Developing Country Studies
ISSN 2224-607X (Paper) ISSN 2225-0565 (Online)
Vol.13, No.1, 2023

Determinants of Income Inequality in Kobo Town

Tewodros Zegeye Belay
MA in Development Economics, Senior Banking Officer at Commercial Bank of Ethiopia

Abstract
Income inequality is when individuals living in a given country earn disproportionate income levels. This accounts for the presence of high poverty, especially for households with lower incomes. Identifying the factors determining the existence of inequality is essential to reduce income inequality. With this objective, this study analyzes the determinants of income inequality in kobo town. To explore the determinants of income inequality, primary data obtained from the sample households for the city is applied. The inequality situation in this town is analyzed using income distributions, proving that income inequality in the city is high, and the Gini index of income is 0.38. then, marginal effects from the ordered logit model proved that households with houses have a higher probability of being in a high-income group than those who have lived in a rented house to be in a high-income stream. The effects of continuous variable reported as high education level is associated with a high probability of being a higher income earner.
Keywords: Egalaterian Line, Gini Index, Income Inequality, Lorenz Curve, Ordered Logit, Kobo Town
DOI: 10.7176/DCS/13-1-01
Publication date: January 31st 2023

Introduction

Income inequality is one of the major development issues throughout the world and a problem for both advanced and non-advanced countries. Classical economists like David Ricardo, Jevons, Baumol, Samuelson, and Stopler first attempted it. Those classical economists focused their attention on how the total income will be divided between the rewards for factors of production, which are wage, rent, profit, and interest. However, they only related the inequality concept with economic variables of growth and poverty in the 1955 work of Simon Kuznets. While in a year, Kuznets developed an inverted U shape Kuznets curve that shows the prevalence of a direct relationship between economic growth and income inequality. At the early stage of Agrarian economic development, there is a high level of economic growth without experiencing high-income dispersion after development (Kuznets, 1955). By using this model as a reference point, many empirical studies are conducted to prove the applicability of the Kuznets curve in their own respective countries or regions.

A high level of income disparity is prevalent in countries found in the region of Sub-Saharan Africa, and the most unfair income distributions in the world are found in Sub-Saharan Africa (UNDESA, 2009).

Several works in income inequality literature emphasize the issues of the relationship between income inequality and economic growth but need to consider the causes of income inequality. Hence, many research papers are conducted to analyze differences in income between countries and fail to consider determinants of income inequality within the same country. The determinants of income distributions within countries needed attention. Albeit much has been written on the subject of income inequality, empirical work in the area has traditionally focused on accounting for trends in household-level data within specific countries and instead tries to answer the question of determinants of microeconomic perspectives. Many empirical studies show that income inequality is more vast in the least developed countries than in the advanced countries (Anthony et al., 2009).

Ethiopia is experiencing a very high complex relationship between inequality and growth. Unlike other rapidly growing economies, the country has not experienced a remarkable increase in inequality, as measured by the Gini coefficient, while poverty reduction occurs rapidly. Structural transformation and poverty reduction require the implementation of reforms that would result in a rise in income disparities in addition to the expansion of the economy. Ethiopia's Gini coefficient is 0.33 world bank 2015. So this shows that Ethiopia is grouped among the major egalitarian countries in the world. This can be done through a visible expression of highly equal production distribution in rural areas because of land regulations that prevent farmers from consolidating land. Hence, most of the population still lives in rural areas; small rural Gini contributes to the low national measure. In the urban area, the reverse is true. After a decline in inequality between 2004 and 2010 (by 6.2 percentage points), most up-to-date developments indicate that the income gap is widening again. The key factors behind the increasing income disparities are the rising skill premia for advanced education and changes in household characteristics (younger households, living alone, or in couples with no children). (IMF, rep 2015)

Income inequality is not overblown in rural parts of Ethiopia due to annually earned agricultural income. Urban areas experienced high differences in living standards ranging from those who cannot fulfill their daily basic needs to wealthy individuals who are billionaires. Since Poverty reduction is the core objective of the Ethiopian government, it has implemented different poverty reduction policies and strategies to reach middle-income countries in the coming 20 years (GTP 2, 2015), and relative poverty reflects inequality in the

distribution of income. To overcome the inequality problem in urban areas of Ethiopia, first, studying the factors contributing to this income gap is important (MoFEC, 2016).

Income inequality shows individuals or households how far away from or nearest to the egalitarian equi-distribution line. It is a situation in which some individuals accumulate higher income which enables them to live a luxurious life, while others are in a difficult situation, even if they cannot fulfill their daily basic needs. This inequality is the means for the low level of welfare in society. It matters for growth as well.

When the population of the income shares of the top 20 %(the rich) increases, the GDP growth declines over the medium term; It states that the benefits do not trickle down. In contrast, an increase in the income share of the bottom 20 percent (the poor) is connected to higher GDP growth. The population's poor and middle-class matter a lot for growth across various interrelated economic, social, and political channels. Higher inequality results lowering growth by depriving of the ability of lower-income households of living healthily and accumulating physical and human capital (IMF, 2015). Increasing accumulation of incomes also reduces aggregate demand and slow growth, which is why the wealthy spend a lower fraction of their incomes than middle and lower-income groups. To reduce this unfair distribution of income between households, then studying and distinguishing factors that determine income inequality is necessary.

Research and development agenda has been given less attention to urban inequality in Ethiopia, particularly in medium towns like Kobo. Ethiopia's annual urban population growth rate is estimated to be above 4.3 % (PASDEP, 2006). In connection with this, income inequality in urban areas is growing, and its effect on urban poverty in developing countries like Ethiopia is very high. Hence, it is the means for low-level and unequal growth rates and spreading poverty in a given country; numerous papers worldwide have been conducted to identify the causes behind inequality in least-developed countries. The problem of inequality is severe in the least developed countries when compared and contrasted with the developed ones. Barro (2000), when we talk about the study of the determinants of income inequality and growth, was among the first to show a structurally different relationship between inequality and growth between developing and advanced ones. By dividing the sample into two groups, he found the relationship is structurally different. In higher-income developed countries, the low level of inequality may be associated with a higher growth rate. In contrast, higher income inequality is associated with lower growth in the least developed countries.

Several factors can influence inequality. For example, one would naturally think that the education level will account for income inequality. Certain social-economic groups do not have access to quality education in the United States, especially at the secondary school level. Those individuals with a low level of education are investigated to be low-income earners. There is much less income disparity in countries that provide higher-quality secondary education across the economic spectrum (WB, 2015). In addition to education, increased demand for high-skilled workers adds to a widening wage gap. Companies are investing more heavily in developing a high-skilled workforce, driving wages up for high-skilled workers. This leads to de-emphasizing or automating low-skilled functions, pushing wages for low-skilled workers down.

The inequality decomposition model and regression-based Decomposition of income inequality from household-based collected data have been used by Eskinder (2011) to see factors contributing to income inequality. Relevant variables which determine income inequality are identified. The analysis shows that the education level of households, Age, Public sector, and informal private worker employees have a significant and positive contribution to per capita consumption expenditure. In contrast, household size and age square significantly and negatively contribute. After controlling for the effects of unexplained variables, the Decomposition of factors has indicated that the household head's attributes account for one-third of the total income variance.

Since no previous research has been conducted to analyze the determinants of income dispersion in kobo town, this study tries to test the significance and impact of variables theoretically suggested as determinants of income inequality. Variables such as family size, experience, remittance, house ownership, dependency ratio, level of education, and the sectorial dummy for public and private workers are determinants of income inequality in kobo town. To prove the significance of variables, descriptive and econometric analysis methods were applied to the primarily collected data from randomly selected households.

Having the general objective of examining the problem of income inequality among the residents of Kobo town, this paper attempts to address pertinent questions to the subject, including what determines the existing inequality in the distribution of income in kobo town. What are the levels of income inequality in kobo town? Which socioeconomic features of households account for income inequality? Is there a causal relationship between human capital development and income inequalities in Kobo town? And their theoretical implications.

Literature Review and Theoretical Contexts

Inequality is when different people have different degrees of income or consumption. Income inequality indicates the extent to which individuals or households are dividing income between an unfair manner and the distribution of total national income among households in a disproportionate way.

Most inequality measures do not depend on the mean of the distribution, and this property of mean independence is considered a desirable property of an inequality measure (Gebeyaw, 2015). Instead, inequality is concerned with measuring distribution. There are two measures of income distribution: the personal or size distribution of income and the functional distribution of income. Size distribution refers to the distribution of income according to the size class of persons, like the share of total income accruing to the poorest specific percentage or the richest specific percentage of a population regardless of the sources of that income. Economists commonly use this measure to measure the magnitudes of inequality in the population. It deals with individual persons or households and their total incomes. How they received that income is not considered. What matters is how much each earns, irrespective of whether the income is derived solely from employment or other sources such as interest, profits, rents, gifts, or inheritance.

Moreover, the locational and occupational source of income is ignored. Whether the source of income is urban or rural or from the agricultural, manufacturing, or service sector needs to be considered. If two individuals receive the same personal income, they are classified together, irrespective of the fact that one may work more hours a day than the other. However, the second one collects interest in his inheritance. (Olivier, 2010)

Economists and statisticians, therefore, like to arrange all individuals by ascending personal incomes and then divide the total population into distinct groups or sizes. A standard method is to divide the population into successive quintiles (fifths) or deciles (tenths) according to ascending income levels and then determine what proportion of the total national income is received by each income group (Abebe, 2016).

