

HUMAN CAPITAL AND INCOME INEQUALITY LINKAGE IN SUB-SAHARAN AFRICA: PANEL DATA ANALYSIS (1984–2016)

Gashaw Getaye Molla* 

*Department of Economics, Debark University, Ethiopia

ABSTRACT

Income inequality means that one segment of the population has a disproportionately large share of income compared to the other. Disparities in income and wealth have tended to dominate the discussion on inequality because they contribute directly to individuals and families' well-being and shape the opportunities people have in life. Therefore, addressing income inequality is essential to inspire each country's population's human and productive potentials to bring development. Therefore, this study examines the relationship between income inequality and human capital using static panel data analysis. Specifically, the study employs fixed effect panel data analysis using Least Square Dummy Variable for 25 sub-Saharan African countries. The World Bank data series was widely used as the data source for macroeconomic variables, while the Gini index has obtained from the Standardized World Income Inequality Database. The empirical results reveal that human capital in terms of secondary school enrollment rate has a negative impact on income inequality. The study also found a U-shaped relationship between real gross domestic product per capita and inequality, and it does not support the well-known concept of the Kuznets curve.

Keywords: *Income Inequality; Human Capital; Panel Data; Random Effect; Fixed Effect.*

JEL: *C10, Q0, A10*

To cite this document: Molla, G. G. (2021). Human Capital and Income Inequality Linkage in Sub-Saharan Africa: Panel Data Analysis (1984-2016) (2021). *JDE (Journal of Developing Economies)*, 6(2), 186-200

ARTICLE INFO

Received: October 24th, 2020

Revised: March 1st, 2021

Accepted: November 4th, 2021

Online: November 25th, 2021

*Correspondence:

Gashaw Getaye Molla

E-mail:

gizagetaw121@gmail.com

Introduction

Income inequality is today an important economic fact and has long been a topic of interest among economists. Inequality in income is an economic problem first, and it also becomes a political and social problem in modern society. Therefore emphasizing addressing income inequality is an amoral issue, but it is also necessary to inspire human and productive potentials of each country's population to bring development towards a socially sustained path ([United Nations, 2013](#)).

Sub-Saharan Africa has long been viewed as the most inequality region globally, next to Latin America ([Cord et al., 2013](#)). The literature emphasises human capital in terms of education as one of the significant factors affecting the degree of income inequality ([Johansen, 2014](#)). Policymakers usually justify higher education spending as a very effective tool for reducing income inequality. However, theoretical studies suggest that the relationship between

education and income inequality is not always clear. For instance, the human capital model of the income distribution, stemming from the work of [Schultz \(1961\)](#) and [Teixeira \(2014\)](#), implies that the distribution of earnings (or income) is determined by the level and distribution of schooling across the population.

[Knight & Sabot \(1983\)](#) also emphasise the complicated effect of human capital accumulation on income distribution by “composition” and “wage compression” in a dual economy. They argue that an expansion of education has two different effects on the earnings distribution. The “composition” effect increases the group’s relative size with more education and initially raises income inequality but eventually lowers it. On the other hand, the “wage compression” effect decreases the premium on education as the relative supply of educated workers increases, thereby lowering income inequality. The positive effect of human capital has been widely recognised in the literature, which suggests that human capital is essential for economic growth and favourable for individuals and societies. The literature highlights education as one of the factors affecting the level of income inequality, which is the focus of this research.

The research uses secondary school enrollment rate as a proxy to human capital and the Gini coefficient to measure income inequality. The Gini coefficient measures the extent to which income distribution among countries deviates from a perfectly equal distribution. A Gini coefficient equal to zero expresses perfect equality, and a Gini coefficient equal to 1 expresses perfect inequality ([Todaro, 1994](#)).

According to [UNDP \(2017\)](#), even though sub-Saharan Africa achieved an average reduction in its unweighted Gini coefficient from about 0.47 to 0.43 between 1991 and 2011, the region remains one of the most unequal in the world – with 10 of its countries listed among the 19 most unequal in the world. This paper tries to examine the potential sources of this inequality in income by using panel data from 1984- 2016 for sub-Saharan Africa countries.

The empirical literature on the relationship between human capital and variation in income or earnings is not as rich as the literature on returns to education. But, some empirical literature generates a controversial conclusion. Among these, [Park \(1996\)](#), [De Gregorio & Lee \(2002\)](#), and [Johansen \(2014\)](#) finds a positive relationship between inequality in education and inequality in income and suggests broad-based equitable access to education as an effective policy to achieve social equity for developing countries. Conversely, [Checchi \(2004\)](#) finds a strong negative linkage between human capital and measured income inequality. In this paper, the researcher contributes to this empirical puzzle by using panel data to systematically describe, identify and analyse the variation in outcomes of empirical studies. Based on available panel data, the study aims to comprehensively study the linkage between human capital and income inequality in twenty-five (25) sub-Saharan African countries. Specifically, the study is designed to investigate the effect of human capital in the secondary school enrollment rate on income inequality in sub-Saharan Africa and test the inverted U-shaped relationship between income inequality and GDP per capita (i.e. testing Kuznets hypothesis) for Sub-Saharan Africa.

