct phase of economic and social transformation, ultimately leading to full industrialisation and modernisation. He emphasised that the internal factors within a country are the primary drivers of underdevelopment. Contrary to Marxist perspectives, Rostow highlights the importance of adopting modern values and

institutions to achieve development. This theory suggests that informality will fade away as industrialisation and modernisation take place (Clement, 2015).

Boeke's theory of economic dualism originated in 1953. It offered significant insights into the coexistence of two distinct economic sectors within a society. Drawing from his research in the Dutch East Indies, Boeke highlighted the cultural specificity of institutions and argued against the top-down imposition of external frameworks. He asserted that indigenous institutions persist alongside introduced ones, resulting in a dualism of prevailing institutions within society. This theory posited the existence of a traditional subsistence sector (informal economy) and a modern formal sector within economies. The traditional sector, characterised by labour-intensive activities, coexists with the capital-intensive modern sector, with the latter absorbing surplus labour as the economy develops. Boeke's theory has laid the groundwork for subsequent research, including the works of Lewis (1954) and Harris and Todaro (1970), and has greatly contributed to the understanding of the formal-informal dichotomy in economies (Clement, 2015).

The structural theory, also known as Neo-Marxist theory, posits that the informal economy is a result of deliberate strategies employed by capitalists to reduce costs, enhance competition, and weaken union control. This perspective, rooted in Karl Marx's (1848) ideology, views globalisation as a primary driver of informality, leading to increased subcontracting activities and reinforcing informal economic practices. Unlike modernisation theory, neo-Marxists argued for a strong link between formal and informal economies, considering both as serving the interests of capitalist development (Huang et al., 2020). This theory challenges traditional views by suggesting that industrialisation and modernisation may increase informality.

Neo-liberalism, championed by proponents like De Soto (1989), attributes informality to a response against excessive state regulation. Entrepreneurs opt for informal operations to evade the costs and bureaucratic hurdles of formal registration. This ideology views informality as a form of resistance against government overreach, allowing for greater entrepreneurial freedom. Embraced during global economic crises, neo-liberalism advocates deregulation, free-market development, and reduced social expenditure to foster formal employment and economic growth (Yusuff, 2011)

The review above shows that there has been a theoretical gap with no recent theory to the best of the researcher’s knowledge explaining the size of informality in terms of newly postulated determinants being put forth. Those that have so far been identified by the existing theories could not have been adequate, not to talk of being exhaustive, particularly as the inclusion of all factors they predict to be relevant in a regression equation for the size of the informal sector still often leave a lot of variation in the size of informality in form of unexplained residuals that seek further postulated determinants to explain them. Therefore, there must still have existed a sizeable number of factors that are unidentified by these theories. Thus, there is a need for new theoretical prepositions on factors that explain the size of the informal sector in an economy.

2.2 Empirical Review

The empirical studies reviewed offer valuable insights into the multifaceted nature of the informal economy across various regions and contexts. This review is classified into empirical studies that are on non-African countries and those that are on African countries.

In the context of empirical studies on non-African countries, Krakowski (2005) highlighted the significance of labour regulation and good governance in reducing informality across 106 countries using data covering 199 to 2000. He applied the log-odd function estimation technique and found out that different levels of tax rates are compatible with similar sizes of the informal economy while the degree of labour regulation and indicator of good governance are both important in explaining the informal economy around the world. Chong and Gradstein (2007) studied the determinants of the size of the informal sector across countries for the period 1990 to 2000 using Ordinary Least Squares (OLS) and instrumental variables estimation approaches. They found that income inequality increases informality, while institutional quality decreases it. Paula and Scheinkman (2007) investigated the determinants of informality in Brazil in 2003 using probit estimation techniques for analysis and found that various measures of the formality of suppliers and purchasers as well as value-added tax are correlated with the formality of firms in Brazil. Loayza (2007) studied the causes and consequences of informality in Peru by employing data from 1985 to 2004 and employing the OLS method with Robust Standard Errors. The result showed regulatory freedom and education as factors reducing informality in Peru.

Also, Dogrul (2012) examined the factors that determined the informal sector employment in Turkey between 2001 and 2006 by adopting a multinomial logit model and noted significant gender differences in Turkey's informal employment as a determinant of the size of the informal sector. Berdiev et al. (2018) examined the effect of economic freedom on the shadow economy for 100 countries covering the period 2000-2015 using two-way country and time-fixed effects elasticity estimates. The result showed that economic freedom, particularly regulation freedom, reduces the shadow economy. Ilyas et al. (2020) investigated the important determinants of informal enterprises in Pakistan by surveying markets in Lahore while employing Bartlett's method to test the null hypothesis and pointed to high tax rates and regulatory burdens as drivers of informality in Pakistan. Huang et al. (2020) used the random effects method to examine the causes of the urban informal sector in China and indicated that economic structure and globalisation impact China's urban informal sector. Syed et al. (2021) assessed the effect of digital financial fit on the size of the shadow economy for 2004-2018 using CUP-FM and CUP-BC estimation and found that digital finance reduces the shadow economy but increases financial instability in Asia. Schneider (2022) analysed the informal economy's size across 157 countries from 1991 to 2017 and linked trade openness, GDP per capita, and economic freedom to reduced informality globally. Finally, Zhanabekov (2022) applied the Bayesian Model Averaging (BMA) to investigate the determinants of informality across developed, developing and post-Soviet Union countries between 2006 and 2014 and demonstrated that economic growth, trade openness, and institutional quality decreased informality across various countries.