The Lorenz curve

The Lorenz curve shows the actual quantitative relationship between the percentage of income recipients and the percentage of the total income received during a given year. The more the Lorenz line curves away from the diagonal, which shows the line of perfect equality, the greater the degree of inequality. The extreme case of perfect inequality is when one person receives all the national income. In contrast, everybody else receives nothing, which the congruence of the Lorenz curve with the bottom horizontal and right-hand vertical axes would represent. Because no country exhibits perfect equality or perfect inequality in its income distribution, the Lorenz curves for different countries will lie somewhere to the right of the diagonal. The greater the degree of inequality, the greater the bend and the closer to the bottom horizontal axis the Lorenz curve will be.

Gini coefficient (index)

The Gini coefficient measures the extent to which income and consumption expenditure distribution among individual households within an economy deviate from a perfectly equal distribution. A Lorenz curve plots the cumulative percentage of total income received against the cumulative number of recipients, starting with the poorest individual households. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality expressed as a percentage of a maximum area under the line. . Thus, a Gini of 0 represents perfect equality in the distribution of income, while a Gini index of 1 implies the presence of perfect unequal distribution.

General Entropy Index

The common inequality indicators mentioned above cannot be used to assess the significant contributors to inequality by different subgroups of the population and regions or by income source. The GE class has essential advantages:

First, they can be decomposed into within- and between-group inequality over space and time. The within-group inequality shows that overall inequality is attributed to the change in income distribution by considering the group as a population. The between-group inequality index helps to examine how much of the overall inequality is due to changes in the mean income of each group by assuming all members earn/consume the average amount equally. Second, different entropy class measures are sensitive to different parts of the distribution.

The Generalized Entropy class of indicators, including the Theil indexes, can be decomposed across these partitions in an additive way, but the Gini index cannot. To decompose Theil's T index (i.e., GE (1)), out of the total income of the population Y, Yi the income of a subgroup from the total population N, the following formula is necessary (Abebe, 2016).

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{N}\sum\left(\frac{yi}{y}\right)^{\alpha} - 1\right] \text{ Where } \alpha \neq 0,1$$

$$G(0) = \frac{1}{N}\sum \ln\left(\frac{yi}{y}\right) \text{ And}$$

$$GE(1) = \frac{1}{N}\sum\left(\frac{yi}{y}\right)\ln\left(\frac{yi}{y}\right)$$

Where GE (0) = mean log deviation measure, also called the Tails index

GE (1) = first-level general entropy index

The second standard measure of income distribution used by economists, the functional or factor share distribution of income, attempts to explain the share of total national income that each factor of production (land, labor, and capital) receives. Instead of looking at individuals as separate entities, the theory of functional income distribution inquiries into the percentage that labor receives as a whole and compares this with the percentages of the total income distributed in the form of rent, interest, and profit (i.e., the returns to land and financial and physical capital). Although specific individuals may receive income from all these sources, that is not a matter of concern for the functional approach.

A sizable body of theoretical literature has been built around the concept of the functional income distribution. It attempts to explain the income of a factor of production by the contribution that this factor makes to production.

Supply and demand curves are assumed to determine the unit prices of each productive factor. When these unit prices are multiplied by quantities employed on the assumption of efficient (minimum-cost) factor utilization, we get a measure of the total payment to each factor. For example, the supply and demand for labor are assumed to determine its market wage. When this wage is then multiplied by the total level of employment, we get a measure of total wage payments, also sometimes called the total wage bill. In this paper, theories on the distribution of income are explained into two broad categories of classical distribution theories of David Ricardo and the demand-pull theories of the income distribution.

The Ricardian theory

From a neo-classical point of view, the earliest type of reasonable wage circulation hypothesis in its propelled shape is found in David Ricardo's work. In the primary section of the introduction to his authoritative work, Ricardo (1817) illuminates the fundamental extent of his request "To decide the laws which direct this dissemination wage between rent, profit and wages, is the foremost issue in Political Economy." Ricardo states that the entire "creation of the earth" is separated by paying rent, taking after the standard of diminishing negligible profitability of land. The most profitable sections of land are being used, to begin with, says Ricardo, so that through time, less gainful sections of land remain, and the best sections of land are the priciest (Oliver, 2010).

The appropriation between the rest of the wages and profit is then made. Ricardo does not consider diminishing marginal productivity of labor and capital until the rise of marginalist schools creates minimal investigation. Instead, Ricardo supports Malthus' standard of the populace to build up a hypothesis that says "specialists will duplicate to make the wage bill grow faster than benefits." Accordingly, the offer of benefit will undoubtedly diminish in time and driving at last to a circumstance Ricardo called the stationary State. With benefits being pressed out of the framework. Industrialists would have no motivators to extend their actions, and economic development would arrive at an end. Ricardo's understanding reached is tested toward the finish of the nineteenth century. The new financial school's position, just later known as the marginalist or neoclassical school, is best abridged by two quotes. Stanley Jevons, one of the primary defenders of the new school in Europe, clarifies the extent of his new approach in an accompanying way: "The clear conclusion is that the main any expectation of accomplishing a simple arrangement of Financial aspects is to excursion aside, once and for ever, the mazy and ludicrous suspicions of the Ricardian School (Oliver, 2010). The fact of the matter is with the French School, and the sooner we perceive this reality, the better it will be for the world" (Jevons 1871 as referred to by Oliver (2010)). On the opposite side of the Atlantic, John Bates Clark, the originator of the American branch of neoclassical economists and "a focal figure in the development of the minimal profitability investigation of appropriation" (Baumol, 1985), states the finish of the approach as "It is the reason for this work to demonstrate that a characteristic law controls the dispersion of pay to society that this law

The marginalist revolution was born to escape a "mazy" political economy by replacing it with a more sophisticated economic science characterized by good mathematics. Ricardo's decreasing marginal productivity principle was generalized; all factors of production now had the same (diminishing returns) property and could be put under the umbrella of a unique production function. The most famous of those production functions is that proposed by Cobb and Douglas, $Y = AL^aK^b$, for which one verifies that the labor share W/Y is indeed the parameter a, which is assumed to be constant. The constancy of the labor share implies that workers are being paid real wages at the level of their labor productivity.

Three significant points are concluded from the analysis presented thus far. The first conclusion of the marginalist school is that everybody will be paid according to his/her contribution to the production process, which is acceptable to most economists at the current time. Modern economics is still under the influence of such a conclusion. This is undoubtedly a significant and reassuring idea that John Bates Clark was right. However, this implied that the inequality question was a non-issue: first, because it was embodied in a more extensive theory, that of the functional income distribution, and second because intrinsic fairness characterized such a system. Much of the literature on inequality as we know it today was indeed not born at that time, only much

later (Frankin, 2007).

The second conclusion is that the relative factor shares are constant through time. As such, this conclusion may appear striking as Solow himself expressed "skepticism" about it (Solow, 1958). Ricardo had already foreseen the way out of the stationary State and improved productivity or technology, but he still needed to develop the idea further in connection with income distribution. Several cases for productivity improvements have been proposed. Harrods-neutral technological change is an improvement affecting only labor productivity; Solow-neutral technological progress affects only capital productivity, while Hicks-neutral technological progress affects both factors (Giovannoni, 2006).

Corry (1966) attempts to summarize the effects of technological change on the relative shares of income in the neoclassical framework. His conclusions indicate that the inclusion of technological progress is undoubtedly a welcome development but that this did not lead to a significant leap forward in economic research:

"I have argued that the basic models can be made to fit the broad facts of history. But suppose the facts had been different, would we have rejected the models? I think not. These models are really better described as frameworks for handling the relative share problem and only make specific predictions with added restrictions. Thus, what I have called the neoclassical approach does not in general predict the course of relative shares. The State of economic knowledge does not enable us to predict the direction of impact of innovation on relative shares".

The third conclusion, shared by the Ricardian and marginalist tradition, is that everything is a matter of a simple universal law, a law of Nature. Such a law would be inescapable, only possible to be disturbed by technology ascribed to an exogenous factor, as reflected in the "Solow residual." Much of the later economic research was concerned with providing more details than just a natural law. This body of research was mainly influenced by Keynesian economics and amounted to a fundamental reformulation of the classical factor share theory.

Demand-Based Theories

Non-classical approaches need to touch base at comparative conclusions regarding natural laws. It concentrates on four unique methodologies, specifically those of Stolper and Samuelson (1941). The way persuades such a significantly limited decision that those are oft-cited hypotheses whose convenience will become more evident through whatever remains of this paper. We picked those speculations not because they are dissimilar complex to-accommodate hypotheses but because they have, in like manner, to relate practical circulation to request conditions in the economy individually the venture share, unemployment, macroeconomic strategies, and exchange separately.

As a starting point, it may be essential to note that those three approaches, despite being loosely referred to as "Keynesian," ought to be better called, at best, "heterodox," for Keynes said very little about income distribution. Keynes (1936) begins by expecting that "propose the actual wage is equivalent to the peripheral result of work. Keynes' adherents endeavored, pretty much rationally, to lay out the elements of a "Keynesian" hypothesis of pay circulation.

Starting with a model of salary circulation regularly disregarded (it brings bits of knowledge drawn from the universal exchange hypothesis). The center of the hypothesis comprises the models from Ricardo (1817), Heckscher (1919), Ohlin (1933), and Stolper and Samuelson (1941) (as referred to in Oliver, 2010). We talk about this assemblage of work first because it has established in neoclassical financial matters that every one of the four models lay on suspicion of impeccable rivalry and complete work of assets. However, the Stolper and Samuelson hypothesis expresses that exchange factors, imports, and fares are requested factors that impact salary appropriation. The piece comprises three stages. In the first place, Ricardo predicts that exchange will occur taking after similar preferences and that exchange evens out the relative costs of exchanged merchandise. Second, Heckscher and Ohlin take note that nations will send out merchandise that uses their copious component of creation because those items end up being less expensive. A bounteous work nation will trade concentrated work products. Third, Stolper and Samuelson demonstrate that relative value meetings will profit the proprietors of the plenteous component of a generation more than the proprietors of the rare creation variable. Subsequently, Stolper and Samuelson's commended article has pay circulation suggestions: not each monetary gathering benefits similarly from the exchange.