Literature Review

Human Capital and Income Inequality

Why is human capital substantial when we look at income inequality? First is the contribution to having a good life. Second is its relation to economic growth and the distribution of income. In human capital theory, [Teixeira \(2014\)](#) showed that acquiring education increases individuals’ skills, competencies, and productivity. Since wages equal workers’ productivity in

a competitive labour market, higher productivity will lead to a higher wage. It means that a more educated society holds greater welfare. Since the conception of this theory, it has been the focus of increasing research. It has encouraged the production of many empirical and theoretical studies.

Acknowledging a causal relationship between education and earning is a well-established result, but it is less clear-cut when analyzing the link between income inequality and educational attainments. On the one hand, rising wage inequality should encourage investments in education mainly because it raises the return to education. [Topel \(1997\)](#) observes a faster skill accumulation as a result of rising returns. This increase in the supply of skills should eventually mitigate the increase in inequality.

On the other hand, when income inequality increases, it also affects households' resources to finance education. The intergenerational theory claims that there exists a perfect correlation between income and education distributions. It causes that barriers, e.g. liquidity constraints and family background, might prevent the investment in education for that part of the population belonging to the bottom of the income distribution. If this is persistent, then the population will be trapped in low education and income levels for more than one generation.

The accumulation of human capital has also been essential for economic growth and favourable for individuals and societies. The positive effect on the individual is that the more educated people are, the better the labour market term in wage and employability. [Lochner & Moretti \(2004\)](#) find that other positive effects will be better health, fertility, well-being and a lower chance of engaging in crime. In imperfect credit markets, [Galor & Zeira \(1993\)](#) show that wealth distribution affects investments in human capital.

Developing an overlapping generation model with intergenerational transmissions suggests that the initial distribution of wealth is crucial to determine individuals' education choices and the aggregated output in both the short and the long run. [Banerjee & Newman \(1993\)](#) end up with similar conclusions. Their theoretical model suggests that the initial distribution of wealth shapes the pattern of occupational (educational) choice. [Filmer & Pritchett \(1999\)](#) perform an empirical analysis using household surveys for 35 countries, where they use the poverty index as their proxy for the household's economic status. They find that the poverty index is correlated with reduced school attainment in the poorest 40 per cent of the population. [Checchi \(2003\)](#) analyses the issue using an unbalanced panel for 108 countries from 1960-1995. His main finding is a robust negative correlation between income inequality and secondary education enrolment. This dataset makes it possible to use econometric methods to address the problem of reverse causality, i.e. are people more educated because they have a higher income or do people have a high income because they have higher education?

More recently, it has increasingly been acknowledged that some of the unequal economies in the world are in Africa. Using the Gini coefficient as the measure of within-country income inequality, the average Gini coefficient in Africa is 0.43, which is 1.1 times the coefficient for the rest of the developing world, at 0.39. Furthermore, the upper bound of the continent's range of Gini coefficients exceeds that of the developing world, indicating that extreme inequality is also a distinct feature on the African continent. Using another measure of income inequality shows that, on average, the top 20 per cent of earners in Africa have an income over ten times that of the bottom 20 per cent.

Over the past two decades, evidence from all over the world has shown the harmful effects of high levels of inequality on everything from economic growth to poverty reduction,

social unity and public health. A similar pattern has been shown in sub-Saharan Africa, especially regarding the influence of growth on reducing poverty. Reducing inequality is not only helpful but essential. High inequality is divisive and socially corrosive (Wilkinson & Pickett, 2010).

The economy of sub-Saharan Africa has grown at an unprecedented pace over the past decade. Seven of the ten countries with the highest growth rates worldwide are in Africa. However, growth has been concentrated in particular sectors of the economy and specific geographical areas within countries. The benefits of this growth have not been broadly shared and have left out large sections of the population. Poverty has not fallen as much or as fast as expected, and economic inequalities have remained high. There are, of course, significant differences between the countries in the region and their directions of inequality.

There is broad agreement that the average economic inequality in sub-Saharan Africa is the highest in the world after Latin America (Damodar N, 2004; Milanovic, 2003). The average level of income inequality in sub-Saharan Africa declined from the 1960s to the 1980s. On the other hand, it increased in the 1990s and fell again in the 2000s. Empirical research on inequality started in 1955 when Simon Kuznets published his study on the economic growth-income inequality relationship. The general conclusion suggested for developed countries using insufficient data was that the relative distribution of income, as measured by annual income incidence is rather broad classes, has been moving toward equality. The reduction in income inequality in developed countries was accompanied by significant rises in real income per capita. However, for underdeveloped countries, income inequality increased (Kuznets, 2019).

Johansen (2014) investigates and analyzes the effect of human capital on income inequality. This study uses educational attainment as a proxy for human capital to investigate its effect. The dataset used for the empirical investigation contains data on 123 countries from 1960 to 2010. This dataset makes it possible to use econometric methods to address the problem of reverse causality, i.e. are people more educated because they have a higher income or do people have a high income because they have higher education? A two-least square estimation is used to address this problem of endogeneity using parents' education as an instrument. The instrumental variable estimation results present a positive and significant relation between improved educational attainment and income inequality.