In the case of the studies on African countries Folawewo (2006) examined the determinants of informal labour demand in Nigeria in the second half of 2003 using the OLS and Instrumental Variable IV techniques and identified that in Southwestern Nigeria, factors like the establishment year, investment, profit levels, and ownership nature affect informal labour demand, with worker characteristics such as age, education, gender, experience, and tenure also playing crucial roles. Traore (2012) investigated the determinants of informal sector activity with a special focus on informal self-employment in Burkina Faso in 2007. He applied the OLS and Instrumental Variable IV techniques to estimate and found that in Burkina Faso, education reduces the likelihood of self-employment, while urban living and single status also decrease it, whereas access to credit and engagement in non-tradable services increase self-employment. Aikaeli and Mkenda (2014) analysed the determinants of informal employment in Tanzania by applying a multivariate logit regression model to 2013 survey data and noted that income and capital constraints are critical for micro and small entrepreneurs, with higher education and income-reducing informal employment, while women are more likely to be informally employed.

In a broader approach, Ouedraogo (2017) evaluated the relationship between governance, corruption, and the size of the informal economy in Sub-Sahara Africa between 2000 and 2010. He employed the OLS method with a stepwise estimation procedure, to estimate four models and the result revealed that corruption and per capita income were positively associated with informality, whereas good governance and low unemployment reduced it, while economic freedom can have a mixed effect on it. Ladan and Williams (2019) employed a descriptive method of analysis to analyse the data collected by a survey in 2012 in Zamfara, the result indicated that informality arises from both exclusion from and voluntary exit from the formal sector. Keneck-Massil and Noah (2019) examined the relationship between education and the informal economy in Africa for the period 2000-2015 using the Two-Stage Least Squares (2SLS) estimation method and observed that vocational education financial development and robust institutional frameworks reduce informality, while secondary education social globalisation and higher tax rates increase it.

Njangang et al. (2020) investigated the relationship between financial development and the size of the informal economy for the period 1991-2015. The paper adopted different estimation techniques like OLS, fixed effects, and system generalised method of moments and reported that in Sub-Saharan Africa, financial development reduces informality, with GDP per capita growth, human capital, domestic investment, and savings also negatively impacting it, while population growth and remittances increase it. Anthelme (2021) employed the maximum likelihood technique and found that in Ivory Coast for the period 1991-2018, government spending, inflation, and trade openness reduce informality, while higher taxes and unemployment rates increase it. Esaku (2021) examined the relationship between the shadow economy and trade openness in Uganda for the period 1991-2015. The paper adopted the Auto Regressive Distributed Lag (ARDL) model and found that trade openness, GDP per capita, financial development, and institutional quality reduce informality, while the tax burden increases it. Serawitu (2021) investigated the intensity of informal competition among firms in Ethiopia. To do this, he employed the estimated linear probability model (LPM) and identified that

corruption, high taxes, and limited credit access in Ethiopia in 2015 increased informal competition, while larger firm size reduces it. Ejiogu et al. (2022) noted that across 47 African countries, corruption, political instability, crime, and access to finance influence informality, with better access to finance and infrastructure increasing the informal sector size, while crime and political instability decrease it. Ningaye and Ketu (2023) employed OLS, Fixed Effects and System GMM estimation techniques on data covering the period of 2003 to 2018 and observed that across 42 African countries covered, improved infrastructure reduces informality. Finally, Nkemgha (2023) using similar techniques like OLS, fixed effects, instrumental variables, and the system GMM, investigated the impact of international trade on the informal economy in 24 countries of Sub-Sahara Africa from 2000 to 2020 and found that international trade reduced informality.

Collectively, these studies reveal that so far, the same variables or factors are recycled from study to study. Therefore, there is a need to explore alternative determinants because what has been identified does not explain in totality the multi-faceted nature of informality in an economy.

Thus, there is a need for an empirical study that will further explain informality in terms of new factors, which is what the present study seeks to explore.

3.0 Methodology

3.1 Theoretical Framework

The theoretical framework for this paper, which was non-existent before now, to the best of the study's knowledge, is based on the theory of ethnic fractionality put forth by this study proposes a link between the level of ethnic diversity within an economy (ethnic fractionalisation) and the size of its informal sector. The core argument is that a more fragmented society with distinct ethnic groups can lead to a larger informal economy. This is because it is difficult to move from both formal and informal sectors in one's ethnic domain into the formal sector in other ethnic domains but relatively easy to move from both formal and informal sectors in one's ethnic domain to the informal sector in other ethnic domain, with this being the source of the asymmetry. As movement across one's ethnic boundary is a normal feature in an economy, this means that such movements are more prone to swell the number of economic agents joining the informal sector the greater the ethnic diversity that exists and, hence the greater the incidence of ethnic fractionalisation. In the limit whereby there is just a single ethnicity in an economy, this asymmetric effect will cease to exist. This phenomenon is peculiar to individuals, as it also pertains to businesses, particularly small businesses, seeking to relocate from their existing ethnic base to another ethnic base. Such a small business will be more received in the informal sector of the new ethnic base than the formal sector, thereby acting as an incentive for it to go informal (as opposed to formal) there where fewer obstacles and resistance exist.

This theory predicts that as the number of distinct ethnic groups increases (higher ethnic fractionalisation), the size of the informal sector will also grow due to the limited opportunities for formal sector mobility across ethnic divides

3.2 Model Specification

The theoretical framework provided in the previous section sheds light on the link between the size of the informal sector and ethnic fractionalisation. Specifically, this theoretical framework predicts that ethnic fractionalisation should have a positive effect on the size of the informal sector. Also, based on the espoused theory of dualism as well as several other theories that are not espoused but merely reviewed in Section 2, many other factors affect the size of informality, as previously confirmed by many existing empirical studies and these are included to serve as the control variables in the model specified for estimation in this paper. Four of such factors are considered important and, hence, singled out as the control variables. These are Economic development, unemployment, educational attainment and institution.

