



THE ANALYSIS OF THE MAIN FACTORS INFLUENCING INCOME INEQUALITY: A CASE OF SOUTH AFRICA

Olimjonova Malika Sadriddin qizi

Executive Assistant

Westminster International University in Tashkent

ma.olimjonova@wiut.uz

Abstract

This study investigates the main determinants of income inequality in South Africa over the period 1970–2020. Despite sustained economic growth since the end of apartheid, South Africa remains one of the most unequal countries in the world, with a Gini coefficient of approximately 63%. Using an econometric approach, the study examines the effects of inflation, literacy, unemployment, education, government expenditure, government debt, taxes, and population growth on income distribution. The empirical findings reveal that inflation, literacy, unemployment, education, and government expenditure are statistically significant determinants of income inequality. Among these factors, educational attainment has the strongest impact on increasing income disparity, while higher government expenditure contributes to reducing inequality. In contrast, government debt, taxes, and population growth are found to have no significant effect. The results highlight the importance of targeted economic and social policies aimed at addressing structural inequalities and promoting inclusive growth in South Africa.

Keywords: Income inequality, South Africa, Gini coefficient, Inflation, Unemployment, Education, Government expenditure, Economic growth, Income distribution, Econometric analysis.

Introduction and Motivation

Income inequality impacts countries at all stages of development and with various sociological, economic and demographic characteristics. According to Dabla-Norris et al. (2015), the gap between the rich and the poor in advanced economies has been at its greatest level for decades, although inequality dynamics in emerging markets and



developing nations have been more volatile. Among these, South Africa, being classified as an upper-middle income country, has historically been recognized as one of the world's most unequal societies, which has experienced stabilized economic growth since 1994, yet its effect on inequality has been detrimental. Therefore, many studies have been performed to identify the determinants of income inequality, but few have analyzed it in the context of South Africa. As listed by the World Bank, South Africa has the highest Gini Coefficient (a standard measure of income dispersion) in the world being equivalent to 63%.

Hence, the main objective of conducting this empirical research is to investigate the key drivers behind South Africa's high inequality in income distribution as only few time series observations have been carried out in the context of a single country, especially for transitional economies.

Admittedly, nations effectively handling economic disparity have more harmonious civil and political communities, more consistent growth, and higher levels of happiness, as well as reduced crime and other forms of social instability. Therefore, investigating the root causes of income inequality in South Africa is essential to proposing policy actions that could lead to the economic welfare of the nation.

The rest of the paper is organized into 6 sections. Section 2 presents empirical review of literature while Section 3 focuses on the methodology exploited while conducting the empirical study. The empirical results on the determinants of inequality and five OLS assumptions are presented and discussed in Section 4 and 5 respectively prior to the final part, which provides concluding remarks.

Literature review

Several empirical studies have attempted to investigate the determinants of income distribution from various perspectives. The economic, social and demographic variables such as inflation, literacy rate, unemployment, education, government debt and expenditure, income tax and population growth have been shown in the literature to explain the dynamics of income inequality.

When there is inflation, price levels tend to outpace the increase in money wages. On this basis, it was concluded by several researchers such as Laidler and Parkin (1975) and Fischer and Modigliani (1978) that inflation contributes to high income inequality by hurting the poor relatively more than the affluent. Additionally, Ferreira and Litchfield (2000) discovered a strong positive relationship between inflation and inequality in Brazil. However, according to Ho-Yin Yue and Shatin (2011), inflation



reduces income inequality in two ways: first, by transferring income from nominal lenders to nominal borrowers; second, redistributing income through the tax system.

In a study on the correlation between income and literacy, Verner (2005) and Rehman (2017) found a weak, nonlinear relationship up to a threshold, which in turn led to a positive relationship having increased Gini and one's chances of increasing income as well as a country's quality of human capital.

