İSTANBUL BİLGİ UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
ECONOMICS MASTER'S DEGREE PROGRAM

THE RELATIONSHIP BETWEEN ECONOMIC DEVELOPMENT AND INCOME INEQUALITY: KUZNETS CURVE

Huriye AYHAN

115622006

FACULTY MEMBER, PhD Fatma DİDİN SÖNMEZ

İSTANBUL

2018

The Relationship between Economic Development and Income Inequality: Kuznets Curve

Ekonomik Kalkınma ve Gelir Eşitsizliği Arasındaki İlişki: Kuznets Eğrisi

Huriye Ayhan

115622006

Tez Danışmanı: **Dr. Öğr. Üyesi Fatma Didin Sönmez** ....

Istanbul Bilgi Üniversitesi

Jüri Üyesi: **Doç. Dr. Serda Selin Öztürk** ....

Istanbul Bilgi Üniversitesi

Jüri Üyesi: **Doç. Dr. Yasemin Özerkek** ....

Marmara Üniversitesi

Tezin Onaylandığı Tarih : 07.06.2018

Toplam Sayfa Sayısı: 44

Anahtar Kelimeler (Türkçe)	Anahtar Kelimeler (İngilizce)
1) Kuznet Hipotezi	1) Kuznets Hypothesis
2) Gelir eşitsizliği	2) Income Inequality
3) Ekonomik Kalkınma	3) Economic Development
4) Ekonomik Büyüme	4) Economic Growth
5) Panel Veri	5) Panel Data

ACKNOWLEDGEMENT

Firstly, I would like to express my gratitude to my advisor Asst. Prof. Dr. Fatma Didin Sönmez for continuous support and advisory. Her guidance helped me progress in every step of research.

I would like to thank to Faculty Member, PhD Burak Eroğlu for supporting me in econometrics part. I also special thanks to Assoc. Prof. Dr. Serda Selin Öztürk for her helpful comments and suggestions in econometrics part. Besides, I would like to thank to committee members: Assoc. Prof. Dr. Yasemin Özerkek, and Assoc. Prof. Dr. Selin Serda Öztürk for accepting to join my thesis defense.

Finally, I must express my very profound gratitude to my parents for providing me with support and continuous encouragement throughout writing this thesis. Also, I am very grateful to Mahmut Kıraç for his endless love and as well as never-ending support. This research would not have been possible without their supports.

Thank you.

TABLE OF CONTENTS

List of figures

List of tables

ABSTRACT

In this study, the relationship between economic development and income inequality in the context of the Kuznets inverted-u hypothesis has been examined for 45 countries from 1995 to 2014. According to the Kuznets hypothesis, income inequality increases during the first stages of economic development and then decreases with economic development once it reaches a certain point.

This relationship has been examined by panel data method. While income inequality measured by the Gini coefficient, economic development has been measured both by the Human Development Index and gross domestic product per capita separately. In the literature, economic development usually measured by gross domestic product. In this study, measurement of economic development is both gross domestic product and Human Development Index which is the multidimensional index that includes the education index, the health index, and income index.

There is an u-shaped relationship between economic development and income inequality the model by which economic development is measured both by GDP per capita and by the Human Development Index. Findings in the study have shown that the equation measured by the Human Development Index, better explains the relationship between economic development and income inequality. According to the study results, there was a U-shaped relationship between economic development and income inequality as opposed to the Kuznets inverted-u hypothesis for 45 countries surveyed from 1995 to 2014.

Keywords: Kuznets Hypothesis, Income Inequality, Economic Growth, Economic Development, Panel Data

ÖZET

Bu çalışmada ekonomik kalkınma ve gelir eşitsizliği arasındaki ilişki Kuznet ters-u hipotezi bağlamında 45 ülke için 1995-2014 yılları arasında incelenmiştir. Kuznet hipotezine göre, gelir eşitsizliği ekonomik kalkınmanın ilk aşamalarında artarken belli bir noktaya ulaştıktan sonra ekonomik kalkınma ile birlikte azalmaya başlar.