Drawing from economic theory and empirical research, it is hypothesised that economic development will have a negative effect on informality. This is because the modernisation theory by Rostow (1960) posits that economic development reduces the size of the informal sector. As economies become more developed, informality decreases. Empirical studies have tested this theory but have yielded mixed results. Chong and Gradstein (2007), Berdiev et al. (2018), Keneck-Massil and Noah (2019), Njangang et al. (2020), and Bulut (2023) found that economic development negatively impacts informality, while Huang et al. (2020) found a positive effect. Unemployment is considered a determinant of informality because, according to modernisation theory, the formal sector's inability to create sufficient jobs leads to excess labour seeking survival in the informal sector, resulting in a positive relationship between unemployment and informality. Empirical studies generally support this positive effect of unemployment on informality. For example, Dougherty and Escobar (2013), Igudia et al. (2016), Huang et al. (2020), Anthelme (2021), and Gezer (2022) found a positive relationship, while Medina et al. (2017) reported a negative effect. For this paper, it is expected that unemployment will exert a positive influence on informality. Additionally, educational attainment reflects the skills of the labour force. Higher-skilled workers are more likely to secure formal sector jobs, whereas lower-skilled workers are more likely to engage in informal employment. Empirical studies by Loayza (2007), Dabla-Norris et al. (2008), Berdiev et al. (2018), and Dougherty and Escobar (2013) have found that education negatively affects informality. Based on this logic and empirical evidence, it is expected that educational attainment will have a negative impact on informality. Governance indicators consist of the traditions and institutions by which authority in a country is exercised. This is also expected to play a crucial role, with stronger governance associated with reduced informality. Therefore, this paper expects a negative effect of the composite governance indicator on informality.

Based on the theoretical framework and variable justification above, the model to be estimated is thus specified as:

$$IF_{it} = \beta_0 + \sum_{j=1}^{p} \sigma\, IF_{it} + \sum_{j=0}^{q1} \alpha_{it}\, ETFR_{it-j} + \sum_{j=0}^{q2} \gamma_{it}\, EDEV_{it-j} + \sum_{j=0}^{q3} \delta_{it}\, \mathrm{UNE}_{\,it-j} + \sum_{j=0}^{q4} \mu_{it}\, EDU_{it-j} + \sum_{j=0}^{q5} \alpha_{it}\, WGIC_{it-j} + U_{it} \quad (1)$$

$$i = 1, 2, - - -, N; t = 1, 2, - - -, T.$$

where: IF = Size of Informal Sector (DGE, MIMIC); ETFR = Ethnic Fractionalisation; EDEV = Economic Development; UNE = Unemployment; EDU = Educational Attainment; WGIC = World Governance Indicator Composite. The error term U_{it} comprises of individual country effects μ_i and the Gauss Markov error term ε_{it}.

The baseline model in equation (1) needs to be re-parameterize so that the short-run and long-run dynamic panel model structure can be obtained as follows:

$$IF_{it} = (\beta_0 + \Psi_i\, IF_{it-1} + \beta_{1i} ETFR_{it-1} + \beta_{2i} EDEV_{it-1} + \beta_{3i}\mathrm{UNE}_{it-1} + \beta_{4i} EDU_{it-1} + \beta_{5i} WGIC_{it-1}) + \sum_{j=0}^{p-1} \eta_{ij}^{*}\, \Delta IF_{it-j} + \sum_{j=0}^{q_1-1} \sigma_{ij}^{*}\, \Delta ETFR_{it-j} + \sum_{j=0}^{q_2-1} \gamma_{ij}^{*}\, \Delta EDEV_{it-j} + \sum_{j=0}^{q_3-1} \delta_{ij}^{*}\, \Delta\mathrm{UNE}_{it-j} + \sum_{j=0}^{q_4-1} \omega_{ij}^{*}\, \Delta EDU_{it-j} + \sum_{j=0}^{q_5-1} \varphi_{ij}^{*}\, \Delta WGIC_{it-j} + \mu_i + \varepsilon_{it} \quad (2)$$

The short-run slope coefficients for each cross-section are equal to σ_{ij}^{*}, γ_{ij}^{*} … and φ_{ij}^{*} for ETFR, EDEV, UNE, EDU, and WGIC, respectively. By assumption, ΔIF_{it-j}= $\Delta ETFR_{it-j}$= $\Delta EDEV_{it-j}$= $\Delta\mathrm{UNE}_{it-j}$= ΔEDU_{it-j}= $\Delta WGIC_{it-j}$= 0 in the long run. Therefore, the long-run coefficients for each of the variables are obtained as $-\frac{\beta_{1i}}{\Psi_i}$, $-\frac{\beta_{2i}}{\Psi_i}$ … and $-\frac{\beta_{5i}}{\Psi_i}$ for ETFR, EDEV, UNE, EDU, and WGIC, respectively. The error correction equivalent of equation (2) is given by:

$$\alpha_i \theta_{it-1} + \sum_{j=0}^{p-1} \eta_{ij}^{*}\, \Delta IF_{it-j} + \sum_{j=0}^{q_1-1} \sigma_{ij}^{*}\, \Delta ETFR_{it-j} + \sum_{j=0}^{q_2-1} \gamma_{ij}^{*}\, \Delta EDEV_{it-j} + \sum_{j=0}^{q_3-1} \delta_{ij}^{*}\, \Delta\mathrm{UNE}_{it-j} + \sum_{j=0}^{q_4-1} \omega_{ij}^{*}\, \Delta EDU_{it-j} + \sum_{j=0}^{q_5-1} \varphi_{ij}^{*}\, \Delta WGIC_{it-j} + \mu_i + \varepsilon_{it} \quad (3)$$

where θ_{it} is the linear error correction term (ECT), which is given as:

$$\theta_{it} = \Delta IF_{it} - \phi_0 - \phi_{1i} \Delta ETFR_{it-j} - \phi_{2i} \Delta EDEV_{it-j} - \phi_{3i} \Delta\mathrm{UNE}_{it-j} - \phi_{4i} \Delta EDU_{it-j} - \phi_{5i} \Delta WGIC_{it-j} \quad (4)$$