Gonzalez and Menendez (2000) examined the correlation between unemployment and labor income inequality in Argentina (1991-1998) and discovered that unemployment accounted for 43 percent of the entire increase in disparity. In addition, Tregenna and Tsela (2008) also investigated a close positive relationship between income inequality and unemployment rate in South Africa.

Education has also played a fundamental role in the debate over how to reduce income inequality. Knight and Sabot (1983) and Crenshaw and Ameen (1994) confirmed the hypothesis stating that education is negatively correlated with income inequality while being statistically significant. On the contrary, Barro (1999) studied the impacts of various factors on income disparity and found out a negative relationship for primary education attainment, while for higher education it was positive.

Sezgenç (2019) found out that public debt has a strong negative effect on income distribution using OLS method for the panel data from 1990 to 2016. Similarly, Tung (2020) also ascertained the negative correlation between these variables in 17 developing countries.

The research work carried out by Calderon and Serven (2004) utilized panel data on Latin American countries to examine whether income inequality is influenced by government expenditure, and concluded that state-owned infrastructure initiatives serve as the main source of fairly-paid employment among low-skilled labor force, thereby leading to reduced income inequality.

Moreover, Joumard et al. (2012) discovered that taxes on income led to 25% decrease in income disparity if measured through the concentration coefficient by about 25% in the late 2000s. This finding was supported by Leibbrandt and Wegner (2011) who investigated that direct taxation in post-apartheid South Africa contributed to a decrease in income inequality levels.

Finally, Moller et al. (2009) and Rey (2018) investigated whether a rapid rate of population growth could lead to increased levels of inequality as economies with various segments of labor force involve both highly skilled and unskilled workers, resulting in a joint migration stream that causes the degrees of income inequality and occupational differentiation to expand.



Methodology and Data Description

The present empirical research used the annual time series data which involves the period of 50 years spanning from 1970 to 2020. All the selected variables in the model were adopted from the World Development Indicators, which is believed to contribute to the validity of the empirical findings. Moreover, the equation is estimated using ordinary least square (OLS) technique with the selected data on South Africa. Since the focal aim of this study is to investigate the leading causes of income inequality in South Africa, “Gini Index”, which measures the area between the equality line and the Lorenz curve, was selected as the measurement of income inequality as well as the dependent variable, and 8 explanatory variables were tested throughout the process of the regression analysis.

As there is a linear relationship between the exogenous and each of the chosen indigenous variables, the empirical study employs the following equation:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7 + \beta_8x_8 + E$$

Where y denotes “Gini Coefficient”, x_1 -inflation rate, x_2 – literacy rate, x_3 - unemployment, x_4 -education, x_5 – government debt, x_6 –government expenditure, x_7 – taxes and x_8 – population growth.

Research Hypothesis:

Based on the empirical literature, this study proposes a model to analyze the drivers of South African income disparity. The current study intends to examine the following hypothesis based on the objective:

H_0 : there is a significant relationship between all of the independent variables and the dependent variable

H_a :there is a significant relationship between at least one of the independent variables and the dependent variable

The table below demonstrates more detailed description of each variable used for the analysis along with the anticipated results of the selected independent variables on income inequality based on the findings of previously conducted research studies.

Table 1. Descriptive Statistics



Variable	Obs	Mean	Std. Dev.	Min	Max
<i>giniindexw~g</i>	50	.6119388	.0228477	.57	.68
<i>inflationc~c</i>	50	8.952703	4.41351	-.6920301	18.65493
<i>literacyra~g</i>	50	91.30308	2.885515	84.04667	95.02297
<i>unemploye~r</i>	50	29.64433	3.237229	22.41	34.18036
<i>educationa~h</i>	50	3.145396	2.930658	-1.93478	8.24519
<i>centralgov~g</i>	50	38.96464	5.243882	28.13011	44.86862
<i>generalgov~o</i>	50	16.63861	2.364931	11.11166	20.65017
<i>taxesoninc~a</i>	50	58.32835	4.021355	51.62518	70.19032
<i>population~w</i>	50	1.977093	.5725277	1.217762	2.758618