There are two lessons most applicable to our investigation into pay circulation. If Heckscher and Ohlin are correct, we ought to watch the meeting of element costs among exchanging countries. Second, more noteworthy exchange openness in more capital-escalated nations should prompt a lower wage share. Once more, those are hypothetical assumptions that are testable against the information. However, the outcomes ought to be deliberately translated. An adjustment in salary appropriation in nations whose generation structure is generally described by flawed rivalry (Krugman, 1979). Krugman cannot depict a Heckscher-Ohlin-Samuelson impact to the degree that their analysis is based on perfect competition in the labor market.

Another vital model of income distribution is that of Kaldor. Kaldor (1956, 1957, as cited by Oliver) spells

out six famous "stylized facts" to be understood as mainly prevailing over the long run: (1) labor productivity is constant, (2) capital productivity is constant, (3) constancy of the capital-labor ratio, (4) constancy of the distribution of income, (5) relative stability of the real interest rate and (6) there is a significant disparity of labor productivity growth between countries. The central feature of Kaldor's income distribution theory rests on algebra, allowing him to derive the famous equation, which states that the profit share is a function of the investment rate and of the capitalists' and workers' propensity to save. Kaldor was strongly criticized for this work because of the specific framework surrounding Kaldor's equation. His assumptions, such as the above, were deemed overly restrictive to the extent that they relegated Kaldor's conclusions to be valid only in the case of total employment. The impossibility of unemployment is certainly not a Keynesian feature. However, two other conclusions of Kaldor are accurate: capitalists get the profits they spend on investments.

Most importantly, Kaldor's algebra shows that income distribution relates to the investment rate, a specific demand variable. Thus we would expect, in empirical studies, to find the profit share to go up during booms and down during depressions, which is a proposition testable against the data. Goodwin also uses algebra, but it makes for a much more original model of income distribution. The model is that of a predator-prey, where the unemployment rate and the labor share chase each other, forming a circular phase diagram. The predator and the prey never completely exhaust each other, yet when the wage share gets "too high," it becomes detrimental to labor, and unemployment rises. The wage share acts as the predator until unemployment reduces wage costs, and the circle can take place another time. Like in the Kaldor model, Goodwin's model rules out the possibility of persistent unemployment: income distribution will eventually adjust to restore the correct level of employment somewhere around full employment; this is only a matter of time. The Goodwin model can be deemed such to the extent that unemployment is a starting point of Keynesian economics. However, unlike Keynes (1936), Goodwin's unemployment is not allowed to persist. The benefit of Goodwin's model is that, like Kaldor's, it relates income distribution to the business cycle and, more precisely, to a cyclical effect on unemployment. Thus, we should expect to see unemployment rise when the labor share is "above equilibrium," again a proposition we can take to the data.

The last model in terms of "natural laws" may be found in Kalecki (1954). Contrary to Kaldor and Goodwin, Kalecki analyzed imperfect competition from the beginning Competition and information is imperfect, so there is no natural tendency for the economy to converge towards any stationary state or entire employment situation. What does matter for Kalecki is the markup over production costs, the amount of unused capacity of labor and capital, and the economic policy of the State. The latter is central to the Kaleckian analysis, for it is the only way to restore full employment or a "fairer" income distribution if there is a political willingness to do so. The former is just as important Kalecki relates the markup to the degree of monopoly, to sure profitability "norms" which are primarily imposed in the banking sector, and to the relative power of labor unions, in a class struggle framework. Thus for Kalecki, the further we are from perfect competition, the more the income distribution is likely to be influenced by the economic policy if we consider the markup as exogenously given. As for the previous theories, Kalecki's position is testable against the data; do we observe income distribution affected by economic policies? The difficulty here lies in what measure to represent the stance of economic policy government spending and budget deficit.

Literature Review

Many Studies around the 1990s based on cross-sectional comparisons generally supported the hypothesis that income inequality initially increases and then declines as the economic growth of a country rises. Many recent studies, however, based on larger data sets, have consistently refuted the inverted-U curve of Kuznets on the ground that either there is no systematic relationship between the two or the relationship is country-specific.

Even if the Kuznets hypothesis failed to be universal, country-specific studies show mixed results, at least in the current situation. Okidi (2004) for Uganda, Fuwa (2003) for the Philippines, and Round (2001), using cross-country data from 35 African countries, all found a positive relationship between inequality and growth. Wan (2006), on the other hand, using micro data from china, and Iranian (2005), using 82 countries' cross-country evidence, found a quadratic relationship between inequality and growth, a rising inequality followed by declining inequality as per-capita growth over time. This result supports the Kuznets hypothesis.

Barro (1999), using a three-stage least squares estimator which treats the country-specific terms as random, finds that the effect of inequality on growth is negative in developing countries but is positive in rich countries, which is a real U-curve as opposed to Kuznets.

Authors Galor and Zeira (1993) deal with the Macroeconomy and the income distribution in the UK. This research investigates a theoretical linkage between income distribution and macroeconomics through investment in human capital. It also shows that income and wealth distribution are related to long-run macroeconomic issues, economic growth, and sectorial adjustment. The findings in this paper show that the distribution of wealth can significantly affect aggregate economic activity, both in the short and long run.

Afonso (2008) studied the determinants of income distribution and efficiency of public spending in

Bangladesh. The paper surveys the impact of public spending, education, and institutions on income distribution in advanced economies. They also assess the efficiency of public spending in redistributing income by using the Data Envelopment Analysis nonparametric approach. They see that public policies significantly affect income distribution through the social spending channel and indirectly using high-quality education/human capital and sound economic institutions.

Wan (2006) researched inequality –growth nexus in the short and long run: Empirical evidence from China. They argue that the conventional approach of data averaging is problematic for exploring the growth-inequality nexus. He introduces the polynomial inverse lag framework so that the impacts of inequality on investment, education, and, ultimately on, growth can be measured at precisely defined time lags.

Perdiz (2010) wrote about World's Growth and Inequalities. This article focuses on the relevance of the choice of measure (or meaning) of inequality. In the short term, economic growth may be accompanied by the simultaneous rise of some aspects of inequality and the fall of others. In the long term, economic growth will hardly cause a robust increase in inequality because inequality has reached historic highs.

Kookshin (1997) wrote about trends in and determinants of income distribution in Korea. He observes that disagreeing with the official statistics, the size distribution of income in Korea has not improved steadily since the late 1970s but deteriorated worst ever in the late1980s. The high rise of real estate prices, which caused an overall sense of relative deprivation, was also a significant root of worsened income distribution in the 1980s. Alejos (2003) observe the contribution of the determinants of income inequality in Guatemala. Their study decomposes income inequality in Guatemala into factors related to human capital, ethnic and gender discrimination, occupational structure, and non-labor income. Article results show a significant variation between the determinants at the national level, and in different socioeconomic groups, agriculture and livestock workers show as most.

Giovanni and Walter (2009) try to investigate the determinants of income distribution in the United Kingdom using a conditional distribution estimation approach. The results of this study show the effects of socioeconomic and demographic features of income inequality among households in the United Kingdom. Moreover, in this study, the inter-family difference is identified as the root cause of a high-income dispersion level at the household level.

Simon(2013) conduct a research paper on the main factors behind the high level of income inequality in sub-Saharan African countries using up-to-date panel data sets from countries between 1990 and 2010. Based on the results from random effect regression, variables such as government expenditure, the level of education, and the existence of democracy are investigated to be essential variables in reducing income inequality. However, foreign aid is found to increase income dispersion since it does not benefit poor households.

Ssewanyana et al. (2004) work on the determinants of income inequality in Uganda revealed that education level has a positive relationship with income level; male-headed households have more income when compared with female-headed households.in addition to this age of the primary breadwinner and household size have a negative relationship with households' income level.

Okidi (2004) wrote a paper entitled "Understanding the determinants of income inequality in Uganda." Their article is interesting because Uganda experienced gradual and sustained growth and poverty decline in the last ten years. The benefits of growth, however, are distributed in different ways. This study provides insights into a deepening understanding of the determinants of income inequality in Uganda. Decompositions by subgroups revealed that household characteristics are influential components of overall inequality, a finding also supported by the results based on the regression analysis.

Zelda (2004) utilizes the OLS estimation technique and regression-based decomposition method to investigate determining factors of income inequality in Botswana. The study uses primary data collected from the household survey in Botswana. The results provide an efficient way to quantify the roles of household variables on income inequality in a multivariate context. Results in this paper indicate that variables like secondary school education, training, value-added tax, number of children, and the number of adult working individuals in the household contribute significantly to income inequality among households in Botswana. On the other hand, variables such as age and primary education owning livestock from one to ten in number contribute a lot to reducing the household-level dispersion of income in Botswana.