The coefficients of the ECT α_i are the error-correcting speed of adjustment while $-\phi_0 - \phi_{1i}, -\phi_{2i} \ldots and - \phi_{5i}$ represent the long-run parameters of the model. When α_i is significant and negative and less than one in absolute value, it implies that co-integration exists (Kilishi et al., 2021).

3.3 Methods of Estimation

To determine the effect of ethnic fractionalisation on the size of the informal sector, both descriptive and inferential analyses were carried out. The descriptive analysis involves the use of summary statistics to describe each variable, while the inferential analysis involves the use of the Panel Auto Regressive Distributed Lag (PARDL) technique to conclude the study. The choice of the PARDL is based on the panel unit root test, which shows variables to be both stationary and with unit root, and also cointegrated in the long run.

3.4 Data

The data used for this study are panel ones spanning 1993 to 2021 for 41 Sub-Saharan African countries. These are Angola, Benin, Botswana, Burkina Faso, Burundi, Cape Verde, Cameroon, Central African Republic, Chad, Comoros, Congo (Dem. Rep.), Congo (Rep.), Cote d'Ivoire, Equatorial Guinea, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Tanzania, Togo, Uganda, Zambia, and Zimbabwe. The start of the time frame was chosen because there were no data on the dependent variables before 1993 and 2021 is the last year for which data are available. Also, out of the 46 Sub-Saharan African countries, 40 countries were covered because they are the only Sub-Saharan African countries for which statistics are available. The definition of the variables employed, their sources and how they were measured are described below.

3.5 Measurement of Variables

Informality is measured using two different variants (viz: Dynamic General Equilibrium and Multiple Indicator Multiple Cause). This is because, among the measures of informality, these two are the ones that measure the size of informality for the whole economy. The Dynamic General Equilibrium (DGE) DGE is a form of indirect measure of informal sector size which considers how optimising households will allocate labour between formal and informal economies in each period and how the allocation changes over time. The already generated data, which are expressed as a percentage of GDP, are available for 158 economies (out of which 40 are SSA economies) throughout 1950 - 2020 from the data source, which is Elgin et al. (2021). The MIMIC is a form of structural model that similarly uses an indirect approach and takes into account multiple possible causes of informal activities and also captures multiple outcome indicators in measuring the size of the informal sector. Like the DGE, this too is similarly measured as estimates of informal output as a percentage of official GDP and already generated data are available for 160 economies (out of which 40 are SSA economies) between 1993 and 2020 from the same data source as those of Elgin et al. (2021).

Ethnic Fractionalisation is measured by the Historical Index of Ethnic Fractionalisation (HIEF) and it corresponds to the probability that two randomly drawn individuals within a country are not from the same ethnic group. The HIEF dataset contains an ethnic fractionalisation index for 165 countries across all continents (out of which 36 are for SSA economies). The dataset covers the annual period from 1945-2013. The new or post-

2013 dataset is a natural extension of previous ethnic fractionalisation indices and it allows its users to compare developments in ethnic fractionalisation over time. Theoretically, fractionalisation indices range from 0, when all individuals are members of the same group to 1, when each individual belongs to his or her group (Dražanová, 2020). This data was sourced from Harvard Dataverse online and the missing years are extrapolated for using the Stata software.

Economic Development (EDEV) is proxied by the GDP per capita at 2017 PPP constant US $, and it is sourced from the World Bank's World Development Indicator (WDI) online. The unemployment rate (UNE) is measured as unemployment about the total labour force, sourced from the World Bank's WDI online. Educational attainment is proxied by the literacy rate sourced from the World Bank's WDI online. The governance indicators composite (WGIC) is measured as an average of the individual governance indicators for six dimensions of governance (viz: Voice and Accountability; Political Stability and Absence of Violence/Terrorism; Government Effectiveness; Regulatory Quality; Rule of Law; Control of Corruption). The data is sourced from the World Bank's World Governance Indicators (WGI) online. It ranges from -2.5 to 2.5, with higher values indicating better outcomes.

4.0 Research Results

4.1 Descriptive Analysis

This section presents and evaluates the descriptive statistics for each of the variables employed in the study. The descriptive statistics are summarised in Table 1, which shows the mean, minimum value, maximum value and standard deviation.