Table 2. Variable Description

<i>Variable name</i>	<i>Explanation of the variable</i>	<i>Expected results</i>	<i>Data source</i>
<i>Gini Coefficient</i>	Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality.	Dependent variable	World Development Indicators
<i>Inflation</i>	Inflation, consumer prices (annual %)	Positive and statistically significant	World Development Indicators
<i>Literacy rate</i>	Literacy rate, adult total (% of people ages 15 and above)	Positive and statistically significant	World Development Indicators
<i>Unemployment</i>	Unemployment, total (% of total labor force) (national estimate)	Positive and statistically significant	World Development Indicators
<i>Education</i>	Educational attainment, at least Bachelor's or equivalent, population 25+, total (%) (cumulative)	Negative and statistically significant	World Development Indicators
<i>Government debt</i>	Central government debt, total (% of GDP)	Negative and statistically significant	World Development Indicators
<i>Government expenditure</i>	General government final consumption expenditure (% of GDP)	Negative and statistically significant	World Development Indicators
<i>Taxes</i>	Taxes on income, profits and capital gains (% of total taxes)	Negative and statistically significant	World Development Indicators
<i>Population growth</i>	Population growth (annual %)	Positive and statistically significant	World Development Indicators



Results and Discussion

This section of the paper provides the empirical results obtained through running a multiple linear regression model with the use of “STATA 14” software, followed by detailed interpretations and discussions of the results on the relationship between income inequality and each of the eight independent variables.

Table 3. OLS Estimation results.

Number of observations	50
<i>F</i> (8, 9)	3.80
<i>Prob</i> > <i>F</i>	0.0315
<i>R-squared</i>	0.7714
<i>Adj R-squared</i>	0.5683
<i>Root MSE</i>	.01564

	Coefficient	St. Error	t	P> t	[95% Conf. Interval]	
<i>giniindexworldbankestimatesipovg</i>						
<i>inflationconsumerpricesannualfpc</i>	.0057759	.0019482	2.96	0.016	.0013688	.010183
<i>literacyrateadulttotalofpeopleag</i>	-.0061739	.001935	-3.19	0.011	-.0105512	-.0017967
<i>unemploymenttotalofiotallaborfor</i>	-.0047316	.0018811	-2.52	0.033	-.008987	-.0004761
<i>educationalattainmentatleastbach</i>	.018622	.0054688	3.41	0.008	.0062507	.0309933
<i>centralgovernmentdebttotalofgdp</i>	-.0041432	.0025373	-1.63	0.137	-.0098829	-.0015966
<i>generalgovernmentfinalconsumptio</i>	-.0177012	.0072521	-2.44	0.037	-.0341066	-.0012957
<i>taxesonincomeprofitsandcapitalga</i>	-.0028855	.0019489	-1.48	0.173	-.0072941	.0015231
<i>populationgrowthannualsppopgrow</i>	-.0080561	.0231399	-0.35	0.736	-.0604023	.0442901
<i>_cons</i>	1.850252	.344308	5.37	0.000	1.071373	2.629131



The table provided above shows that the time series data involved 50 observations in total. As regards the R-squared, it produced quite a high value of 0.7714, meaning 77.14% of the variation in Gini could be explained by the variations in independent variables. Based on 95% confidence interval, the independent variables whose “p” values were equal or less than 0.05 were considered statistically significant in this particular study.