The study carried out by Olawumi (2012) through the co-integration technique briefly shows which variables are relevant to affect income dispersion and which are not. The empirical findings in the study revealed that Gini Coefficient is very high in Nigeria, indicating a high level of income inequality. Also, the employment rate, inflation rate, Gross Domestic Product, and social spending were true determinants of income distribution in the Nigerian economy during the period under review (1977-2005). The study also found that both the growth rate of output and government health expenditure exhibited an inverse relationship with the Gini coefficient of income distribution in the Nigerian economy, while employment rate, inflation rate, and government education expenditure had a direct relationship with the Gini coefficient of income distribution in the Nigerian economy

Maina (2006) conduct a study to analyze the relationship between economic growth and income inequality

in Kenya. Using time series data from 1950 to 2006 and a simple OLS estimation technique, the study shows that income inequality, measured by the Gini coefficient, is negatively related to growth. This does not follow Kuznets's hypothesis since Kenya is a low-income country, which would simultaneously increase economic growth and income inequality. Our case is different. This can be explained by the social problems associated with inequality. These social problems include stealing, civil wars, and political instability.

Gebeyaw (2015) studied the link between financial institution development and income equality in Ethiopia. Econometrics from work (analysis) was employed as a general method of analysis using secondary data collected from the national bank of Ethiopia (NBE), central statistical authority (CSA), MOFED, and EEA for 33 years in the period of 1980-2012. The Engle Grander two steps procedures were followed to estimate the long-run and short-run parameters for the variables included in the model.

The empirical result of this study reveals that broad money, GDP, credit, and openness are the main determinants of income distribution (inequality) in Ethiopia in the long run. In the short run, only a broad money-to-GDP ratio is the primary determinant of income equality. Tassew (2009), in his study on poverty and inequality analysis in Ethiopia, found that even if income inequality remained unchanged in rural areas, there was a substantial increase in urban areas' income dispersion. Moreover, the results in this paper show that in Ethiopia, income growth reduces poverty, and increases in inequality increase poverty; the income-poverty elasticity lies in the range of -1.7 to -2.2. In rural Ethiopia, the increase in consumption has led to a reduction in headcount poverty.

Another study made by Ephrem (2006) on the analysis of Economic Growth, Income Distribution, and Poverty in Ethiopia using the Computable General Equilibrium Model uses time series data to prove the applicability of Kuznets's inverted U-shape hypothesis in Ethiopia. Moreover, the CGE model considers market interaction, that is, the effects of pricing outcomes of one market in other markets, and its effects, in turn, creating ripples throughout the whole economy, perhaps even to the extent of affecting the price-quantity equilibrium in the original market.

Abebe (2016) conducted a study using a general entropy index to analyze the determinants of income inequality among sampled households who find themselves at the bottom and top of the income or consumption distribution in urban centers of the South Wollo administrative Zone of Ethiopia. The findings of this study reveal factors such as; access to education, the informal sector economy and urban agriculture, job opportunities for youths, urban investment, and urban infrastructures on narrowing household income distribution in Kobo town.

Alemayehu and Tadele (2009) use SAM; Given the limitations of CGE models, econometric estimations are carried out to handle the household's agricultural production function using the Ethiopian Rural Household Survey (ERHS). Simulation exercises are performed to explore the impact of increasing agricultural productivity on poverty and inequality. Then, the mean income elasticity of the Foster-Greer-Thorbeck poverty indices is estimated to examine its poverty impact. Quality effects are explored by estimating the Gini index for rural and urban households using the Generalized Quadratic Lorenz curve.

Eskindir (2011) shows the binding effect of income inequality in poverty reduction using household-level data collected from the Bench-Maji zone, SNNP of southwest Ethiopia. Investigating the determinants of income inequality using an equality decomposition approach uses data collected from 120 sampled rural households living in the Shekou district of this zone. Finally, the result of this paper indicates that the Gini coefficient of the study area is 0.39, which shows that the income distribution in the study area is inequitable. The relative contribution of each source of income to income inequality is indicated as crop production at 0.35, livestock at 0.01, and nonfarm incomes at 0.03. The result shows that much income disparity is attributed to crop production. It was found that the other income sources have an inequality-decreasing effect; a raise in income from nonfarm income and livestock is favorable for income distribution. Landholding, land allocated for perennial crops, and livestock are household variables with higher inequality weight. An increase in education and livestock variables reduces the income gap, whereas land holding, land allocated for perennial crops & annual crops, and household size widen the gap. Concerned institutions in improving rural equity should pay great attention to nonfarm income-generating activities and improve livestock productivity.

Based on the above literature review and concepts and the relationship between income inequality and its determinants is shown below.

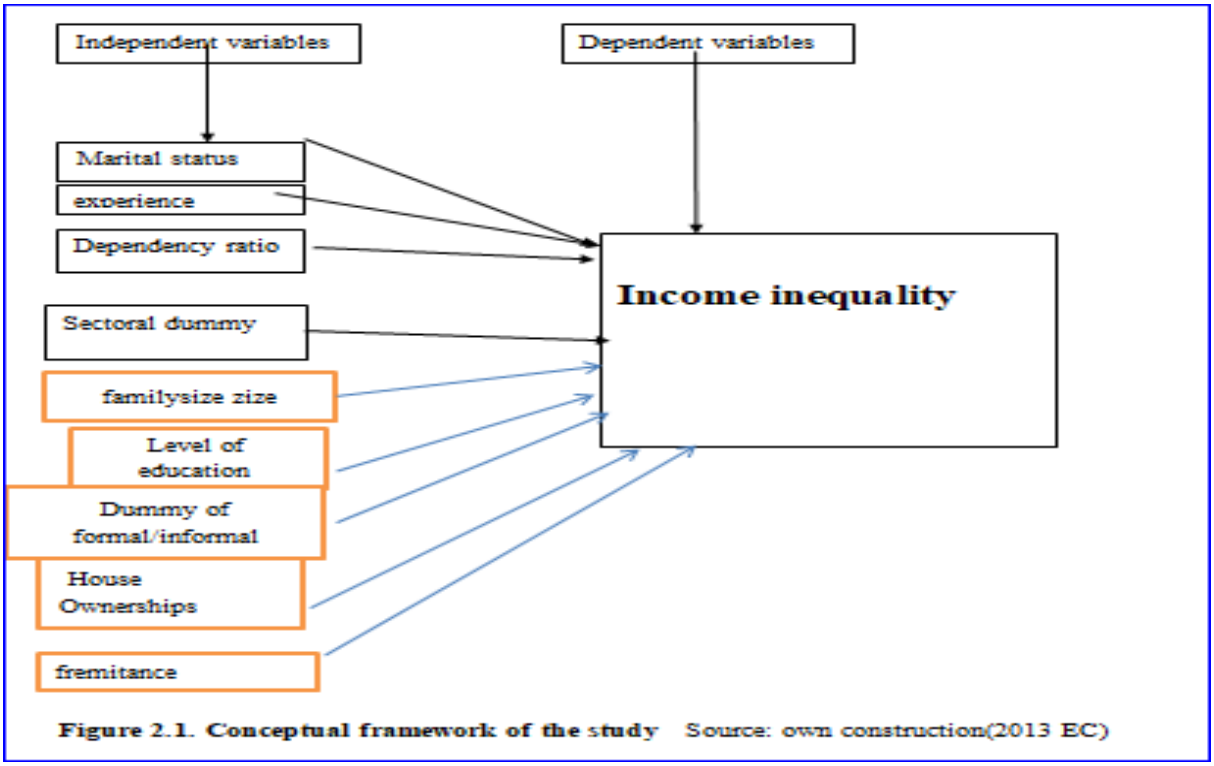


Figure 2.1. Conceptual framework of the study Source: own construction(2013 EC)

METHODOLOGY AND MODEL SPECIFICATION

Description of the study area Kobo town is located in Northeastern Ethiopia, 571km from Addis Ababa, the capital city of the Federal Democratic Republic of Ethiopia under the Amhara Regional state. The town has a longitude and latitude of 12°09′N 39°38′E with an elevation of 1468 meters above sea level. It is the administrative center of Kobo woreda. Kobo is part of a mid-altitude area between the Ethiopian Highlands to the west and the Afar Depression to the east. The town has one rural and four urban kebeles. The town's total population is estimated at 54782, comprising 15254 rural and 39528 urban dwellers (CCA,2013).

Ten percent of the people are engaged in agriculture; others are involved in non-agricultural activities such as petty trading, small-scale industries, handicrafts, and a small number of civil servants (kobo Micro and Small-Scale Enterprise Office, 2013).

In the educational sector, there are four kindergartens, eight primary schools, and three secondary schools; These 15 institutions have 11597 students. The town also has one technical and vocational education training center. Primary and secondary schools have about 487 teachers and 350 sections (CCA, 2013). When we see the quality of education, students-to-class ratio, students-to-teacher ratio, books, and availability of qualified professionals, they are generally in poor conditions. The health policy is being implemented by emphasizing the prevention of diseases mainly to prevent epidemic diseases in rural kebeles; trained health extension workers are assigned.

Generally, there is one hospital, one health center, and about 241 health professionals both in the hospital and. health center. Its coverage reached 63.7%; however, its quality is in question (KCA, 2013).

The city administration report of 2012 put the number of residential houses at about 6400. However, comparing the houses with the town's total population, the residential house problem is acute. Moreover, there is high unemployment, and most youths lack work-creating habits and entrepreneurial skills. In addition, there are limited affordable recreational and youth centers in the town. Hence, most youths often engage in unproductive harmful habits like chewing chat and smoking shisha. These and other problems aggravate the social problem of the town (KCA, 20113).

Sources of Data and Methods of Analysis

The type of data used in this study is primary data collected by purposive sampling techniques from the town. The fact that most town residents rely on rainfall agriculture contributes to insignificant income inequalities in the town. So 381 sample households are selected from four kebeles living in kobo town. Econometric data analysis methods are applied to study the determinants of income inequality. The significance and impact of variables are also analyzed by investigating the marginal effect and odds ratio obtained from the ordered logit model.