Table 1: Summary Statistic Table

Variable and Description	Mean	Std. Dev.	Min	Max
DGE – Informality size as % of official GDP	39.158	8.564	20.3	63.9
MIMIC - Informality size as % of official GDP	40.784	7.858	20.9	63.3
ETFR - Historical Index of Ethnic Fractionalisation	0.668	0.221	0.044	0.89
EDEV - GDP per capita, PPP (2017 constant international $)	7.926	0.852	6.80	9.67
UNE – Unemployment as % of total labour force	7.62	6.165	1.294	21.155
EDU - Literacy rate, % of people aged 15 and above	38.768	22.668	6.449	86.8
WGIC – Composite Governance Indicator - in units scaled between -2.5 to +2.5	0	1	-1.645	2.201

Source: Author's computation (2024).
Explanatory Notes: Std Dev = standard deviation, Coeff of var = coefficient of variation, Min = minimum, max = maximum

Table 1 reveals the mean and standard deviation of dynamic general equilibrium or DGE to be 39.158% and 8.564%, with a minimum value of 20.3% and a maximum value of 63.9%. The mean and standard deviation of multiple-indicator multiple-cause or MIMIC were 40.784% and 7.858% respectively, with a minimum value of 20.9% and a maximum value of 63.3%, while the mean and standard deviation of ethnic fractionalisation or ETFR were 0.668 and 0.221, with a minimum value of 0.044 and a maximum value of 0.89. Also, the mean and standard deviation of per capita income at 2017 PPP constant US $ or EDEV were 4132.615 and 4245.873 respectively, with a minimum value of 796.109 and a maximum value of 15907.9. The mean and standard deviation of Unemployment as % of the total labour force or UNE were 7.62 and 6.165 respectively, with a minimum value of 1.294 and a maximum of 21.155. As for the Literacy rate as % of people aged 15 and above or EDU, it had a mean and a standard deviation of 38.768 and 22.668 respectively, with a minimum value of 6.449 and a maximum of 86.8. finally, the mean and standard deviation of the composite world governance indicator or WGIC were 0 and 1 respectively, with a minimum of value of -1.645 and a maximum value of 2.208.

4.2 Unit Root and Cointegration Tests

To avoid the consequence of having spurious regressions, the panel data unit root test is carried out in this study to examine the stationary nature of each of the variables used in the models. The Im, Pesaran and Shin (2003) unit root test is employed, due to its reliability and common use in literature. The significance level adopted in evaluating the results is 5 per cent while the decision rule is to reject the null hypothesis that a variable has a unit root (i.e., the variable is a non-stationary series) if the p-value is less than or equal to 5 per cent significance level and accept the null hypothesis if otherwise. The test results for unit root are presented in Table 2.

Table 2: Panel Unit Root Test Results

Variables	z-statistic	p-value	Order of Integration	Conclusion
DGE	3.6519	1.000	I(1)	Unit Root
	-9.2133	0.000	I(0)	
MIMIC	3.1618	0.999	I(1)	Unit Root
	-14.2262	0.000	I(0)	
ETFR	-11.5156	0.000	I(0)	Stationary
	NA	NA	NA	
EDEV	6.2972	1.000	I(1)	Unit Root
	-11.8646	0.000	I(0)	
EDU	3.4017	1.000	I(1)	Unit Root
	-4.5005	0.000	I(0)	
UNE	-2.9939	0.001	I(0)	Stationary
	NA	NA	NA	
WGIC	2.4157	0.9921	I(1)	Unit Root
	-13.9947	0.000	I(0)	

Source: Author's computation (2024).

Explanatory Notes: the following are the meanings of the acronyms: DGE = Dynamic General Equilibrium, MIMIC = Multiple Indicator Multiple Cause, ETFR = Ethnic Fractionalisation, YNQ = Income Inequality, INF = Inflation, EDU = Educational Attainment, UNE = Unemployment and WGIC = WGI Composite. A coefficient is deemed statistically significant if its p-value is less than or equal to

the 0.05 significance level employed in the study, which implies that the null hypothesis of the existence of unit root is rejected and, hence, the variable is regarded as being stationary while the converse holds if the p-value exceeds 0.05 significance level. Finally, "NA" stands for "not applicable".

The results of Table 2 reveal that 4 out of the 8 variables (viz: ETFR, YNQ, INF and UNE) are stationary at level and the remaining 4 variables (viz: DGE, MIMIC, EDU and WGIC) are stationary only after being first-differenced. These conclusions are based on the p-values of the z-statistics that are all less than or equal to 5 per cent for the unit root test conducted on the level form of those 4 variables that are adjudged to be stationary and for the unit root test conducted on the first difference form (but not the level form) of those 4 variables that are found to have a unit root. Specifically, as all the 2 dependent variables (viz: DGE and MIMIC) dependent variables have unit root and their respective explanatory variables are a combination of stationary and unit root series, this implies that using the OLS approach as the estimation technique is prone to produce spurious regression results and also that a follow-up cointegration test is required to confirm whether a long run relationship exists among the variables featuring in each equation.

In the same vein, a cointegration test is carried out since the results from the unit root test reveal that the applicable dependent variables have unit roots and some of the explanatory variables are stationary at levels while others are stationary at first difference. As a result, it is the Kao test methodology that is considered appropriate and, hence, adopted for carrying out the test. A 5 per cent significance level of the test statistic is adopted in evaluating the results. The null hypothesis of the Kao residual cointegration test is that there is no cointegration, and the decision rule is that, if the p-value of the test statistic is greater than the chosen critical 5 per cent significance level, then, the null hypothesis is accepted so that it is concluded that there is no cointegration and, if otherwise, the null hypothesis is rejected. The test results for cointegration are presented in Table 3.

Table 3: Result of Kao Panel Cointegration Test

Model	t-Statistics	p-value	Conclusion on H_0
Model 1, with DGE as the measure of informality	-5.909	0.000	Rejected
Model 2, with MIMIC as the measure of informality	-2.794	0.003	Rejected

Source: Author's computation (2024).

Explanatory Notes: The decision rule is that a test statistic is statistically significant and, hence, Ho of lack of cointegration is rejected if the p-value is less than 5%, thereby confirming the existence of cointegration. If otherwise, Ho is accepted, which means that there is an absence of cointegration.