To start with, out of 8 variables tested, 5 of them were found to form a high correlation with income inequality dynamics of South Africa which are inflation, literacy, unemployment, education and government expenditure. The estimates suggest that 1% increase in inflation increases the Gini by 0.5% meaning that there is a positive relationship between these variables which is in conformity with the expected result and the findings of Ferreira and Litchfield (2000), Laidler and Parkin (1975) and Fischer and Modigliani (1978). Furthermore, the result of educational attainment did not have the expected sign which was negative, as it shows that 1% rise in the population with high education leads to a 1.8% increase in income disparity while being in line with the research of Barro (2000) who claimed that higher education makes a worker more skilled and productive thereby contributing to a higher wage. However, the result reveals that unemployment in the case of South Africa is negatively associated with the Gini coefficient which contradicts the results of Gonzalez and Menendez (2000) and Tregenna and Tsela (2008). Regarding the impact of government spending, it was also found to be negatively correlated with the Gini index such that 1% increase in the expenditure yields an average of 1.7% reduction in income inequality. The same result was reached by Calderon and Serven (2004) who indicated that a rise in government debt may increase the income of the workers, resulting in redistribution from capitalists to the workers and decrease in personal income inequality. Finally, yet importantly, 1% rise in the percentage of people who are literate causes income inequality to reduce by 0.6%, which is not supported by the results obtained by Verner (2005) and Rehman (2017). The estimated results show that public debt, population growth and tax rate are statistically insignificant in explaining South African income distribution.

OLS assumptions

This section of the paper aims to demonstrate and interpret the empirical findings of a number of diagnostic tests that have been performed to observe whether five of the required OLS assumptions are satisfied or not.

Heteroskedasticity



In order to check for heteroscedasticity (unequal variance of the error terms), Breusch-Pagan/Cook-Weisberg test has been performed.

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
H_0 : Constant variance	
Variables: fitted values of giniindexworldbankestimatesipovg	
Chi2 (1)	= 2.14
Prob> chi2	= 0.1431

While the null hypothesis used in this diagnostic test states that error terms/residuals possess constant variances, the findings of the test suggest that due to a p-value of 0.1431 being larger than chi-test (0.05, significance level of 95%), we fail to reject H_0 and it can be inferred that there does not occur the problem of heteroscedasticity in the regression analysis.

Multicollinearity

Variable	VIF	1/ VIF
<i>educationa-h</i>	21.60	0.046296
<i>generalgov~o</i>	18.43	0.054263
<i>population-w</i>	11.94	0.083753
<i>centralgov~g</i>	11.52	0.086802
<i>inflationc~c</i>	5.23	0.191192
<i>taxesoninc~a</i>	4.55	0.219960
<i>unemploye~r</i>	3.06	0.326267
<i>literacyza~g</i>	1.75	0.572535
<i>Mean VIF</i>	9.76	

As regards multicollinearity, the current study aims to examine the mean change in the selected dependent variable caused by one independent variable while holding others constant. Meanwhile, the essential condition for no multicollinearity states that the mean vif (variance inflation factor) needs to be smaller than 10, which in this case is equal to 9.76, suggesting that the regression analysis is not prone to multicollinearity.

No autocorrelation



	<i>giniin-g</i>	<i>inflat-c</i>	<i>litera-g</i>	<i>unempl-r</i>	<i>educat-h</i>	<i>centra-g</i>	<i>genera-o</i>	<i>taxeso-a</i>	<i>popula-w</i>
<i>giniindexw-g</i>	1.0000								
<i>inflationc-c</i>	-0.3747	1.0000							
<i>literacyra-g</i>	-0.1920	0.1884	1.0000						
<i>unemploye-r</i>	-0.4164	0.5792	-0.2319	1.0000					
<i>educationa-h</i>	0.4860	-0.6974	0.1376	-0.7553	1.0000				
<i>centralgov-g</i>	0.2717	-0.5696	0.0401	-0.7144	0.8823	1.0000			
<i>generalgov-o</i>	0.2183	-0.3960	0.1374	-0.6996	0.8548	0.8796	1.0000		
<i>taxesoninc-a</i>	-0.0984	0.3612	-0.0099	0.5168	-0.6260	-0.7569	-0.8227	1.0000	
<i>population-w</i>	-0.4395	0.7617	-0.1388	0.6810	-0.8954	-0.8278	-0.6499	0.4854	1.000

