Financial Inclusion, Institutional Quality and Economic Growth in Sub-Saharan African Countries

Salam S. Mohammed , Golpet M. Kassem , Muhammad Ali

Under guidance of: Prof. Thirumulanathan D

Prepared by: Aditya Singh Chouhan, Shrikant Dethe

Importance of Financial Inclusion and Institutional quality for economic growth in SSA (Sub-Saharan Africa)

1) Financial Inclusion:
   - Access to financial Services, Economic participation,
   - Poverty Alleviation, Inclusive Growth

2) Institutional Quality:
   - Governance and Rule of Law, Regulatory Environment,
   - Investor Confidence, Social and Economic Development

In SSA, enhancing financial inclusion and strengthening institutional quality are critical for unlocking the region's economic potential. These efforts can lead to higher productivity, job creation, resilience against economic shocks, and improved living standards for the population, thus promoting long-term prosperity.

Hypothesis to be tested

Hypothesis 1: Financial inclusion positively impacts SSA economic growth.

Hypothesis 2: Strong institutional quality correlates with economic growth in SSA.

These hypotheses guide our exploration of the relationships between financial inclusion, institutions, and economic performance in Sub-Saharan Africa.

2 - equations

$$\begin{aligned}\mathrm{EG}_{it} &= \beta_0 + \beta_1\mathrm{EG}_{i,t-1} + \beta_2\mathrm{FI}_{it} + \beta_3\mathrm{IQ}_{it} \\ &\quad + \beta_4(\mathrm{FI}_{it} \star \mathrm{IQ}_{it}) + \beta_i X_{it} + \mu_i + \Omega_{it}\end{aligned} \tag{1}$$

where $\mathrm{EG}_{i,t-1}$ denotes the dynamic component of the relationship.

Now, By adding the control variables, equation (1) becomes:

$$\begin{aligned}\mathrm{EG}_{it} &= \beta_0 + \beta_1\mathrm{EG}_{i,t-1} + \beta_2\mathrm{FI}_{it} + \beta_3\mathrm{IQ}_{it} \\ &\quad + \beta_4(\mathrm{FI}_{it} \star \mathrm{IQ}_{it}) + \beta_5\mathrm{IFR}_{it} + \beta_6\mathrm{TOP}_{it} \\ &\quad + \beta_7\mathrm{UER}_{it} + \beta_8\mathrm{Inv}_{it} + \beta_9\mathrm{LIL}_{it} + \beta_{10}\mathrm{TNRR}_{it} \\ &\quad + \mu_i + \Omega_{it}\end{aligned} \tag{2}$$

Where:

- EG_{it}: Economic growth at time t in country i
- FI_{it}: Financial inclusion indicators in country i at time t
- IQ_{it}: Institutional quality in country i at time t
- X_{it}: Control variables in country i at time t
- μ_i: Unobserved country-specific effects
- Ω_{it}: Error term
- IFR_{it}: Inflation rate in country i at time t
- TOP_{it}: Trade openness in country i at time t
- UER_{it}: Unemployment rate in country i at time t
- Inv_{it}: Investment expenditure in country i at time t
- LIL_{it}: Literacy level in country i at time t
- TNRR_{it}: Total natural resources rent in country i at time t

Data and Methodology

- World Development Indicators (WDI) database
- International Monetary Fund (IMF) Financial Access Survey (FAS)
- Over the period of 2004–2020
- 20 Sub-Saharan African Countries (at least four countries each from the sub-regions were selected)
- Principal Component Analysis
- Unit Root test
- Two-step system Generalized Method of Moments (SysGMM) approach introduced by Blundell & Bond (1998)

Dealing with Missing Values

- Forward filling and Backward filling
  - Used for variables under Financial Inclusion, Institutional Quality
- Mean value imputation
  - Used in Control Variables like Trade Openness (TOP), Investment Expenditure (INV)
- 2 countries (Mozambique, Nigeria) whole dataset missing for TOP, INV
  - Countries similar in nature – data was used: Angola and Cameroon respectively
- Dropped variables:
  - Literacy level (LIL), a control variable was dropped due to high missing data
  - 2 measures of Institutional Quality – Human Rights Protection, Civil liberties were dropped. No data available and covered by other measures

Understanding Principal Component Analysis (PCA)

- It is a statistical technique used to reduce the dimensionality of data while retaining most of its variability. It transforms a set of correlated variables into a new set of uncorrelated variables called principal components.
- The primary goal of PCA is to identify patterns in data, reduce redundancy, and simplify the dataset's complexity, making it easier to explore and analyze.
- Process:

Step 1: Standardize the data (mean = 0, standard deviation = 1) to ensure each variable contributes equally.