Bu ilişki panel veri yöntemi ile incelenmiştir. Gelir eşitsizliği gini katsayısı ile ölçülürken, ekonomik kalkınma hem insani gelişme endeksi ile hem de kişi başına düşen gayri safi yurtiçi hasıla ile ölçülmüştür. Literatürde ekonomik kalkınma genellikle gayri safi yurt içi hasıla ile ölçülmüştür. Bu çalışmada ise ekonomik kalkınma hem gayri safi yurt içi hasıla ile hem de çok boyutlu bir endeks olan sağlık ve gelir endekslerine eşit ağırlık veren İnsani Gelişme Endeksi ile ölçülmüştür.

Ekonomik kalkınmanın kişi başına düşen yurtiçi hasıla ile ölçüldüğü modelde ve insani gelişme indeksi ile ölçüldüğü modelde ekonomik kalkınma ve gelir eşitsizliği arasında ‘u’ şeklinde bir ilişki bulunmuştur. Çalışmada elde edilen bulgular, insani gelişme endeksi ile ölçülen denklemin ekonomik kalkınma ve gelir eşitsizliği arasındaki ilişkiyi daha iyi açıkladığını göstermiştir. Elde edilen çalışma sonuçlarına göre incelenen 45 ülkeye için 1995-2014 yılları arasında ekonomik kalkınma ve gelir eşitsizliği arasında Kuznet ters-u hipotezinin aksine U şeklinde bir ilişki gözlemlenmiştir.

Anahtar Kelimeler: Kuznet Hipotezi, Gelir eşitsizliği, Ekonomik Büyüme, Ekonomik Kalkınma, Panel Veri

INTRODUCTION

Income inequality and economic development always have been one of the most important issues among the economists and politicians. Also, income inequality and economic development are important for economic sustainability. A large of papers have investigated the relationship between economic development and income inequality. Kuznets (1955) was the first study which made a significant contribution in theory of economic development and inequality literature. According to this hypothesis, income inequality rises in the early stage of development and reaches a peak then start to decrease with the later stage of development.

This study investigates the relationship between economic development and income inequality in the context of the Kuznets hypothesis for 45 countries from 1995 to 2014. Income inequality is measured by Gini coefficient. Also, economic development is measured by both GDP per capita which is more traditional way and Human Development Index which is the composite index in separate equations. HDI include three different indexes which are the health index, the education index, and income index. A limited number of countries were used in this study because the data used in the study is problematic especially to compile with the Gini coefficient.

This paper is organized as divided into two sections. The first section examines Kuznets hypothesis and gives a brief overview of literature as theoretically and empirically. Also, income inequality and economic growth measures give in section 1. In the second section, the method of this study (panel data analysis) is examined and Kuznets hypothesis is analyzed as empirically. Besides, model, variables, data, econometrics method, and results are shown in section 2. There is a conclusion part of empirical findings at the end of the study.

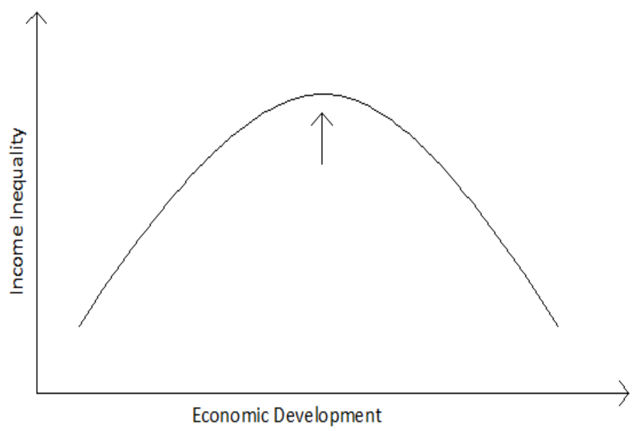
CHAPTER 1

THE RELATIONSHIP BETWEEN ECONOMIC DEVELOPMENT AND INCOME INEQUALITY; KUZNET CURVE

1.1. KUZNETS HYPOTHESIS

Kuznets published a groundbreaking article which title is ‘’Economic growth and income inequality’’ in 1955. Kuznets (1955) claims that income inequality firstly increases in the early stages of economic development and it reaches the peak point then it will decrease in the later stages. This relationship has shown in figure 1.1.