The correlation matrix was performed to examine whether there exists a linear relationship between independent variables (autocorrelation). If the coefficients (between two independent variables) shown in the correlation matrix ranges between 0.9 and 1.0 or -0.9 and -1.0, there appears a high level of positive or negative autocorrelation respectively. In other words, the closer the coefficients are to 0, the lower the level of autocorrelation. Based on the correlation matrix shown above, it can be concluded that no highly correlated independent variables exist in the regression analysis with the exception of such linkages between population growth and educational attainment (-0.8954) or government debt and educational attainment (0.8823).

Zero mean value of the error term

Variable	Obs	Mean	Std. Dev.	Min	Max
sumres	50	11.74251	0	11.74251	11.74251

Similar to the aforementioned OLS assumptions, the diagnostic test performed with the help of “sum of residuals also satisfies the condition for the zero mean value of the error term. In order for the regression model to be unbiased, the average coefficient of the error term/residuals must be equal to 0, which is well explained by std. dev. on the table shown above.

Zero covariance between the error term and regressors

Ramsey RESET test using powers of the fitted values of <i>giniindexworldbankestimatesipovg</i>	
H_0 :	model has no omitted variables
F (3, 6)	= 1.92
Prob > F	= 0.2280

Last but not least, testing for omitted variable bias is essential to fulfill the condition



of zero covariance between the error term and the independent variables.

The H_0 generated through Ramsey RESET test states that the model has no omitted variables with a p-value of 0.2280 being larger than 0.05 suggests that we fail to reject H_0 and conclude that no variables are missing in the model. Hence, it can also be inferred that there is no covariance between residuals and regressors.

Conclusion

A great deal of research has been carried out on income inequality and its determinants, yet only few have associated it with the case of South Africa. Hence, the main objective of the current paper has been to identify the main factors of income inequality in the South Africa for the period 1970-2020. The empirical findings of the econometric study discovered that inflation, literacy, unemployment, education and government expenditure serve as the most statistically significant variables, whose increase leads to a bigger income differentiation. Among these variables, educational attainment has been found as the most influential factor on the increase in income inequality in South Africa, whereas higher government expenditure is an effective tool in reducing income disparity. As for the rest of the variables examined in the model, such as government debt, taxes and population growth, all of them have been found to be insignificant. Finally, by rejecting the null hypotheses, we accept the alternative hypotheses since 5 variables out of 8 were found to have a significant impact on the income inequality in South Africa.

Reference list

- ❖ Barro, R. (1999). *Human Capital and Growth in Cross Country Regressions. Swedish Economic Policy Review*, 6, 237-277. - References - Scientific Research Publishing. Available from:



[https://www.scirp.org/\(S\(i43dyn45teexjx455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=1306834](https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1306834)