Step 2: Compute the covariance matrix or correlation matrix of the standardized data.

Step 3: Calculate the eigenvectors and eigenvalues of the covariance/correlation matrix.

Step 4: Sort the eigenvalues in descending order to identify the principal components.

Step 5: Select the top k eigenvectors corresponding to the largest eigenvalues to form the new feature subspace (where k is the desired number of dimensions).

PCA in our study

1) PCA for Financial Inclusion

Indicators Used: ATMs per 100,000 adults, bank branches per 100,000 adults, ATMs per 1000 km, bank branches per 1000 km, domestic credit to private sector.

PCA transformed these indicators into a single composite index representing overall financial inclusion for each country.

2) PCA for Institutional Quality

Indicators Used: Voice and accountability, political stability, absence of violence, government effectiveness, regulatory quality, corruption control, rule of law, human rights protection, ease of doing business, civil liberties.

PCA created a single composite index reflecting the institutional quality of each country in the study.

Principal Component Analysis (CODE)

```
import numpy as np
import pandas as pd
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler

# Load data
financial_data = pd.read_excel('FI_PCA.xlsx')
```

```
# Selecting relevant columns
selected_columns = ['a_filled', 'b_filled', 'c_filled', 'd_filled', 'e_filled']
data = financial_data[selected_columns]

# Standardizing the data with mean 0 and standard deviation 1
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data)

# Reducing dimensionality using Principal Component Analysis (PCA)
pca = PCA(n_components=1)
composite_index = pca.fit_transform(scaled_data)

# Creating DataFrame for composite index
composite_index_df = pd.DataFrame(composite_index, columns=['Composite Index'])

print(composite_index_df)

# Exporting composite index results
composite_index_df.to_excel('output_FI_PCA.xlsx', index=False)
```

Unit Root test - Theory

Purpose of Unit Root Test: Testing Stationarity and Why is it Important!

- Testing for Stationarity: Null Hypothesis and its interpretation
- Importance in Time Series Analysis: Preprocessing Data and help in Choosing Models
- Key Concepts: Random Walk (common example of non-stationary series) and Differencing

Practical Applications:

- Economics and Finance (in analyzing variables like stock prices, inflation rates, and GDP growth), Climate and Environment (In analyzing trends in climate data, such as temperature or rainfall patterns over time).

Unit Root test – LLC test and IPS test

Table 1. Panel Unit Root Test

Variables	LLC	I(d)	IPS	I(d)
FIINDEX	-3.61168***	I(1)	-3.54626***	I(1)
INSTDEX	-4.60973***	I(1)	-5.76798***	I(1)
L.RGDP	-5.68075***	I(1)	-5.03561***	I(1)
GDPGR	-7.53995***	I(1)	-6.21315***	I(1)
L.PCRGDP	-5.79986***	I(1)	-5.09919***	I(1)
INF	-7.42197***	I(0)	-5.58246***	I(0)
TOP	-4.85710***	1(1)	-4.48596***	1(1)
UER	-3.74430***	I(1)	-2.82003***	I(1)
L.INV	-7.08923***	I(1)	-5.26073***	I(1)
LIL	-5.72671***	I(1)	-3.69455***	I(1)
TNRR	-13.9685***	I(1)	-8.44040***	I(1)

Notes: This table reports panel unit root test results. ***, **, * denote statistical significance at 1%, 5%, and 10% levels, respectively.