Research Methodology

The researcher presents different econometric specifications. Firstly, the OLS estimations on pooled panel data are presented. Secondly, to account for the panel dimension of our panel data, the researcher conducts fixed and random effect models and choose the appropriate one by testing the Hausman test. For empirical analysis, this paper first assumes that income inequality is the function of the secondary school enrollment rate. The model is adopted by including some other explanatory variables and intends to present an analysis of income inequality as a function of human capital (secondary school enrollment rate), trade openness, GDP per capita, squared GDP per capita, inflation, population growth, unemployment rate, the inflow of foreign direct investment and natural resource. Functionally indicated as follow;

$$GINI_{it} = (GDPPC_i, GDPPC_{it}^2, SSE_{it}, OPN_{it}, UNEMP_{it}, FDI_{it}, NR_{it}, POP_{it}, INF_{it}) \quad (1)$$

Where i denote the countries and t denoting the time, the i subscript denotes the cross-section dimension, whereas t denotes the time-series dimension. GINI is the Gini coefficient of income inequality, SSE is secondary school enrollment rate, GDPPC² is per capita income,

GDPPC² is squared per capita income. OPN is trade openness, INF is the inflation rate, POP is population growth, UNEP is the unemployment rate, NR is a natural resource, FDI is foreign direct investment.

Dataset and sources

The data set for income inequality (Gini coefficient) derived from SWIID (Standardized World Income Inequality Database) provides comparable Gini indices of gross and net income inequality. The dataset for the explanatory variable includes human capital (secondary school enrollment rate), trade openness, GDP per capita, squared GDP per capita, inflation, population growth, unemployment rate, and the inflow of foreign direct investment. On the other hand, natural resources are mainly derived from the World Bank - Main Economic Indicators Database. The world Penn tables (WPT), IMF, and UNICTAD were also used as additional data sources. The data set of this study is long balanced panel data as the number of observations (cross-sectional data) is lesser than the number of periods (time series data). The dataset contains 825 observations covering over twenty-five (25) sub-Sahara African countries, listed in appendix1 over the year 1984–2016. This study uses annual observations for the set of the 25 countries. The choices of the countries and the period of 1984–2016 are guided by the availability of data.

Table 1: Summary of Variable Description, Measurement and Expected Sign

No	Variable	Measurement and Symbol	Expected sign
1	Income inequality	Natural logarithm of Gini coefficient (<i>lnGINI</i>)	Dependent
2	Economic growth	Natural logarithm of Real GDP per capita (<i>lnPGDP</i>)	+ve
3	Economic growth ²	Natural logarithm of real GDP per capita squared(<i>lnPGDP²</i>)	-ve
4	Education	Secondary school enrollment rate, (<i>lnSSE</i>)	-ve
5	Inflation	Annual percentage change in consumer price index, (<i>INF</i>)	+ve
6	Population	Growth of population (<i>lnPOP</i>)	+ve
7	Unemployment	Unemployment rate, (<i>lnUNEP</i>)	+ve
8	Trade openness	Sum of total exports and imports as a percentage of GDP, , (<i>lnOPN</i>)	-ve
9	Natural resources	Sum of natural resource divided by GDP, (<i>lnNR</i>)	+ve
10	Foreign direct Inv't	Foreign direct investment inflow as percentage of GDP, , (<i>FDI</i>)	-ve

Methods of Analysis

General Panel Model

In this section, we represent a general panel model that enables us to consider individual heterogeneity (latent time-invariant variables) as in the usually fixed effect and random effect model but permits additional structures for comparison. And we can derive the well-known random effects and fixed effects models by imposing restrictions on this general panel model (Bollen & Brand, 2010). Consider the following equation

$$Y_{it} = \beta_{yx}X_{it} + \beta_{yz}Z_i + \lambda t\eta_i + \varepsilon_{it} \forall_i = 1, \dots, N \text{ and } \forall_t = 1, \dots, T \dots \dots \dots \quad (2)$$

Where Y_{it} is the value of the dependent variable for the i th case in the sample at the t^{th} period, X_{it} is the vector of time-varying covariates for the i^{th} case at the t^{th} time period, β_{yxt} is the row vector of coefficients that give the impact of X_{it} on Y_{it} at time t , Z_i is the vector of observed time-invariant covariates for the i th case with a row vector of coefficients at time t that give the impact of Z_i on Y_{it} .

The η_i is a scalar of all other latent time-invariant variables that influence Y_{it} and λ_t is the coefficient of the latent time-invariant variable (η_i) at time t and at least one of these λ_t is set to one to provide the units in which the latent variable is measured (e.g., set $\lambda_1 = 1$). The ε_{it} is the random disturbance for the i th case at the t th time period with. It also is assumed that ε_{it} is uncorrelated with X_{it} , Z_i and η_i and that for $t \neq s$.

The η_i represents individual heterogeneity that affects the outcome variable. In this research case, the dependent variable is the GINI coefficient as the measure of income inequality within the country. Explanatory variables include human capital proxied by education, trade openness, per capita GDP of each country and the others listed in Table 1. Based on this researcher formulated the random effect and fixed effect model as follows.