Sampling Design and Technique

The sample frame for this study is the four purposely selected kebeles of kobo town (kebele 1, 2, 3and 4) with a population size of 5894,9388,10455, and 13791 respectively, and a total population of four kebele 39514, around kobo 01 kebele have 15254 population(KCA 2013). The total four kebeles of the town contain their respective number of households. The first four kebeles, 1, 2, 3, and 4, contained 1179, 1877, 2091, and 2758 individual households, respectively. The method of sampling used in this study is stratified sampling and random sampling techniques. The four kebeles of the town are used as a stratum, and a systematic random sampling technique is applied to select households from each kebele.

Sample Size Determination

Sample size Selection needs care because using too large a sample size incurs high costs and a shortage of time. Moreover, selecting a smaller sample size does not genuinely represent the population's character and leads to biased results. This study relies upon Yamane's(1973) formula with a 95% confidence level to select a representative sample size from the population. The calculation formula of Taro Yamane is presented as follows.

$$n = \frac{N}{1+N(e)^2}$$

Where：

n= sample size required

N = number of people in the population

e = allowable error (%)

n=7906/1+7906(0.05)^2=381

n=381

Table 1: Sample size classification

Name of strata	Kebele 1	Kebele 2	kebele 3	Kebele 4	Total
Number of selected sample	58	90	100	133	381

Out of the total sample of 381 households, 58 HHs are selected from kebele 1, 90 from kebele 2, 100 from kebele 3, and 133 HHs from kebele 4.

Model selection and specification

Ordered logit Model

The Dependent variable in this model, income, is ordered to take four categories of low-income, lower middle-income, and higher middle-income and high-income group households. So the model that can be selected for this study is one of the discrete choice models presented below. All discrete choice models can be applied when the dependent variables take binary numbers 0 and 1. However, as the category of the dependent variables increases to more than two categories and the values of each category have a meaningful sequential order, the logit and probit models have to be ordered to take a multichoice form. When the dependent variable under the logit model is ordered to take more than two values, the model is known as the ordered logit or proportional odds model. The type of model preferred to be applied in this model is the ordered logistic regression model. Due to the mathematical convenience of logistic regression, the logit model presented above is applied in this study after ordering the dependent variable to take four categories.

Estimation and Inferences

Because of its non-linear characteristics, the best method of estimation for the discrete choice model, except for the linear probability model, is the maximum likelihood estimation technique. The aim of the maximum likelihood function is maximizing the likelihood, which is the probability of occurrence of an exciting event, as OLS estimation aims to reduce the residual sum of squared. Instead of following a normal distribution, each observation under the maximum likelihood estimation technique is treated as Bernoulli or binomial distribution. Using F (XB) probability of success and independent observations, the likelihood function can be formed as follows;

$$L(B) = \prod_{i=0}^{n} [F(XiB)^{yi}(1 - F(XiB))]^{1-yi}$$

By taking the natural logarithm of the likelihood function, it is possible to arrive at the log-likelihood function;

$$lnL = \sum_{i=1}^{n} [yiln(XiB) + (1 - yi)ln\ (1 - F(XiB)$$

Maximization of the likelihood function concerning the parameters B will give if the cumulative distribution function is logistic and probit estimates if the normal distribution is applied.

Pseudo R- squares measures will be applied with a multiple of responses. The formula for Pseudo R^2 is given below;

$$\text{Pseudo } R^2 = 1 - \frac{\varphi_{ur}}{\varphi_0}$$

Where; φ_{ur} is the log-likelihood function for the estimated model and
φ_0 is the log-likelihood function with only intercept.
If the covariates have no explanatory power, the ratio of the above two likelihood functions will become one, and the pseudo-R-squared becomes zero. However, the likelihood function for the estimated model is usually less than that of the likelihood function with only the intercept to imply that the pseudo-R-squared is more significant than zero.

Model Specification

In this study, disposable income is specified as a function of experience, dependency ratio, level of education, remittance, house ownership, and a dummy of a sector that the head of household is working, either public or private and either formal or informal, marital status, and family size. The ordered logit model is applied in this paper. The form of multinomial categories in which the dependent variable is non-linear and where four categories, namely: low-income, low-middle income, high-middle income, and high-income levels, are included under the dependent variable. This classification is based on the world banks' income thresholds in the least developed countries. Those households with a daily income of 1.90 are grouped under the low-income stream since they live below the poverty line. Those individual households with a total monthly income of less than 1300, between 1300 and 5000, between 5000 and 10000, and above 10000 are categorized to be low, lower-middle, upper- middle and high-income earners, respectively.[world bank atlas method 2015]

$Y = f(\text{family size, marital status, formal, dependency ratio, sector, level of education, house ownership , experience, remittance.})$

The ordered logistic regression model of income inequality is specified as follows:

Model

$$Inc = \beta 0 + \beta 1 experience + \beta 2 dr + \beta 3 edu + \beta 4 dpub + \beta 5 dform + \beta 6 fam.s + \beta 7\ mstatus + \beta 8 remitt + \beta 9 house + e$$

Where; inc = income of multichoice for lower, low middle, upper middle and high − income levels

dr = dependency ratio within households

experience = years of experience individuals have

edu = years of education or education level

dpri = dummy for sectors hired 1, for public
0, elsewhere

for = dummy for formal and informal workers 1, formal
0, otherwise

Fam s= the size of household members

Ms= dummy for marital status 1, married
0, otherwise

Remi=remittance 1, have remittance
0, elsewhere

house=house ownership 1, have own house
0, elsewhere

And then, the ordered logit model is specified as follows by stating first the critical assumptions of this model.
In the ordered logit model, there is an observed ordinal variable Y, a function of another continuous and unobserved latent variable Y^*, and the value of Y is determined by the latent variable Y*, which has various threshold levels.

Recall, $Yi^* = \beta 1 experience + \beta 2 dr + \beta 3 edu + \beta 4 dpub + \beta 5 dform + \beta 6 fam.s + \beta 7\ mstatus + \beta 8 remitt + \beta 9 house + e = X\beta + ei$, no intercept term is here

The three categories of the dependent variable Yi can be explained as follows;

$Yi = 1$, if $Yi^* \leq k_1$, where k_1 is the highest value of low income groups

$Yi = 2$, if $k_1 < Yi^* < k_2$, where k2 is the maximum value in lower medium income threshold

$Yi^* = 3$, if $k_2 < Yi^* < k_3$, where k3 when k3 is the maximum value of higher middle income group

$Yi^* = 4$ if $k3 < Yi^* < k4$ when k4 is the highest value of high income group

$Yi^* = n$ when $K_{n-1} < Yi^* <$ on when k_n is the maximum income in high-income group HHS

In general, $Y_i^* = \sum_{k=1}^{n} (\beta_k X_{ki} + ei)$, Ei follows the logistic distribution

After considering all the above assumptions, the logistic regression is formed below.

The general formula for the ordered logit model as presented by (Richard, 2015) with M categories

$$p(Y = M/x = \frac{\exp(Xi\beta - K_{M-1})}{1 + \exp(Xi\beta - k_{M-1})}$$

$$p(Y = 1/x) = 1 - \frac{\exp(Xi\beta - k1)}{1 + \exp(Xi\beta - k1)} = \frac{1}{1 + \exp(Xi\beta - k1)}$$

$$p(Y = 2/x) = \frac{1}{1 + \exp(Xi\beta - k2)} - \frac{1}{1 + \exp(Xi\beta - k1)}$$

$$p(Y = 3/x = \frac{1}{1 + \exp(Xi\beta - k2)}$$

The coefficient of the explanatory variables cannot be interpreted as the OLS technique. The marginal effects and the odds ratio are interpreted as probability changes.

Data Presentation, Analysis, and Results

This study adopts different data analysis methods such as standard deviations, graphs, percentages, means, variance, ratios, and figures to analyze the characteristics and natures of the data.

Description of Dummy variables

Table 2: summary statistics on dummy variables,

Explanatory variables	Type of variable with category	Several obs.	Frequency	Percentage
Marital status	Married=1 single=0	381	287 94	75.33 24.67
Formality of sector	Formal=1 Informal=0	381	228 153	59.84 40.16
Sectorial dummy	Public sector=1 Private sector=0	381	180 241	47.24 52.76
Remittance	Have remittance=1 No remittance=0	381	108 273	28.35 71.65
House ownership	Own house=1 Rent=0	381	233 148	61.15 38.85

(own source computation using stata)

The above table shows the descriptive of dummy variables used as explanatory variables in the model. 381,75.53% of the households are married, and the remaining 24.67% are divorced, widowed, or unmarried. When we see the sources of remittance, 28.35% of households have sources of remittance, and 71.65 % have no source of remittance.

Descriptive statistics of continuous variables

Table 3: Summary of continuous variables

Variables	Type	Mean	Var.	Std.	Min	Max	skewness	Kurtosis
Experience	Continues	10	55.6	7.46	0	30	1.16	3.62
Edu	Continues	9	37.4	6.11	0	19	-0.22	1.57
Family size	Continues	4.5	5.13	2.26	1	13	0.38	2.95
Dr	Continues	55.8	2573.75	50.73	0	250	1.02	3.92

Source (own computation using stata)

As indicated by the above table, the average experience of households is estimated to be 10 with 0 and 30 minimum and maximum ages, respectively. The experience of households varies with the variance and standard deviations of 55.6 and 7.46 numerical figures, respectively. The distribution of experience is measured by the skewness and kurtosis measures. 1.16% of the population is skewed to the right, meaning more population's experience is greater than the mean age of 10 years. Kurtosis measures, If any data set's kurtosis is less than three and it is a platikurtic distribution, the distribution produces fewer and less extreme outliers than the normal distribution.