Unit Root test – our values and sample code

Variables	LLC	I(d)	IPS	I(d)
FIINDEX	-6.1451	I(1)***	-6.3678	I(1)***
INSTDEX	-8.1215	I(1)***	-8.3809	I(1)***
L.RGDP	-3.1336	I(1)***	-4.082	I(1)***
GDPGR	-7.5624	I(1)***	-8.6215	I(1)***
L.PCRGDP	-3.4081	I(1)***	-4.3987	I(1)***
INF	-7.5908	I(0)***	-6.2136	I(0)***
TOP	-8.5019	I(1)***	-7.763	I(1)***
UER	-3.7054	I(1)***	-2.8943	I(1)***
L.INV	-8.3768	I(1)***	-7.6312	I(1)***
TNRR	-13.7548	I(1)***	-8.0102	I(1)***

```
. xtunitroot llc d.TNRR

Levin-Lin-Chu unit-root test for D.TNRR
----------------------------------------

Ho: Panels contain unit roots                 Number of panels  =     20
Ha: Panels are stationary                     Number of periods =     16

AR parameter: Common                          Asymptotics: N/T -> 0
Panel means:  Included
Time trend:   Not included

ADF regressions: 1 lag
LR variance:     Bartlett kernel, 8.00 lags average (chosen by LLC)
------------------------------------------------------------------------
                    Statistic      p-value
------------------------------------------------------------------------
 Unadjusted t        -19.2162
 Adjusted t*         -13.7548       0.0000
------------------------------------------------------------------------
```

Two step-system GMM : Blundell & Bond (1998)

GMM: A generic method for estimating parameters in statistical models

Uses moment conditions that are functions of model parameters and data, such that their expectation is zero at parameters true value

GMM is a dynamic panel estimator.

Key difference between GMM and System GMM lies in their approach to addressing endogeneity in dynamic panel data models.

Unlike traditional GMM, System GMM simultaneously uses moment conditions based on both the levels and first differences of the variables as instruments.

By including moment conditions based on both levels and differences, System GMM allows for more efficient estimation and better control of endogeneity.

Two step-system GMM : Blundell & Bond (1998)

$$\begin{aligned} EG_{it} = & \beta_0 + \beta_1 EG_{i,t-1} + \beta_2 FI_{it} + \beta_3 IQ_{it} \\ & + \beta_4 (FI_{it} * IQ_{it}) + \beta_5 IFR_{it} + \beta_6 TOP_{it} \\ & + \beta_7 UER_{it} + \beta_8 INV_{it} + \beta_9 LIL_{it} \\ & + \beta_{10} TNRR_{it} + \mu_i + \omega_{it} \end{aligned}$$

Variables

Dependent: LRGDP or LPCRGDP or GDPGR

Endogenous: L.LRGDP or L.LPCRGDP or L.GDPGR

Explanatory: FIINDEX, INSTDEX, FIINDEXINSTDEX

Control: INF, TOP, UER, LINV, LIL, TNRR

Instruments

Internal: L.LRGDP or L.LPCRGDP or L.GDPGR

External: INF TOP UER LINV TNRR

Two step-system GMM

	2-Step System GMM	2-Step System GMM	2-Step System GMM
Variables	LRGDP	LPCRGDP	GDPGR
L.LRGDP	0.817***		
	(0.0481)		
L. LPCRGDP		0.727**	
		(0.0488)	
L.GDPGR			-0.161*
			(0.0589)
FIINDEX	0.261*	0.0185	1.594
	(0.102)	(0.0249)	(4.062)
INSTDEX	-0.0605**	0.0245***	3.003
	(0.0200)	(0.00607)	(1.436)
FIINDEXINSTDEX	-0.137**	-0.00415	0.348
	(0.0475)	(0.00548)	(3.623)
INF	-0.000121	-0.000681***	-0.0473
	(0.000295)	(0.000222)	(0.0573)
TOP	0.000116	-0.000321	-0.0240
	(0.000129)	(0.000257)	(0.0259)
UER	0.00246	0.000105	-0.175
	(0.00140)	(0.000567)	(0.0982)
LINV	0.0694***	0.0464***	7.704**
	(0.0108)	(0.00954)	(2.102)
LIL	-0.241	0.137	-43.36
	(0.142)	(0.0843)	(8.009)
TNRR	0.00104***	0.00173***	0.533*
	(0.000304)	(0.000468)	(0.0570)
DUMMY_GFC	-0.000957***	-0.00476***	-0.991
	(0.00383)	(0.00147)	(0.187)
N	300	300	300
Instruments	50	18	18
Groups	20	20	20
AR(1)	-1.68	-1.49	-1.88
AR(2)	-1.61	-0.56	-1.87
Sargan test	29.55	34.63	57.31
Hansen test	4.17	7.40	8.72