Random Effects Model (REM)

Consider an economic relationship that involves a dependent variable, Y , and several observable explanatory variables represented by, X_k . You have panel data for Y and X_k . The panel data consists of N -units and T -time periods, and therefore you have N times T observations. It can write the random-effects model as

$$Y_{it} = \beta_k \sum X_{kit} + v_i + \varepsilon_{it} \text{ for } i = 1, 2, \dots, T \tag{3}$$

β_k = vectors of the parameter

X_{kit} = vectors of explanatory variable

The classical error term is decomposed into two components. The component v_i represents all unobserved factors that vary across units but are constant over time. The component ε_{it} represents all unobserved factors that vary across units and time. It is assumed that v_i is given by

$$v_i = a_0 + \omega_i \text{ for } i = 1, 2, \dots, N \tag{4}$$

The v_i is decomposed into two components: 1) a deterministic component α_0 , 2) a random component ω_i . Once again, each of the N units has its own intercept. However, in this model the N intercepts are not fixed parameters; instead, they are random variables. The deterministic component α_0 is interpreted as the populations mean intercept. The disturbance ω_i is the difference between the population mean intercept and the intercept for the i^{th} unit. It is assumed that the ω_i for each unit is drawn from an independent probability distribution with the following properties,

$$E(\omega_i) = 0 \quad Var(\omega_i) = \sigma\omega^2 \quad Cov(\omega_i, \omega_s) = 0 \quad Cov(\omega_i, Xstk) = 0$$

The N random variables v_i are called random effects. It can rewrite the random effects model equivalently as

$$Y_{it} = \alpha_0 + \beta_k \sum X_{kit} + \mu_{it} \tag{5}$$

Where $\mu_{it} = \omega_i + \varepsilon_{it}$. An important assumption underlying the random-effects model is that the error term μ_{it} is not correlated with any explanatory variables. If you estimate the random-effects model using the OLS estimator, you will obtain parameter estimates that are unbiased but inefficient. In addition, the OLS estimates of the standard errors and hence t-statistics are incorrect. It is because the OLS estimator ignores the autocorrelation in the error term μ_{it} . We can use a Feasible GLS (FGLS) estimator to obtain unbiased and efficient estimates that consider the correlated auto disturbances.

For this research, the random effect model is specified as follows:

$$\ln GINI_{it} = \alpha + \beta_0 \ln GDPPC_{it} + \beta_1 \ln GDPPC_{it}^2 + \beta_2 \ln SSE_{it} + \beta_3 \ln OPN_{it} + \beta_4 \ln UNEMP_{it} + \beta_5 \ln FDI_{it} + \beta_6 \ln NR_{it} + \beta_7 \ln POP_{it} + \beta_8 \ln INF_{it} + \omega_i + \varepsilon_{it} \quad (6)$$

The disturbance ω_i is the difference between the population mean intercept (α_0) and the intercept for the i^{th} unit.

Fixed Effects Model (FEM)

The fixed-effects model is used whenever you are only interested in analyzing the impact of variables that vary over time. The fixed-effects model explores the relationship between the predictor and outcome variables within an entity. Each entity has its characteristics that may or may not influence the predictor variables. When using the fixed effects model, we assume that something within the individual country may impact or bias the variables, and we need to control for this. Another important assumption of the fixed effects model is that the time-invariant characteristics are unique to the individual country and should not be correlated with other individual characteristics. Each entity is different. Therefore, the entity's error term and the constant should not be correlated with others (Hausman & Taylor, 1981).

Consider an economic relationship that involves a dependent variable, Y , and the observable explanatory variables, X_k , and one or more unobservable confounding variables. You have panel data for all variables (dependent and explanatory). Again the panel data consists of N -units and T -time periods, and therefore you have N times T observations.

The classical linear regression model without an intercept is given by

$$Y_{it} = \beta k \sum X_{kit} + \mu_{it} \text{ for } i = 1, 2, \dots, N \quad \text{and} \quad t = 1, 2, \dots, T \quad (7)$$

Where Y_{it} is the value of Y for the i^{th} unit for the t^{th} time period; X_k is the value of X_k for the i^{th} unit for the t^{th} time period, and μ_{it} is the error for the i^{th} unit for the t^{th} time period.

The fixed effects regression model, which is an extension of the classical linear regression model, is given by

$$Y_{it} = \beta k \sum X_{kit} + v_i + \varepsilon_{it} \quad (8)$$

Where $\mu_{it} = v_i + \varepsilon_{it}$. The error term of the classical linear regression model is decomposed into two components. The component v_i represents all unobserved factors that vary across units but are constant over time. The component ε_{it} represents all unobserved factors that vary across unit and time, idiosyncratic error. It is assumed that the net effect on Y of unobservable factors for the i^{th} unit that are constant over time is a fixed parameter designated α_i . Therefore, the fixed effects model can be rewritten as

$$Y_{it} = \beta k \sum X_{kit} + \alpha_1 + \alpha_2 + \alpha_3 \dots \dots \alpha_N + \varepsilon_{it} \quad (9)$$

The unobserved error component has been replaced with fixed parameters, $\alpha_1 + \alpha_2 + \dots + \alpha_N$, one parameter for each of the N units in the sample. These parameters are called unobserved effects and represent unobserved heterogeneity. For example, α_1 represents the net effect on Y of unobservable factors that are constant over time for unit one, α_2 for unit two... α_N for unit N . Therefore, in the fixed-effects model, each unit is used for the unit in the sample has its intercept. These N intercepts control for the net effects of all unobservable factors that differ across units but are constant over time. It can be Two alternative but equivalent estimators to estimate the parameters of the fixed-effects model. 1) Least squares dummy variable estimator. 2) Fixed effects estimator, but the researcher used Least Squares Dummy Variable