The household's level of education is ordered by the years of schooling they spend their time on education, starting from the illiterates to those who have a doctorate. 0 years illiterate, 1-8 years elementary,9-12 years high school,13-17 diploma, degree, and masters. The average education level of households is a diploma and above. The variance and standard deviations are 37.4 and 6.11, and the kurtosis and skewness distribution measures indicate that about -0.22% of the total populations are below high school drops with a platikurtic distribution of 1.57.

The family size of the households ranges from 1 to 13, with the average household members of five individuals. The family size varies from household to household. The variance of 5.13, out of which 38% of the population have individuals more than the average number of 5 members with a platikurtic distribution of 2.95

A summary of the dependency ratio of the town shows that, on average, one individual is dependent on two working individuals as measured by a mean dependency ratio of 55.8, even if the dependency ratio varies from zero to 250 with variance and standard deviations of 2573.75 and 50.73. In addition, 1.02% of the households face a dependency ratio higher than the mean average value of 54.78, as measured by 1.02 levels of skewness. Unlike other variables described above, the kurtosis of dependency ratio is greater than 3 (3.92) to indicate the presence of leptokurtic distribution, which measures more out layers than the normal distribution.

The Description of the Distribution of Income

The income distribution can be analyzed using descriptive summary measures such as mean, standard deviation quartile, percentile, and decile ratios, and the Lorenz curve and Gini indexes. This part applies the percentile ratio as a powerful distribution analysis tool. Based on the table presented below, income distribution is summarized using the percentile ratio of income in kobo town.

Based on the data from households, 50% of the individual households of kobo town earns an average monthly income of 8500, as indicated by the fifty percentile ratio. The mean income value is 10449.74 even if an individual household's total income varies from 1000 to 32000 with a standard deviation of 7314.

The income distribution is skewed to the right by 1.22%. This indicates that only one up to two percent of the population earns income more significant than the mean value. The kurtosis for income measures a value of 3.93, which indicates the presence of leptokurtic distribution that measures more out layers than the normal distribution.

Table 4: Income distribution of the bottom 25% and top 25%.

Percentiles	Smallest			
1%	1000	1000		
5%	2000	1000		
10%	3000	1000		
25%	5000	1000		
		Obs = 381		
50%	8500	Mean	10449.74	
	Largest	Std. Dev.	7314.821	
75%	13000	32000		
90%	22000	32000	Variance	5.3507
95%	27000	32000	Skewness	1.222133
99%	32000	32000	Kurtosis	3.933648

Source: own computation using stata

According to table 4.1 above, the poorest 1%, 5%, 10%, and 25% percent of the population earn an average monthly income of 1000,2000,3000, and 5000, respectively. This means that the total income share of the bottom 25% of the population is only 3.93% of the total income.

In opposite to this small proportion of the top income earners are earning the highest proportion of the total income. The wealthiest 1%, 5%, 10%, and 25% of the total population are earning an average income of 32000, 27000, 22000, and 13000, respectively, and this shows that there is a high level of income inequality in kobo town where few top income earners are exploiting the income of the majority of the poor.

In addition to the above distribution measurements, the Lorenz curve and Gini index measures of inequality are applied, which do not depend on the mean of the distribution. Instead, inequality is concerned with distribution. In order to measure the income of households' inequality in the study area, the Lorenz curve and Gini index/coefficient are used.

The Lorenz Curve and Gini Index of Income

The Lorenz curve is one measure of income inequality, indicating how small amount the distribution is far away from the equality line. Any income distribution with a Lorenz curve near the equality line represents relatively equal income distribution. If the Lorenz curve for a given distribution is far away from the line of equality, the distribution is unequal. As indicated in the Lorenz curve graph (fig 4.1), the income distribution in this study area is high, as indicated by the downward bending curve.

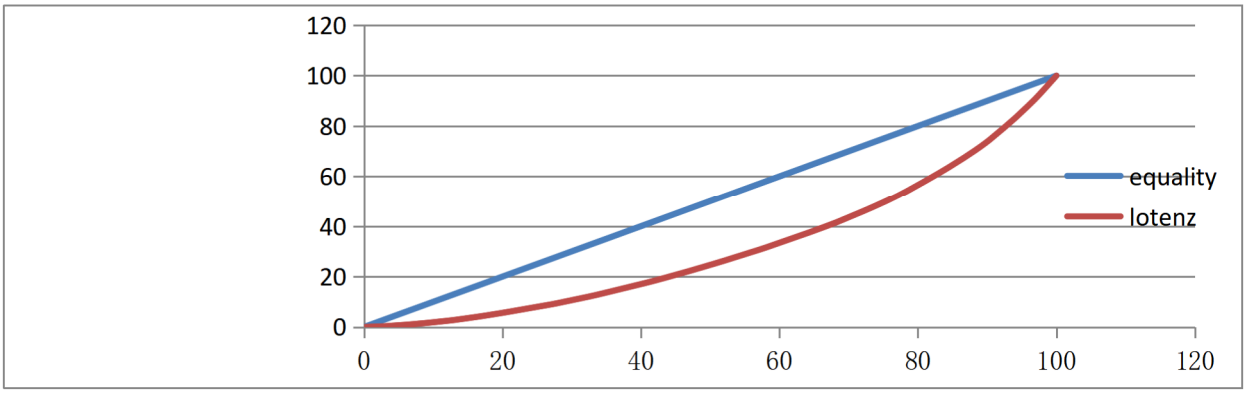


Figure 3: Lorenz curve for income distribution in Kobo.

Gini Index Calculation

Even if Lorenz can measure inequality, it cannot indicate the exact quantitative value of the distribution's dispersion. So Gini coefficient is the best measure of inequality with the exact number to indicate the level of inequality. It continuously measures a value between zero and one (between 0 and 100 when calculated as a percentage). The Gini index is zero when there is equal distribution indicating all individuals under consideration are earning equal income levels. One case is when one individual earns all the income while others earn nothing. World Bank report of countries' results on distributional studies shows that low-income distribution is experienced in a country if its Gini index is between 22 and 30 and is highly unequal if its Gini index is between 31 and 64. Among countries in the world, low-income inequality exists in the Scandinavian countries of Norway and Sweden, with respective Gini indexes of 25.9 and 27.3. In opposition to this high level of income inequality exists Comoros (64.3), followed by the African countries of Botswana and Namibia with a Gini coefficient of 61.3 and 60.4, respectively. Egypt has the lowest level of income inequality among all African countries, with a Gini index of 30.8. In our country, Ethiopia, a relatively low level of income inequality is recorded, with a Gini index of 33.3 at the national level compared to African countries (WB, 2015).

The excel method of calculating income inequality at the household level with the following formula (American Statistical Association, 2014) is applied to derive the value of the Gini coefficient in this study.

$$\mathbf{gini} = \frac{\sum(\mathbf{2i-n-1})\mathbf{xi}}{\boldsymbol{\mu}\mathbf{n^2}}$$

Where;

i = individual household μ = mean value of income and

n = total sample size xi = income of household i

The Gini coefficient of kobo town is estimated to be 0.379 to indicate the high level of income inequality. So this Gini figure of .38 is more significant than Ethiopia's 0.33 national average Gini coefficient as measured by World Bank (WB, 2015). There are two reasons it is greater than the national average Gini index. First, there is a high-income inequality in urban areas of Ethiopia and a relatively low-income inequality in rural counterparts due to annually earned equal agricultural income. So, a high-income inequality level exists in urban areas compared with the national average since the average is low inequality in rural areas. Second, one of the sampling techniques applied in this study is the stratified sampling technique which considers households from the lowest and highest income groups.

Econometric Analysis

This study applies ordered logistic regression with a maximum likelihood estimation technique. Under this logistic regression model, the significant effect of explanatory variables on income is analyzed using the marginal effects of the maximum likelihood estimators. Before estimating models, some diagnostic tests about the relevance and Nature of data are presented below.

Diagnostic Tests

Multi-collinearity problem: Multi-collinearity is a phenomenon that may be observed in multiple linear regressions because of either little variation or high correlation among these explanatory variables. It results in a high estimate of standard error (small t-ratios) high R-square. An economic term used to measure the degree of multi-collinearity in the model is the Variance Inflation Factor (VIF). That is given by $VIF = \left(\frac{\delta^2}{\sum Xi^2(1-R^2)}\right)/\frac{\delta^2}{\sum Xi^2} = \frac{1}{1-R^2}$

From this equation, it is observed that there is a positive relationship between the variance inflation factor and R-square.

The decision rule for the Variance Inflation Factor is that if VIF is less than ten, the model has no multi-collinearity problem. If it is more significant than ten, the problem of multi-collinearity exists.

As presented in the table below, the variance inflation factor for all variables included in the model is less than ten with the VIF mean value of 1.15 or 0.115 level of tolerance, indicating that there are no multi-collinearity problems. For all continuous independent variables, the correlation result is less than 75%, proving that there is no multi-collinearity problem.