It can be observed from Table 3 that the t-statistic is statistically significant in each of the models, as the p-values are all less than 5%. Following the decision rule, the null hypothesis is to be rejected and, hence, it is concluded that a long-run relationship exists among the series featuring in each of the models. This implies that the panel ARDL estimation technique can be employed to derive not only the short-run but also the long-run estimates of the parameters of the models, although it is only the long-run estimates that are of interest in the paper and which, therefore, are the only ones reported.

4.4 Choice among the Variants of Panel ARDL Estimation Method to Adopt

Since the outcomes of the prior unit root and cointegration tests carried out are supportive of the adoption of Panel ARDL estimation methods, tests are first conducted to ascertain the relative appropriateness of the three alternative variants of the Panel ARDL method of estimation too. These three alternative variants are the Mean Group (MG), Pooled Mean Group (PMG) and Dynamic Fixed Effects (DFE). The Hausman test is employed to determine the relative suitability among them. If the p-value is greater than 0.05, indicating no significant difference, the PMG estimator is preferred over MG or DFE. If the p-value is greater than 0.05 it indicates that DFE is superior to MG. The estimator (MG, PMG, or DFE) that is favoured at least twice in three comparisons (MG vs. PMG, MG vs. DFE, PMG vs. DFE) is chosen as the most suitable. The results of the tests are presented in Table 4.

Table 4: Results of Hausman Test Guiding the Choice among the PMG, MG and DFE Variants of Panel ARDL Estimation Method

MODEL	MG & PMG			MG & DFE			PMG & DFE			Conclusion
	z-stat	p-value	Decision	z-stat	p-value	Decision	z-stat	p-value	Decision	
Model 1										
Model 1, with DGE as a Measure of informality	8.62	0.071	PMG	20.33	0.000	DFE	1.49	0.828	PMG	PMG
Model 2, with MIMIC as a measure of informality	2.34	0.673	PMG	0.24	0.993	MG	1.59	0.673	PMG	PMG

Source: Author's computation (2024).
Explanatory Notes: Models 1 and 2 depict the models for each measure of informality. The decision rule is that if the p-value of the Hausman test is greater than 0.05, it is concluded that the PMG estimator is superior to the MG; the DFE estimator is superior to the MG estimator; and the PMG estimator is superior to the DFE estimator.

In Table 4, based on the p-value of the computed Hausman test statistic and the adoption of a 0.05 significance level, PMG is seen to be the best for estimating both models (viz: Model 1 and Model 2) since it is selected twice out of the three comparisons.

4.5 Estimates of the Panel Regression Equations

The PMG variants of Panel ARDL, which are applicable due to the result of the Hausman tests, are employed as the estimation method. The long-run estimates are presented for the two models in Table 5. Concerning all the model estimates, the coefficients, t-statistics and p-values are reported in the first, second, and third columns respectively. The coefficient is considered statistically significant if the corresponding p-value is less than 0.05 and statistically insignificant if the p-value is greater than 0.05 cut-off.

Table 5: Long-run Panel ARDL Estimates of the Regression Equations for DGE and MIMIC

Variable	DGE			MIMIC		
	Coefficient	z - stat	p-value	Coefficient	z - stat	p-value
ETFR	3.358	2.64	0.008	15.716	4.97	0.000
EDEV	-0.001	-4.21	0.000	-0.004	-8.93	0.000
UNE	0.254	3.79	0.000	0.326	4.43	0.005
EDU	-0.118	-4.71	0.000	-0.164	-5.93	0.000
WGIC	-3.978	-14.16	0.000	-4.967	-9.07	0.000
constant	20.361	6.73	0.000	10.254	16.47	0.002
ECT_{t-1}	-0.113	-6.75	0.000	-0.121	-3.51	0.000
R^2	0.289			0.283		
Adjusted R^2	0.285			0.280		
F Statistic	77.68	-	0.000	75.57	-	0.000

Source: Author's computation (2024).

The R^2 for the two models estimated are 0.289 and 0.283 respectively while the *F* statistics are 77.68 and 75.57 respectively with a p-value below the cut-off of 0.05 significant level. This shows that the models are of good fit and have high explanatory powers. Also, the coefficients of ECT_{t-1} in the models estimated were -0.113 and -0.121 respectively and are negative and statistically significant at 5% as their p-value is less than 0.05 significant level. This shows the speed of adjustment from short-run deviations to long-run equilibrium.

Coming to the evaluation of the performance of each explanatory variable, the coefficients of ETFR are positive and statistically significant in both models, with a coefficient of 3.358 and 15.716 respectively and p-values of less than 0.005. Also, the coefficients of EDEV are negative and statistically significant in both models, with a coefficient of -0.001 and -0.004 respectively and p-values of less than 0.005. As for UNE, the coefficients are positive and statistically significant in both models with a coefficient of 0.254 and 0.326 respectively and p-values of less than 0.005. In the case of EDU, the coefficients are seen to be negative and statistically significant, in both the DGE and the MIMIC models, with coefficients of -0.118 and -0.164 and p-values of 0.000 and 0,000 respectively Finally, the coefficients of WGIC are negative and statistically significant, with a coefficients of -3.978 respectively and -4.967 and p-values of less 0.005.

5. Discussion of Result and Implication of Findings

The presented in the previous section provide evidence of the positive effects of ETFR on both the DGE and the MIMIC measure of informality. This is in line with the expectations of this paper. Thus, confirming that a higher degree of ethnic fractionalisation corresponds to a larger size of the informal sector. An increase in ethnic fractionalisation will increase the size of the informal sector using the DGE variant as a measure of informality. Similarly, an increase in ethnic fractionalisation will increase the size of informality according to the MIMIC variant of the informality measure. This validates the proposition that ethnic fractionalisation has a positive effect on informality.