Estimator for this research. The least-squares dummy variable estimator involves two steps. In step 1, create a dummy variable for each of the N units in the sample. These N dummy variables are defined as follows.

$$D_{ki} = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}$$

In step 2, run a regression of the dependent variable on the N dummy variables and the explanatory variables using the OLS estimator. For a model with N units and explanatory variables, the step 2 regression equations without an intercept are

$$Y_{it} = \alpha_1 D_{1i} + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \dots + \alpha_N D_{Ni} + \beta_k \sum X_{kit} + \epsilon_{it} \tag{10}$$

or with an intercept is

$$Y_{it} = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \dots + \alpha_N D_{Ni} + \beta_k \sum X_{kit} + \epsilon_{it} \tag{11}$$

It can run the least-squares dummy variable regression with dummy variables for all units and no intercept or $N-1$ units with an intercept. It yields estimates of the N fixed-effects intercept parameters and the slope parameters.

For the fixed effects model, the following model is estimated.

$$\ln GINI_{it} = \alpha + \beta_0 \ln GDPPC_{it} + \beta_1 \ln GDPPC_{it}^2 + \beta_2 \ln SSE_{it} + \beta_3 \ln OPN_{it} + \beta_4 \ln UNEMP_{it} + \beta_5 \ln FDI_{it} + \beta_6 \ln NAT_{it} + \beta_7 \ln POP_{it} + \beta_8 \ln INF_{it} + v_i + \epsilon_{it} \tag{12}$$

In this equation v_i is interpreted as $\alpha_1, \alpha_2, \dots, \alpha_N$, one parameter for each of the N units in the sample.

Hausman Test for Fixed and Random Effects Models

In many situations, you may be uncertain whether the unit dependent unobserved effects (v_i) are correlated with one or more of the explanatory variables, and therefore uncertain whether the fixed-effects model or random-effects model is most appropriate. In these situations, you can use a Hausman test to test whether the unit dependent unobserved effects (v_i) are correlated with the explanatory variables.

For the Hausman test, the null and alternative hypotheses are as follows.

1. Defining the null and alternative hypotheses:
 - H_0 : v_i is not correlated with X_{it} , i.e. the appropriate model is Random effects. There is no correlation between the error term and the independent variables in the panel data model, i.e. the covariance between the error term and the covariates is zero. $Cov(\alpha_i, X_{it}) = 0$
 - H_1 : v_i is correlated with X_{it} , i.e. the appropriate model is fixed effects. The correlation between the error term and the independent variables in the panel data model is statistically significant, i.e. the covariance between the error term and the covariates is different from zero. $Cov(\alpha_i, X_{it}) \neq 0$
2. The statistic, computed above is compared with the critical values for the χ^2 distribution for k degrees of freedom, where k is the number of slope parameters in the model. The null hypothesis is rejected if the Hausman statistic is bigger than its critical value.

Econometric Analysis

Panel Regression Results

Table 2: Summary of Panel Regression Results

Dependent variable: LNGINI				
Variable	1. FE (time effect)	2.FE(country effect)	3.FE (Time+Country)	4. RE
InGDPPC	-.3272656***	-.24923436**	-.2100478	-.23808356**
	.08231022	.08583135	.11373489	.084999
	-3.98	-2.90	-1.85	-2.80
InGDPPC2	.02810948***	.02160878***	.02094079*	.02068348***
	.00565181	.00589218	.00815599	.00582662
	4.97	3.67	2.57	3.55
InSSE	.02663216	-.02252748*	-.0083884	-.03018569**
	.01492695	.01127492	.01197505	.01006189
	1.78	-2.00	-0.70	-3.00
InOPN	-.00726325	.03361064**	.03229016**	.02886882**
	.01569399	.01041174	.0110293	.01007954
	-0.46	3.23	2.93	2.86
InUNEMP	.08201649***	-.00038794	-.00802485	.00753424
	.00970076	.0113769	.01209941	.01111238
	8.45	-0.03	-0.66	0.68
FDI	-.00474534**	-.00335426***	-.00337689***	-.00351448***
	.00146209	.0007503	.0007649	.00075081
	-3.25	-4.47	-4.41	-4.68
InNR	.0514744***	-.00988485**	-.00895861*	-.00835505*
	.00409511	.00355175	.00363788	.00352827
	12.57	-2.78	-2.46	-2.37
InPOP	-.04378859***	-.08118482***	-.09541226	-.06016002***
	.00588996	.02076898	.05722093	.01501078
	-7.43	-3.91	-1.67	-4.01
INF	.0007569**	.00041202**	.00041395**	.0004491**
	.00025463	.00013936	.00014339	.0001393
	2.97	2.96	2.89	3.22
Cons	4.9860426***	5.8651631***	5.7269187***	5.4380316***
	.28902711	.39942736	.66958577	.37695961
	17.25	14.68	8.55	14.43
No of obs.	825	825	825	825
R-squared	0.4864	0.8748	0.8789	0.2154
Adjusted R ²	0.4595	0.8697	0.8685	
Hottest				
P-value	0.0000	0.0839	0.0103	
Time effects	Yes	No	Yes	No

Dependent variable: LNGINI				
Variable	1. FE (time effect)	2.FE(country effect)	3.FE (Time+Country)	4. RE
P-value	0.8701		0.7747	
Country ef- fects	No	Yes	Yes	No
P-value		0.0000	0.0000	
Random ef- fects	No	No	No	Yes

Source: Research Data, Processed

The first-panel regression, i.e. regression number one, includes only time fixed effects. In this regression, each coefficient is statistically significant except for trade openness and secondary school enrollment rate. However, an F-test of the time fixed effects variables results in a p-value of 0.8701. The results, in conclusion, are that time fixed effects are not present in this panel data. Hence, it cannot reject the null hypothesis of all time dummy parameters except zero.