Two step-system GMM : our values

Variables	2-Step System GMM LRGDP	2-Step System GMM LPCRGDP	2-Step System GMM GDPGR
L.LRGDP	0.1279891		
	0.1299437		
L. LPCRGDP		0.1344321	
		0.1749724	
L.GDPGR			-0.3348746
			0.0940735
FIINDEX	0.2793871	0.2304331	-1.595289
	0.0951163	0.065035	23.52127
INSTDEX	0.1648276	0.0525114	-5.14774
	0.1018046	0.139105	6.410885
FIINDEXINSTDEX	-0.7655828	-0.6747864	-134.8739
	0.4497715	0.5834737	104.0144
INF	0.000022	-0.0001254	-0.0086063
	0.0001223	0.0001003	0.0107851
TOP	-0.0004594	0.0000741	-0.0047384
	0.0006464	0.0004796	0.0855155
UER	-0.0066469	-0.0154628	-1.551537
	0.0110068	0.0132382	0.9130297
LINV	0.0593053	0.0542555	9.960085
	0.0385957	0.0418884	4.632566
TNRR	0.0007946	0.0008346	0.1742934
	0.0014882	0.0023922	0.1385138
DUMMY_GFC	-0.0142079	-0.0046433	-2.697133
	0.0388328	0.0723498	6.002934

N	300	300	300
Instruments	20	20	20
Groups	20	20	20
AR(1)	-2.15	-1.7	-1.65
AR(2)	-0.32	-0.41	-0.06
Sargan test	11.32	12.94	7.87
Hansen test	12.61	14.66	13.62

Diagnostic tests for 2-step sysGMM

1) Two tests for instruments validity

Hansen (1982) J test and Sargan (1985) test of overidentiying restrictions, tests the overall hypothesis of overall validity of the instruments used

- **Failure** to reject these null hypothesis (p-value > alpha) give support to the choice of instruments

Satisfied for all our 3 tests

```
Sargan test of overid. restrictions: chi2(10)    =  12.94  Prob > chi2 =  0.227
  (Not robust, but not weakened by many instruments.)
Hansen test of overid. restrictions: chi2(10)    =  14.66  Prob > chi2 =  0.145
  (Robust, but weakened by many instruments.)
```

Diagnostic tests for 2-step sysGMM

2) Tests for autocorrelation/serial correlation of the error term

Tests the null hypothesis that the differenced error term is first and second order correlated

- **Failure** to reject these null hypothesis (p-value > alpha) of no second order serial correlation implies that the original error term is serially uncorrelated and the moment conditions are correctly specified

(that is value of AR(2) > 0.05)

Satisfied for all our 3 tests

```
Arellano-Bond test for AR(1) in first differences: z =  -1.70  Pr > z =  0.090
Arellano-Bond test for AR(2) in first differences: z =  -0.41  Pr > z =  0.678
```

Reasons for different values obtained

1) Raw Data:
   - Not completely available (2 Institutional Quality, Literacy level – dropped variable)
2) Missing value generation :
   - 2 countries whole dataset missing
   - We used forward and backward fill, mean imputation

Conclusion

- Financial inclusion has a significant positive relationship with per capita real GDP. This was observed by author as well.

Reason: As more economic agents have access to quality and formal financial products, there is a rise in their access to credits and propensity to undertake savings and investments, which increases economic activities and promotes economic growth.

- The findings reveal mixed but significant effects of institutional quality on economic growth. The results show that institutional quality has a positive relationship with economic growth, particularly where economic growth is measured as per capita real GDP and real GDP (in the case of author it was per capita real GDP and GDP growth rate)

Thank You