Table 5: Variance inflation factor of the continuous variables

Variable	Vif	1/vif
Family size	1.25	0.802242
Experience	1.21	0.824736
Dr	1.13	0.887078
Edu	1.01	0.986698
Mea vif	1.15	0.115

(own source computation using stata)

Table 6: Contingency correlation for discrete independent variables (all less than 75%).

	house	remittance	formal	sector	marriage
var					
house	1.0000				
remittance	0.2504	1.0000			
formal	0.1820	-0.0075	1.0000		
sector	-0.0116	-0.0003	0.0030	1.0000	
marriage	0.1435	-0.0048	0.0901	0.1269	1.0000

(source: own computation using stata)

Heteroskedasticity

It is a condition in which the assumption of constant variance is not fulfilled because of significant variations in the size of entities for which data are collected and outliers, which can increase or decrease variance. With this problem, it is not good to make an economic interpretation for coefficients of independent variables since it has a negative effect.

In this study, by using Breusch-Pagan / Cook-Weisberg test for heteroskedasticity, there is no problem with Hetroskedasticity since the probability is greater than 5% and the chi-square is insignificant (chi2 (1) = 0.02, Prob > chi2 = 0.88) which fails to reject the null hypothesis (Ho: Constant variance).

Normality test

According to the Shapiro-Wilk W test for normality, the error term is said to be normally distributed, or the null hypothesis, which says the error term is normally distributed, is rejected when the value of SW is less than the relevant critical value.

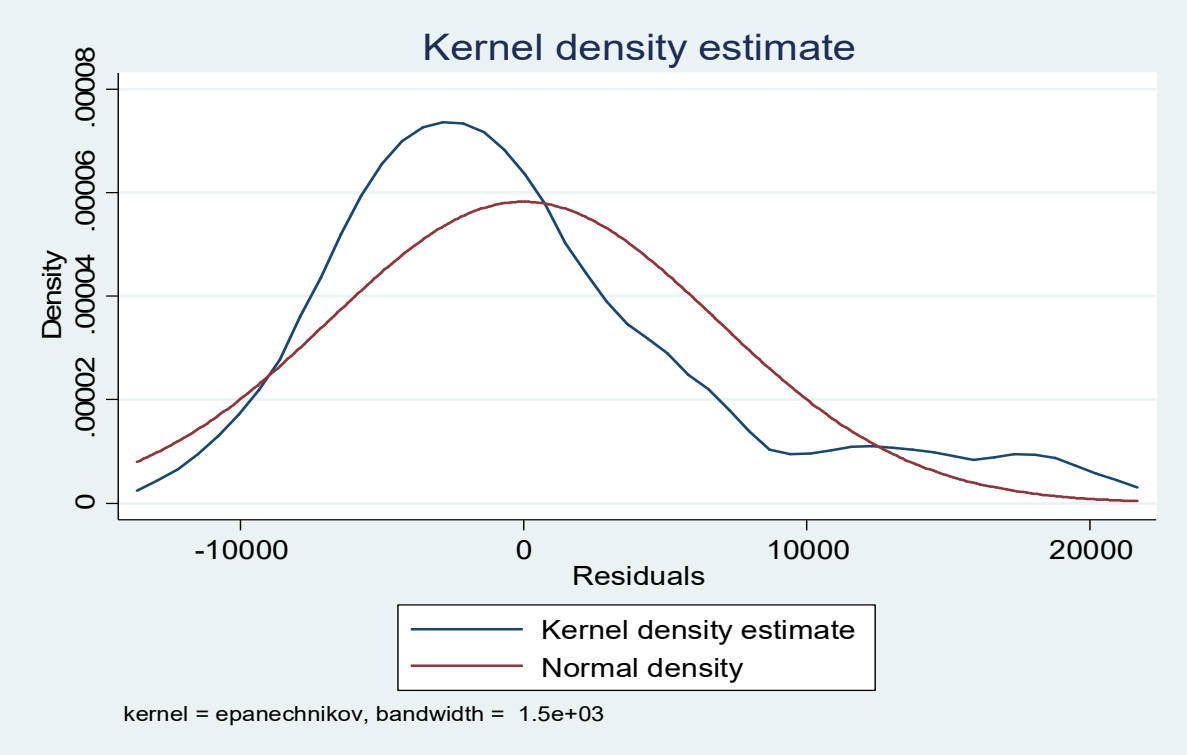
On the other hand, the error term is normally distributed if it has zero mean and constant variance. In this model, the normality test shows that the mean value of the error term (probability) is zero in an approximate value. Previously the model was tested for heteroscedasticity, and its variance was proved constant. However, in this test, the null hypothesis that there is normal distribution is rejected; hence the error term is not normally distributed.

Table 7: Shapiro-Wilk W test for normal data

Variable		Obs.	W	V	Z	Prob>z
Error		381	0.92435	19.940	7.105	0.00000

(source: own computation using stata)

Using a null hypothesis which says H_0 : there is a normal distribution, their distribution is proved non-normally distributed by rejecting the null because of a p-value less than 5%, as shown in the above table. So this abnormality problem may result from the effect of relevant variables excluded from the model, and testing for omitted variables is necessary. However, the low large number and the central limit state that the normality problem does not matter if the sample size under the study is large enough with a value greater than 200 and the sample size is 381, which is greater than 200(Lumley, 2009).



(source: own computation using stata)
Figure 4: kernel densities for normality test

Test for Omitted variables
Ramsey RESET test using powers of the fitted values of y

Ho: model has no omitted variables	
F(3, 368) =	4.77
Prob > F =	0.0028

The probability is greater than 5%, which is 8.2%, and the null hypothesis will be failed to reject, and there is no omitted variable in the model. The distribution is not normal, but there is no problem with omitted variables. All the above tests are equivalent for both the OLS estimation and the ordered logit model, but the following test is valid only for the latter.

Model Specification Test
It tests whether we need new variables in our model to run a new regression with observed Y against Yhat and Yhat-squared as independent variables. The null hypothesis is that there is no model specification error. If the p-value of *hatsq* is insignificant (0.30), we accept the null and conclude that our model is correctly specified.

Test for Proportional Odds Assumption
A commonly used model for analyzing data with categorical, ordinal outcomes is logistic regression, assuming proportional odds. This assumption sometimes called the assumption of parallel lines, is that the effect of independent variables on the ordinal dependent variable is uniform over all of the categories or levels of the dependent variable income groups in this study. That is, a one-unit increase in independent variables has the same effect on the probability of a response being in a higher category regardless of category.

The Stata program brant was used to test the proportional odds assumption. Brant, 1990 (as cited by Fikadu, (2009) proposed a test of proportional odds assumption for the ordinal logistic model by examining the separate fits after estimating the logit model. It also provides tests for each independent variable. When only one independent variable exists in the model, the results of the omnibus and individual tests are the same. The Brant wald test of parallel regression assumption yields a chi-square statics of 0.259, greater than 5% probability. Indicating that the proportional odds assumptions for the whole model (with the null hypothesis H0: there is parallel regression) are held since the null hypothesis of parallel regression exists is failed to be rejected. Accordingly, the effect of the explanatory variable on the dependent variable is the same irrespective of the

categories to show a consistent effect of regressors across the categories or levels of income.

The Estimation Result of the Ordered Logit Model

Table 8:Ordered logistic regression results

Variable	Coef	Std. err	Z	P>|z|	[95% Conf.	Interval]
House	-.4457689	.2228309	-2.00	0.045***	-.8825095	-.0090283
Remittance	-.1536314	.2247322	-0.68	0.494***	-.5940984	.2868357
Experience	.0379221	.0155498	2.44	0.015	.007445	.0683992
Edu	.0787218	.0163618	4.81	0.000*	.0466532	.1107904
Formality	.0523915	.1990302	0.26	0.792	-.3377005	.4424834
Sector	-.2291561	.1943284	-1.18	0.238	-.6100327	.1517205
Marriage	.7795264	.2438368	3.20	0.001	.3016151	1.257438
Dr	.003653	.0018238	2.00	0.045**	.0000785	.0072276
Family size	.0031856	.050764	0.06	0.950	-.09631	.1026813
/cut1	-2.476747	.4450621			-3.349053	-1.604441
/cut2	.360568	.3162939			-.2593568	.9804927
/cut3	2.148743	.3317294			1.498565	2.798921

*, **, *** is the 1%, 5%, and 10% level of significance of variables
(source: own computation using stata)

Number of obs = 381 Wald chi2(9) = 51.66 Prob > chi2 = 0.0000
Log pseudolikelihood = -417.65541 Pseudo R2 = 0.0625

The results table reports three cut-points: cut1, cut2, and cut 3 are ancillary parameters that indicate the thresholds between the income groups. These are the estimated cut points on the latent variable, Y*, used to differentiate the adjacent levels of categories of income levels. When the response category is 1, the latent variable falls at or below the first cut point. When the response category is 2, the latent variable falls between the first and second cut points; when the response category reaches three if the latent variable is between the second and the third cut point. Finally, income will be above the third cut point if the response category is four.

The Chi-Square test with n-1 degrees of freedom is LR chi2 (9) = 51.66, indicating that the predictor is ordered logit regression coefficient is statistically significant and different from 0 in their combined effect. As a result, the entire model with the predictor provided a better fit (reject the null hypothesis H0: all coefficients are zero since p=0.000 is less than 5%). Under this logistic regression model, the significant effect of explanatory variables on income is analyzed using the coefficients, odds ratio, and marginal effects of the maximum likelihood estimators.

Unlike in the OLS, the coefficients of this model will tell us the relationship between income and its explanatory variables, whether positive or negative. However, it cannot tell the magnitude by what amount income will change as explanatory variables change; the magnitude can be distinguished using marginal effect and odds ratio values.