The following panel regression (2) includes country fixed effects instead of time fixed effects. Including the country fixed effects increases the adjusted R2 to 0.8697. However, the statistical significance of the coefficients suffers from the inclusion of country fixed effects. Though conducting an F-test for the validity of country effects returns an F-value of 124.95 and a p-value of 0.0000. The results of the F-test indicate that country effects are valid and should be included in the model specification. Only one variable (unemployment) is now statistically insignificant. After controlling for country effects, secondary school enrollment and trade openness have changed signs, but they have become significant.

Regression three includes both time and country effects. Much of the results appear similar to the country effects model in regression two. The magnitudes of the coefficients have changed, and secondary school enrollment, unemployment rate, the initial level of GDP per capita and population are no longer statistically significant. The adjusted R2 value does not change more. Testing the fixed effects coefficients using an F-test again fails to reject that the time effects are different from zero with a p-value of 0.7747. Furthermore, the country effects test allows for rejecting that the country effects coefficients are equal to zero with a p-value of 0.0000. These two tests suggest that time effects are not present but that country effects are present in the data.

Regression four uses a random-effects model instead of fixed effects, as seen in the previous three regressions. The result of the random effects model is similar to the country effects model only in the sign of the coefficients. Still, the magnitudes have changed dramatically except for the unemployment rate. A comparison to the country effects model (regression two) is needed since the random effects regression (regression four) results appear to be relevant. The test (Hausman test) is presented in the previous section for this comparison. And the result showed that the random-effects model is not the correct specification, but instead, that including country effects is the proper model to use. And in the random-effects model, all except the unemployment rate are statistically significant coefficients. So, the introduction of country effects improves the overall explanatory power of the model. But, the unemployment rate is no longer significant, and the sign has changed. The coefficient for the unemployment rate is likely no longer significant as unemployment is relatively constant over time within each country. As a result, unemployment is controlled by the inclusion of country fixed effects.

Panel Regression Interpretation

The signs and magnitudes of the coefficients of the independent variables are interpreted based on a result obtained by the one-way fixed-effect model in panel regression specifications. Because the two ways are fixed-effect model is not appropriate due to the absence of time fixed effect. According to the results, the economic growth rate measured by GDP per capita has a statistically significant and negative relationship with income inequality as the t-value is -2.91. This negative relationship result is inconsistent with the hypothesis that previously mentioned economic growth and income inequality has a positive relationship at the initial. That is, increasing the size of the economic pie contributed to equal distribution of income. It is possible to say that getting a bigger economy can help equalize income. It is an exciting result because it provides several questions and topics for further study. First, governments have an incentive to spend their resources more on boosting the economy to reduce income inequality. But, the squared GDP per capita has a statistically significant and positive relationship with income inequality as the t-value is 3.67. So, this finding does not support the Kuznets inverted U-shape hypothesis. The result shows that income inequality has a negative relationship with GDP per capita and a positive relationship with squared GDP per capita, i.e. they have a U-shaped relationship. Despite the limitation in a number of observations and data limitations, the finding supports the results of several previous types of research of [Kentworthy \(1999\)](#) and [Bénabou \(2000\)](#).

The primary variable, which secondary school enrollment rate (a measure of human capital), has a statistically significant and negative relationship with income inequality as the t-value is -2.00, which supports the prior expectation. This result supports the finding of [Barro \(2000\)](#) that founds a negative relationship between primary and secondary school enrollments and income inequality. This negative relationship between secondary enrollments and income inequality may be thought of as inherently connected to development. An increase in the supply of educated workers tends to diminish the wage gap and, thereby, decreases income inequality.

Testing the Kuznets Inverted-U Hypothesis

The GINI coefficient usually measures income inequality while the GDP per capita characterizes the level of economic development. Most studies performed linear regression of GINI index on the logarithm of GDP per capita and its square term, i.e. the relationship between GINI index and GDP per capita is assumed in the form;

$$\ln GINI_{it} = \beta_0 \ln GDPPC_{it} + \beta_1 \ln GDPPC_{it}^2 + \alpha_0 + \varepsilon_{it} \quad (13)$$

Where GINI is shown, Gini-coefficient and economic growth are by GDP. In this hypothesis, $\beta_0 > 0$ and $\beta_1 < 0$ is regularly predictable in testing intended for the Kuznets inverted U-curve. In the above equations, α_0 is the constant term. Moreover, ε_{it} is the residual of the equation.