- ❖ Calderón, C. and Servén, L. (2004). *The Effects of Infrastructure Development on Growth and Income Distribution. The World Bank, Washington DC. - References - Scientific Research Publishing.* Available from [https://www.scirp.org/\(S\(lz5mqp453edsnp55rrgjt55.\)\)/reference/referencespapers.aspx?referenceid=2688444](https://www.scirp.org/(S(lz5mqp453edsnp55rrgjt55.))/reference/referencespapers.aspx?referenceid=2688444)
- ❖ Crenshaw, E. and Ameen, A. (1994). The Distribution of Income Across National-Populations: Testing Multiple Paradigms. *Social Science Research*, 23(1), pp.1–22.
- ❖ Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F. and Tsounta, E. (2015). Causes and Consequences of Income Inequality: A Global Perspective. *Staff Discussion Notes*, 15(13), p.1. Available from: <https://www.imf.org/external/pubs/ft/sdn/2015/sdn1513.pdf>
- ❖ Ferreira, F.H.G., Leite, P.G. and Litchfield, J.A. (2008). THE RISE AND FALL OF BRAZILIAN INEQUALITY: 1981–2004. *Macroeconomic Dynamics*, 12(S2), pp.199–230.
- ❖ Fischer, S. and Modigliani, F. (1978). *Towards An Understanding of the Real Effects and Costs of Inflation.* National Bureau of Economic Research. Available from: <https://www.nber.org/papers/w0303>
- ❖ González, M. and Menendez, A. (2000). *The Effect of Unemployment on Labor Earnings Inequality: Argentina in the Nineties**. Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.536.8968>
- ❖ Joumard, I., Pisu, M. and Bloch, D. (2012). Tackling income inequality: The role of taxes and transfers. *OECD Journal: Economic Studies*, 2012(1), pp.37–70. Available from: <https://econpapers.repec.org/article/oececokac/5k95xd6l65lt.htm>
- ❖ Knight, J.B. and Sabot, R.H. (1983). Educational Expansion and the Kuznets Effect. *The American Economic Review*, 73(5), pp.1132–1136. Available from: <https://www.jstor.org/stable/1814679?seq=1>



- ❖ Laidler, D. and Parkin, M. (1975). Inflation: A Survey. *Economic Journal*, [online] 85(340), pp.741–809. Available from:
https://econpapers.repec.org/article/ecjeconjl/v_3a85_3ay_3a1975_3ai_3a340_3ap_3a741-809.htm

- ❖ Leibbrandt, M., Wegner, E. and Finn, A. (2011). *The Policies for Reducing Income Inequality and Poverty in South Africa*. Available from:
https://ideas.repec.org/cgi-bin/get_doc.pl?urn=RePEc%3Aldr%3Awpaper%3A64&url=http%3A%2F%2Fopensaldru.uct.ac.za%2Fbitstream%2Fhandle%2F11090%2F79%2F2011_64.pdf%3Fsequence%3D1

- ❖ Moller, S., Alderson, Arthur S. and Nielsen, F. (2009). Changing Patterns of Income Inequality in U.S. Counties, 1970–2000. *American Journal of Sociology*, 114(4), pp.1037–1101.

- ❖ Rehman, S. and Rashidi, M.Z. (2017). The relationship between education and income inequality (the case of Pakistan). *International Journal of Accounting and Economics Studies*, 5(1), p.55.

- ❖ Rey, S.J. (2016). Bells in Space. *International Regional Science Review*, 41(2), pp.152–182.

- ❖ Sezgenç, Ş. and SezginŞ. (2020). Poverty and Income Distribution: Panel Data Analysis with Selected Upper-Middle Income Countries. *Equinox Journal of Economics Business and Political Studies*, [online] 7(2), pp.177–192. Available at:
<https://dergipark.org.tr/tr/pub/equinox/issue/56841/695332>

- ❖ The World Bank (2020). *World Development Indicators | DataBank*. Worldbank.org. Available from: <https://databank.worldbank.org/source/world-development-indicators>.

- ❖ Tregenna, F. and Tsela, M. (2008). *INEQUALITY AND ECONOMIC MARGINALISATION Inequality, unemployment and poverty in South Africa*. Available from:
http://www.tips.org.za/files/u65/inequality_unemployment_and_poverty_-_tregenna_and_tsela.pdf.



- ❖ Tung, L.T. (2020). Can public debt harm social development? Evidence from the Asian-Pacific region. *Journal of International Studies*, 13(2), pp.48–61.
- ❖ Verner, D. (2005). *What Factors Influence World Literacy? Is Africa Different?* papers.ssrn.com. Available from:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=650366
- ❖ Yue, H.-Y. (2011). *INCOME INEQUALITY, ECONOMIC GROWTH AND INFLATION: A STUDY ON KOREA*. Available from:
<https://eidmon.files.wordpress.com/2012/11/infla.pdf>