Also, it shows evidence of a negative effect of EDEV on DGE and MIMIC variants of informality. This is in line with the expectation of this study. Also, previous studies, like Chong and Gradstein (2005) and Berdiev, Saunoris and Schneider (2018), similarly

reported negative effects of the level of economic development on informality. There is evidence of the positive effects of UNE on the two informality measures which is in line with the expectation of the paper and also similar to the result of positive effects of unemployment on informality that had hitherto been reported in several studies that include Celik (2023) and Wondimu and Birru (2020), among others.

In the case of EDU, the result provides evidence of a negative effect of UNE on both variants of informality. This is in line with the expectations of the paper. Previous studies, like Dabla-Norris, Gradstein and Inchauste (2008) and Dougherty and Escobar (2013) also reported similar findings. Finally, the result provides evidence of a negative effect of WGIC on the two measures of informality. This is also in line with the expectations of this paper.

It can be observed from the results in Table 5 that the effect of the variables employed has more effect on the MIMIC measure of informality than the DGE measure. However, whether it is the DGE that is used as a measure for informality or it is MIMIC, the results appear similar, and this attests to the robustness of the proposition and empirical findings of this paper

6. Conclusion and Recommendations

The study investigated the effects of ethnic fractionalisation on informality in Sub-Saharan Africa using two measures of informality (viz: DGE and MIMIC). To achieve this objective, an informality equation was specified, based on the theory of economic dualism and a proposition that is first being put forward in this paper on the role of ethnic fractionalisation as the main theoretical framework. In addition to ethnic fractionalisation, four control variables (viz: Economic development, unemployment, educational attainment and the World governance indicators composite) were also included as explanatory variables in each estimated informality equation, while the two measures of informality mentioned earlier were used as two alternative dependent variables. The paper tested for the presence of unit root concerning each variable, using the Im et al. (2003) unit root test, and found the series to be a combination of I(0) and I(1). This necessitated the use of the Kao panel test for cointegration to examine the long-run relationship among the variables, with the outcome of the test indicating the presence of a long-run relationship among the variables in the two models estimated. The study adopted the Panel ARDL estimation technique, using data covering the period of 1990 to 2022 that were sourced online from Elgin et al. (2021), Harvard Dataverse, the World Bank's World Development Indicators (WDI) and the World Bank's Governance Indicators (WGI).

Based on the above approach, the study found that ethnic fractionalisation has a positive effect on the informal sector, irrespective of which of the two variants of informality measures were analysed. The positive effect of ethnic fractionalisation on the size of the informal sector confirms the theoretical proposition of this paper that ethnic fractionalisation would increase the informal size. Based on the strength and robustness of the evidence backing this conclusion, there is little or no doubt that this paper has contributed to the existing literature on the analysis of the informal economy by swelling

the list of empirically verified factors determining the size of the informal economy by this new factor.

Also, it was found that the control variables (economic development, unemployment, educational attainment and composite governance indicator) all have the expected effects on the size of the informal sector.

In line with the findings above, it is recommended that the authorities should formulate and implement policies that will reduce the role of ethnic fractionalisation, particularly by stemming any impediments that the degree of ethnic fractionalisation that exists to the flexibility of moving across ethnic divides in seeking formal employment and exploring business opportunities in the formal sector to make it similar to the same degree of cross-ethnic-divide flexibility that pertains to the informal sector. In addition, the authorities should strive to develop and implement policies that reduce the unemployment rate but promote economic development, and literacy rate and strengthen institutional quality.

References

Aikaeli, J., & Mkenda, B. K. (2014). Determinants of informal employment: A case of Tanzania's construction industry. *Botswana Journal of Economics*, *12*(2), 51-73.

Amin, A. (2004). Regions unbound: towards a new politics of place. *Geografiska annaler: series B, human geography*, *86*(1), 33-44.

Anthelme, A. N. (2021). Informal Economy in Sub-Saharan Africa: Main Drivers and Estimation of His Size from Ivory Coast.

Berdiev, A. N., Saunoris, J. W., & Schneider, F. (2018). Give me liberty, or I will produce underground: Effects of economic freedom on the shadow economy. *Southern Economic Journal*, *85*(2), 537-562.

Boeke, J. H. (1953). Economics and economic policy of dual societies, as exemplified by Indonesia. *(No Title)*.

Bulut, A. (2023). Informal Economy and Financial Development: The Role of Institutions. *Sosyoekonomi*, *31*(55), 195-209.

Chong, A., & Gradstein, M. (2007). Inequality and informality. *Journal of Public Economics*, *91*(1-2), 159-179.

Clement, C. (2015). *The formal-informal economy dualism in a retrospective of economic thought since the 1940s* (No. 43/2015)

Dabla-Norris, E., Gradstein, M., & Inchauste, G. (2008). What causes firms to hide output? The determinants of informality. *Journal of Development Economics*, *85*(1-2), 1-27.

de Paula, A., & Scheinkman, J. (2007). *The Informal Sector, Second Version*. Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

De Soto, H. (1989). *The other path* (p. 17133). New York: Harper & Row.

Doğrul, H. G. (2012). Determinants of formal and informal sector employment in the urban areas of Turkey. *International journal of social sciences and humanity studies*, *4*(2), 217-231.

Dougherty, S., & Escobar, O. (2013). The Determinants of Informality in Mexico's States. *OECD Economic Department Working Papers*, (1043), 01.