Table 3: Testing Kuznets Hypothesis Model

		Dependent variable: LNGINI		
Variable		Pooled OLS	Fixed Effect	Random Effect
lnGDPPC	Coefficient	-.06009661	-.41094197***	-.39766248***
	St.Err	(.08837141)	(.08813494)	(.08758498)
	t-value	(-0.68)	(-4.66)	(-4.54)
lnGDPPC2	Coefficient	.01134861	.02656432***	.02671485***

Dependent variable: LNGINI				
Variable		Pooled OLS	Fixed Effect	Random Effect
	St.Err	(.00617534)	(.00619206)	(.00614959)
	t-value	(1.84)	(4.29)	(4.34)
Cons	Coefficient	3.574967***	5.2561401***	5.157362***
	St.Err	(.31174819)	(.31365031)	(.31315802)
	t-value	(11.47)	(16.76)	(16.47)
F(2, 822)= 125.20		F(2,798)= 13.48		Wald chi2(2)=21.89
Prob> F=0.0000		Prob> F = 0.0000		Prob> chi2=0.0000
R-squared=0.2335		R-sq:		R-sq:
Adj R-squared=0.2316		within= 0.0327		within =0.0311
Root MSE=.17747		between= 0.1642		between = 0.070
		overall= 0.1050		overall = 0.0375
		sigma_u =.20872486		sigma_u = .16573602
		sigma_e =.07895524		sigma_e = .07895524
		rho = .87482066		rho = .81502984

Source: Research Data, Processed

Table 4: Choosing the Appropriate Model To Test Kuznets Hypothesis

Pooled OLS vs FE	Pooled OLS vs RE	FE vs RE	Decision
H0=pooled OLS Ha=FE	H0=pooled OLS Ha=RE	H0=RE Ha=FE	
F(24, 798) =139.79	chibar2(01)=7845.91	chi2(2) =16.62	Fixed effect model is appropriate
Prob> F = 0.0000	Prob>chibar2=0.0000	Prob>chi2 = 0.0002	
Reject H0 (FE have chosen)	Reject H0 (RE have chosen)	Reject H0 (FE have chosen)	

Source: Research Data, Processed

Since the appropriate model is the fixed-effect model, it should test the hypothesis by using this model. So, as shown in Table 4 above, since (β_0 is negative and significant), economic growth (indicated by GDP per capita) has a negative and significant impact on income inequality at the initial level. And (β_1 , which is the coefficient of squared GDP per capita is positive and significant) at 1percent significance level. This result in a negative relationship between economic growth and income inequality at the initial level. A positive relationship at the latter stage of development does not follow Kuznets inverted U- shape hypothesis since sub-Saharan Africans are a low-income country, making economic growth and income inequality rise at the same time. The hypothesis can be rejected at a 1% significance level as both $\ln G-DPPC$ and $\ln GDPPC^2$ are highly significant. As shown from the above table 4.11, a 1% increase of GDP per capita initially causes 0.41%decrease in the Gini coefficient characterizing income inequality. Still, the impact fades down as the country’s GDP per capita becomes higher. Ceteris paribus, a 1% higher GDP, will cause a 0.026% increase in income inequality.

Generally, the relationship between income inequality and the per capita income level in sub-Saharan Africa is quite different from much of the rest of the world. Inequality appears to be higher at all income levels in the region than elsewhere. It can be explained by social problems related to inequality such as stealing, corruption, political instability.

Since the appropriate model is the fixed-effect model, It should test the hypothesis by using this model. So, as shown in Table 4 above, since (β_0 is negative and significant), eco-

economic growth (indicated by GDP per capita) has a negative and significant impact on income inequality at the initial level. And (β_1 , which is the coefficient of squared GDP per capita is positive and significant) at 1percent significance level. This result in a negative relationship between economic growth and income inequality at the initial level. A positive relationship at the latter stage of development does not follow Kuznets inverted U- shape hypothesis since sub-Sahara Africans are a low-income country, and this would make economic growth and income inequality rise at the same time. The hypothesis can be rejected at a 1% significance level as both $\ln\text{GDPPC}$ and $\ln\text{GDPPC}^2$ are highly significant. As shown from the above table 4.11, a 1% increase of GDP per capita initially causes 0.41% decrease in the Gini coefficient characterizing income inequality. Still, the impact fades down as the country's GDP per capita becomes higher. Ceteris paribus, a 1% higher GDP, will cause a 0.026% increase in income inequality.

Generally, the relationship between income inequality and the per capita income level in sub-Saharan Africa is quite different from much of the rest of the world. Inequality seems markedly higher at all levels of income in the region than elsewhere. It can be explained by the social problems associated with inequality, such as stealing, corruption, political instability.

Conclusion and Policy Implication

The finding indicates that the primary variable, secondary school enrollment rate (a measure of human capital), has a statistically significant and negative relationship with income inequality. An increase in the supply of educated workers diminishes the wage gap and, thereby, decreases income inequality. The economic growth rate measured by GDP per capita has a statistically significant and negative relationship with income inequality, i.e. increasing the size of the economic pie contributed to an equal distribution of income. But, the squared GDP per capita has a statistically significant and positive relationship with income inequality, i.e. as per capita income becomes too high, income inequality rises. And the result does not found to support the Kuznets inverted U-shape hypothesis. Increased economic development (per capita income) tends to lower inequality before a threshold of income is reached. After this point, the curve turns, so increased development widens inequality.