Starting with the negative coefficients, the level of income changes in the opposite direction to the dependency ratio; it increases as the dependency ratio falls and decreases as the dependency ratio increases. This result agrees with the case of households in Tominiain and Koutalia towns of Mali (Brenda, 2013)

Other variables like family size, experience, remittance, ownership of the house, level of education, being hired in the private sector and being male-headed positively correlate with income level. After considering this relationship, the magnitude by what level of probability the income of households will be affected as explanatory variables change can be explained using the odds ratio.

The odds ratio value is the ratio of the probability of success to failure. In this study, the exciting event is to be in the upper-income class. An odds ratio value greater than one indicates that their probability of being in a high-income stream is greater than the probability of falling under this category. When the explanatory variables change by one unit, the log of odds that the individuals became high-income earners will change by the coefficient amount. The effect of variables with magnitude can be explained using the marginal effect's values presented below.

Estimation Of Marginal Effects After the Ordered Logit Model

Table 9: Marginal effect results

Var	dy/dx (y=1)	p>z	dy/dx (y=2)	p>z	dy/dx (y=3)	p>z	dy/dx (y=4)	p>z
House*	.0076675	0.069	.0716776	0.039	.0226079	0.152	-.1019529	0.049
Remittance*	.0028439	0.524	.0257241	0.499	.0057842	0.462	-.0343523	0.490
Experience	-.0006793	0.034	-.0062541	0.014	-.0016389	0.111	.0085722	0.016
Edu	-.0014101	0.008	-.0129827	0.000	-.0034022	0.022	.017795	0.000
Formality*	-.0009432	0.791	-.0086611	0.793	-.0022193	0.789	.0118236	0.792
Dr	-0000654	0.092	-.0006025	0.046	-.0001579	0.119	.0008258	0.045
Family size	-.0000571	0.950	-.0005254	0.950	-.0001377	0.950	.0007201	0.950
Sector*	.0041378	0.238	.0378868	0.248	.0096342	0.256	-.0516588	0.235
Marriage*	-.172492	0.026	-.1386479	0.003	-.0073423	0.585	.1632395	0.001

(source: own computation using stata)

(*) dy/dx is for discrete change of dummy variable from 0 to 1

The above marginal effect results explain the probability of dummy variables first. Being a house owner is associated with0.07167%more likely to be in the low, middle-income group, and -0.1019529% less likely to be in the high-income stream, 71.6%, and 10.19% of probability being in lower middle-income earners and the high-income group households respectively. This probability result reflects that the probability of own households being in a higher income group is less than that of living in rented households. The marginal effects for the dummy of marital status show that married households are associated with -17.2% and -13.8% being less likely to be in low and lower-middle-income groups and 16.3% more likely to be in the high-income groups, respectively. The probability of married households being in a high-income group is greater than that of single household heads.

In addition, the effect of a unit change in continuous explanatory variables on the dependent variable income is also presented in the marginal effects table. Based on these results, a unit change in the level of education is associated with 0.14%, 1.29, and 0.34% less likely to be in the low and lower middle and upper-middle income earner households and 1.77% more likely to be in high-income earner households. The level of education is the significant variable to affect the level of income, and a unit upward movement in the level of education enables the household heads to be in the high-income earning groups. This result is similar to studies made by Ssewanyana, Okido, Angeni, and Burundi (2004), which proves the direct relationship between the level of education and income level,

One unit rise in experience is associated with 0.067%, 0.062% probabilities less likely to be in low and lower middle income, and 0.085% of probability more likely in the high-income groups. This shows that the probability of experienced household heads being under high income is more significant than becoming low and lower middle income. According to the above result, a unit change in dependency ratio is associated with 0.06% of household heads less likely to be in lower middle income and 0.82% more likely to be under the category of high income.

Conclusion

Income inequality is when individuals living in a given country earn a dispersed and unequal income. This accounts for the presence of high poverty, especially for households with lower incomes. On the other hand, reducing the level of income inequality can influence the process of poverty reduction at large since fair income distribution can reduce poverty and boost productive resources. First, identifying the factors determining the existence of inequality is essential to reduce income inequality.

With this objective, this study analyzes the determinants of income inequality in kobo town. For the successful accomplishment of objectives, secondary data from the kobo city Administration and primary data obtained from surveying the town's households are applied. The descriptive methods of analysis from the primarily collected data from 381 sampled households show that out of the total households in the town, 48.3% are male-headed, and the remaining 51.7% are female-headed households. The marital status condition is explained by 75.3% married and 24.6% divorced and widowed household heads. Regarding house ownership, only 28.3% of households live in their own houses, and the rest, 71.6%, live in rented houses. The descriptive statistics of continuous variables also show that there is a five average number of individuals in each household with a mean dependency ratio of 55.8% and an average education level of completing preparatory school.

The inequality situation in this town is analyzed using income distributions. The income distribution is explained by minimum and maximum values of 1000 and 32000, with mean and standard deviation values of 10499.74 and 7314.8, respectively. Out of the total population, the poorest 1%, 5%, 10%, and 15%are earning 1000, 2000, 3000, and 5000 incomes, respectively. On the other hand, the top 1%, 5%, and 10% of households earn an income of 32000, 27000,22000, and 13000, respectively.

The widely used measures of income inequality, like the Lorenz curve and Gini index, are also applied in this study. This distribution is summarized using a quantitative value indicator inequality measure of the Gini index is estimated to be 0.038. In this town, the Gini index is greater than the national average index of 0.33 because of the high level of income inequality in urban areas than in rural areas and the stratified sampling technique applied in the study, which includes households from all income groups.

In addition to the descriptive method of data analysis, the econometric method is also applied with an ordered logit model with a maximum likelihood estimation technique. Then, the ordered logit model is estimated, and the effect of the explanatory variable on income is explained using marginal effect results. Accordingly, the marginal effects regarding house ownership show that those households who live in their own house are associated with 0.76%,0.71%and 2.26% more likely to be in the lower, lower middle, and high middle-income groups, 10.1% less likely to be in the high-income groups respectively. The probability that marital status being married to be in a high-income group is higher than that of the single 1.7%,1.3and7.3%household heads, and 1.6%, respectively

The effects of continuous variable reported as a unit change in education level are associated with 1.41%, 1.29%and 0.34% less likely to be in the low, lower middle, and upper-middle-income earner households and 0.177% more likely to be in the high-income earner households. The level of education affects the level of income positively. An upward movement in the level of education enables the household heads to be in the high-income earning groups. A one-unit change in experience is associated with 0.67%, 6.2%and1.63% less likely to be in low, lower-middle, and upper-middle-income earner households and 8.5% more likely to be in high-income earner households. The number of years of experience positively affects the level of income. One unit's upward movement in the years of experience enables the household heads to be in the high-income earning groups.

Finally, it is proved that one unit rise in household size is associated with 0.057%, 0.52%, and 0.13% probabilities less likely to be in low, lower middle, and high income and 0.72% more likely in the high-income groups**.**

References

Abebe F.N., (2016), determinants of income inequality in urban Ethiopia: a study of south wollo administrative zones, Amhara national and regional state, International journal of applied research vol.2 (1), pp.550–563, Andhra University

Afonso C.,(2008). "determinants of income distribution and efficiency of public spending in Bangladesh," working paper presented to Bangladesh institute of economics, V. 10(2): pp. 52

Alejos K.,(2003). "determinants of income inequality in Guatemala," journal of development economics, V. 5: pp. 22

Alemayehu G., Abebe S., and John W.,(2008). "Growth, poverty, and inequality in Ethiopia: Which way for Pro-Poor Growth?", Journal of International Development 21: 947–970

Alemayehu G., Tadele F.,(2009). "the effect of economic growth on income inequality in Eth Anthony M.,(2009), "income inequality and economic welfare in sub-Saharan Africa', international journal of economics, V.3 pp.5-125

Barro, R.J.,(2000). "Inequality and Growth in a Panel of Countries." Journal of Economic Growth V.5: pp.5–32.

Eskindir Y., (2011). "determinants of income distribution: inequality decomposition analysis in the case of Benchi Maji zone, Sheko district of SNNP of Ethiopia," unpublished MSc thesis presented to Haramaya University graduate studies.

iopia", Andhra University, working paper in Economics

Kenya projects organization (sample size determination using morgan and krejcie table, working paper

Keynes, J. M., (1936), "The General Theory of Employment, Interest, and Money," Harcourt, Brace& World, New-York, 1997 edition by Prometheus Books, Armonk, New York.

Kookshin L., (1997). "determinants of income inequality in Korea," Yale university working paper

Krugman, P., (1979). "Increasing Returns, Monopolistic Competition, and International Trade," Journal of International Economics, Vol. 9, pp. 469–79.

Kuznets, S., (1955). "Economic Growth and Income Inequality," the American Economic Review, Vol.45, No. 1, pp. 1–28.

Tadese wudu – journal of economic and sustainable development Articles

The Determinates of Income Inequality in Urban Ethiopia: The Case of Woldia TownVol 10, No 13 (2019)

Maina K., (2006), "the relationship between economic growth and the applicability of Kuznets curve in Kenya,"

FEC, (2006), Ethiopia: "plan for accelerated and sustained development to end poverty," Addis Ababa, Ethiopia

Okidi U.,(2004), "basic understanding determinants of income inequality in Uganda," M.Sc thesis presented to graduate studies of Makerere university

Olawumi O. and Awe A., 2012, "determinants of income inequality in the Nigerian economy," international

business and management working paper No. 126
Oliver G., (2010), "functional distribution of income inequality and incidence of poverty," working paper, no.58, University of Texas
Perdiz A., (2010), "world growth and inequality," journal of international economics, V.2: pp. 10-57.