Dražanová, L. (2020). Introducing the Historical Index of Ethnic Fractionalization (HIEF) dataset: accounting for longitudinal changes in ethnic diversity. *Journal of open humanities data*, *6*.

Ejiogu, A., Okechukwu, O., Ejiogu, C., Owusu, A., & Adeola, O. (2022). Determinants of informal entrepreneurship in Africa. *International Journal of Entrepreneurship and Small Business*, *45*(1), 35-61.

Elgin, C., Kose, M. A., Ohnsorge, F., & Yu, S. (2021). Understanding informality.CAMA Working Paper N0. 76/2021. http://dx.doi.org/10.2139/ssrn.3916568

Esaku, S. (2021). The long-and short-run relationship between the shadow economy and trade openness in Uganda. *Cogent Economics & Finance*, *9*(1), 1-19.

Folawewo, A. O. (2006). Determinants of informal sector labour demand: An application of alternative methodological approaches to South Western States of Nigeria. *Applied Econometrics and International Development*, *6*(2).147-160

Gezer, M. A. (2022). The causality relationship between military expenditure and GDP in 12 NATO member countries is based on per capita values. *Sosyoekonomi*, *30*(52), 157-181.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American Economic Review*, *60*(1), 126-142.

Huang, G., Xue, D., & Wang, B. (2020). Integrating theories on informal economies: An examination of causes of urban informal economies in China. *Sustainability*, *12*(7), 2738.

Igudia, E., Ackrill, R., Coleman, S., & Dobson, C. (2016). Determinants of the informal economy of an emerging economy: a multiple indicator, multiple causes approach. *International Journal of Entrepreneurship and Small Business*, *28*(2-3), 154-177.

Ilyas, A., Hassan, R. A., Khan, W. A., bin Abang Abdurahman, A. Z., & Fahim, S. M. (2020). Important determinants of informal entrepreneurs. *Journal of Talent Development and Excellence*, *12*(3s), 2697-714.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, *115*(1), 53-74.

Keneck-Massil, J., & Noah, A. (2019). Shadow economy and educational systems in Africa. *Economics Bulletin*, *39*(2), 1467-1478.

Kilishi, A. A., Adebowale, H. A., & Oladipupo, S. A. (2021). The nexus between economic institutions and unemployment: evidence from sub-Saharan Africa. *International Journal of Development Issues*, *20*(1), 74-94.

Krakowski, M. (2005). Determinants of the informal economy: the importance of regional factors. *Available at SSRN 704267*.

Ladan, U., & Williams, C. C. (2019). Evaluating theorizations of informal sector entrepreneurship: Some lessons from Zamfara, Nigeria. *Journal of Developmental Entrepreneurship*, *24*(04), 1-18.

Lewis, W.A. (1954). Economic Development with Unlimited Supplies of Labour. *Manchester School Of Economics and Social Studies*, 22(2), 139-181.

Loayza, N. (2007). The causes and consequences of informality in Peru. *Serie de Documento de Trabajo*, *18*, 1-22.

Loayza, N. (2018). Informality: why is it so widespread and how can it be reduced? *World Bank Research and Policy Briefs*, (133110).

Medina, L., & Schneider, F. (2017). *Shadow Economies around the World: New Results for 158 Countries over 1991-2015* (No. 6430). CESifo.

Nageri, K. I., & Gunu, U. (2020). Corruption and ease of doing business: Evidence from ECOWAS. *Acta Universitatis Sapientiae, Economics and Business*, *8*(1), 19-37.

Ningaye, P., & Ketu, I. (2023). Does infrastructure development matter for the shadow economy in African countries? *International Review of Applied Economics*, *37*(3), 290-310.

Njangang, H., Nembot, L. N., & Ngameni, J. P. (2020). Does financial development reduce the size of the informal economy in sub- Saharan African countries? *African Development Review*, *32*(3), 375-391.

Nkemgha, G. Z. (2023). In the dark or out of the shadows? Trade liberalization and underground economies in Sub-Saharan Africa. *Journal of Economic Integration*, *38*(1), 147-170.

Ouédraogo, I. M. (2017). Governance, corruption, and the informal economy. *Modern Economy*, *8*(02), 256-271.

Portes, A. (1983). The informal sector: Definition, controversy, and relation to national development. *Review (Fernand Braudel Center)*, *7*(1), 151-174.

Rostow, W. W. (1960). The problem of achieving and maintaining a high rate of economic growth: A historian's view. *The American Economic Review*, *50*(2), 106-118.

Schneider, F. (2022). Estimation of Informal Economy: Figures for Developed and underdeveloped countries around the world. *Revista de economía mundial*, (60), 41-65.

Schneider, F., & Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, *38*(1), 77-114.

Serawitu, A. (2021). Determinants of Informal Competition Among Firms in Ethiopia. *International Journal of Health Economics and Policy*, *6*(4), 127-132.

Syed, A. A., Ahmed, F., Kamal, M. A., & Trinidad Segovia, J. E. (2021). Assessing the role of digital finance on shadow economy and financial instability: An empirical analysis of selected South Asian countries. *Mathematics*, *9*(23), 3018.

Traoré, J. A. (2012). *Revisiting the determinants of the informal sector in Burkina Faso* (No. 49006). University Library of Munich, Germany.

Yusuff, O. S. (2011). A theoretical analysis of the concept of informal economy and informality in developing countries. *European Journal of Social Sciences*, *20*(4), 624-636.

Zhanabekov, S. (2022). Robust determinants of the shadow economy. *Bulletin of Economic Research*, *74*(4), 1017-1052.