Firstly to forward the appropriate recommendation, the researcher classifies the explanatory variable as equalizing and dis-equalizing factors. Among the variables equalizing factors (that have income inequality reducing impact) include; the distribution of human capital (mainly secondary education), real GDP per capita, foreign direct investment, the share of natural resources in GDP and population. And the dis-equalizing factors (that have income inequality increasing impact) include trade openness and inflation. From the finding of this research, we have seen that human capital (secondary education) has a significantly significant contribution to reducing income inequality in the study area (SSA). So, improving literacy rates (secondary education) is crucial for a sustainable solution to income inequality. Increasing participation, improving learning, and enhancing the relevance of secondary education, especially in rural areas of sub-Saharan Africa, is a better solution to narrow the income distribution gaps. Based on the results, it is recommended that the sub-Saharan African countries give equal access to education (especially at secondary) attention if their objective is to improve income distribution. Therefore, it is necessary to improve the conditions of educational institutes and remove disparities because disparities in educational institutes cause differences in the skills of people or students, income level, and professional opportunities.

References

- Banerjee, A. V., & Newman, A. F. (1993). Occupational choice and the process of development. *Journal of Political Economy*, 101(2). <https://doi.org/10.1086/261876>
- Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5(1). <https://doi.org/10.1023/A:1009850119329>
- Bénabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, 90(1). <https://doi.org/10.1257/aer.90.1.96>
- Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces*, 89(1). <https://doi.org/10.1353/sof.2010.0072>
- Checchi, D. (2003). Inequality in incomes and access to education: A cross-country analysis (1960-95). *Labour*, 17(2). <https://doi.org/10.1111/1467-9914.00226>
- Checchi, D. (2004). Does Educational Achievement Help Explain Income Inequality? In *Inequality Growth and Poverty in an Era of Liberalization and Globalization*. <https://doi.org/10.1093/0199271410.003.0004>
- Cord, L., Lucchetti, L., Castelán, C. R., & Ratzlaff, A. (2013). *Prosperity and Poverty in Latin America and the Caribbean*. Voxeu. <https://voxeu.org/article/prosperity-and-poverty-latin-america-and-caribbean>
- Damodar N, G. (2004). *Basic econometrics*. The Mc-Graw Hill.
- De Gregorio, J., & Lee, J. W. (2002). Education and income inequality: New evidence from cross-country data. *Review of Income and Wealth*, 48(3). <https://doi.org/10.1111/1475-4991.00060>
- Filmer, D., & Pritchett, L. (1999). The effect of household wealth on educational attainment: Evidence from 35 countries. *Population and Development Review*, 25(1). <https://doi.org/10.1111/j.1728-4457.1999.00085.x>
- Galor, O., & Zeira, J. (1993). Income distribution and macroeconomics. *Review of Economic Studies*, 60(1). <https://doi.org/10.2307/2297811>
- Hausman, J. A., & Taylor, W. E. (1981). Panel Data and Unobservable Individual Effects. *Econometrica*, 49(6). <https://doi.org/10.2307/1911406>
- Johansen, A. L. (2014). The effect of human capital on Income Inequality: An Econometric Analysis. *Economics Letters*, 124(1).
- Kentworthy, L. (1999). *Welfare States, Real Incomes, and Poverty* (No. 370). <https://core.ac.uk/download/pdf/39413248.pdf>
- Knight, J. B., & Sabot, R. H. (1983). Educational Expansion and the Kuznets Effect. *The American Economic Review*, 73(5).
- Kuznets, S. (2019). Economic growth and income inequality. In *The Gap Between Rich and Poor: Contending Perspectives On The Political Economy of Development*. <https://doi.org/10.4324/9780429311208-4>
- Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from pris-

- on inmates, arrests, and self-reports. *American Economic Review*, 94(1). <https://doi.org/10.1257/000282804322970751>
- Milanovic, B. (2003). *Is Inequality in Africa Really Different?* (No. 3169). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=636588
- Park, K. H. (1996). Educational expansion and educational inequality on income distribution. *Economics of Education Review*, 15(1). [https://doi.org/10.1016/0272-7757\(95\)00000-3](https://doi.org/10.1016/0272-7757(95)00000-3)
- Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review*, 51, 1–17.
- Teixeira, P. N. (2014). Gary Becker’s early work on human capital – collaborations and distinctiveness. *IZA Journal of Labor Economics*, 3(1). <https://doi.org/10.1186/s40172-014-0012-2>
- Todaro, M. P. (1994). *Economic Development* (5th ed.). Longman.
- Topel, R. H. (1997). Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality. *Journal of Economic Perspectives*, 11(2). <https://doi.org/10.1257/jep.11.2.55>
- UNDP. (2017). *Income Inequality Trends in Sub-Saharan Africa: Divergence, Determinants and Consequences*. <https://www.africa.undp.org/content/rba/en/home/library/reports/income-inequality-trends-in-sub-saharan-africa--divergence--dete.html>
- United Nations. (2013). *Inequality Matters*. <https://www.un.org/esa/socdev/documents/reports/InequalityMatters.pdf>
- Wilkinson, R., & Pickett, K. (2010). *The Spirit Level: Why Greater Equality Makes Societies Stronger*. Bloomsbury Press.