Figure 1.1.Kuznets Curve



Source:Weil D.(2013) Economic Growth:Third Edition, pp.388

As it is shown in figure 1.1, there is the same directional relationship between economic development and income inequality in the early phase of development but income inequality starts to decline in later stages.

According to Kuznets hypothesis, there are two reasons for increase in income inequality in early stage of development. The first reason is only group that can make savings is upper-income groups. For example; the wealthiest 5 percent of the population makes 2/3 of total savings in USA. The poorest twenty percent of society cannot make any savings. Thus upper-income groups earn high revenues from savings and income inequality is increased (Kuznets, 1955, p. 7). Second reason depends on industrial structure. With the economic development, there has been a shift from the agricultural sector to industrial sector. Thus the weight of urban population is increased. Kuznets assumes that firstly the average per capita income of rural population is usually lower that of the urban, secondly inequality in urban population is higher than rural population (Kuznets, 1955, pp. 7,8). With increasing urban population, income inequality is increased in the first phase of economic development. Moreover, productivity in agriculture is less rapidly than the industrial sector. All of these reasons income inequality increase with increasing weight of industrial sector in early stages of development. Subsequently, income inequality decreases with population growth, new entrepreneurship, technological progress, urbanization, political pressures on redistributive policy.

Kuznets has employed USA, England and Germany's data to test the relationship between economic development and income inequality. Countries data include different years.

Kuznets says about his work ‘‘this paper is perhaps 5 percent empirical information and 95 percent speculation, some of it possibly tainted by wishful thinking.’’ (Kuznets, 1955, p. 26). Because Kuznets data set is limited. Thus, there are many studies which have examined Kuznets hypothesis in the literature. Although most empirical workings have accepted the Kuznets hypothesis, some workings have refuted it. Also, some studies have found unidirectional causality from income inequality to economic growth.

1.2. LITERATURE REVIEW

Kuznets (1955) has analyzed limited data sets belong to Germany, England and, United States. Kuznets hypothesis is the first systematic study on economic development literature. In the literature, there are many studies about Kuznets hypothesis. While some studies examine the Kuznets hypothesis theoretically, a number of studies investigate empirically.

1.2.1. Theoretical Studies on Kuznets Hypothesis

Deininger & Squire (1998) explains the relationship between economic growth and income inequality via political channels. Median voter decides income distribution policy according to political channels.

According to Tribble (1999), there are 2 turning points so he explains Kuznets curve as an S-curve instead of an inverted-u curve. While the first turning point implies which shifting from agricultural sectors to industrial sectors at the lower levels of development, the second turning point shows that transition from industrial sectors to services sectors at higher levels of development. The cubic function has utilized to show the S-curve hypothesis. Time-series data for USA from 1947 to 1990 has employed to testing Kuznets hypothesis. As a result, the relationship between income inequality and growth explain as the S-curve hypothesis. Income inequality firstly increase then decrease with economic development, it rises again at advance development. List & Gallet (1999) also supported this S-curve hypothesis using different methodology.

1.2.2. Equations Used in Kuznets Hypothesis

In order to test Kuznets hypothesis, quadratic and cubic functions were used in the literature. These functions are shown below.

$$GINI = \alpha_0 + \alpha_1 GDP + \alpha_2 GDP^2 + \varepsilon \quad (1)$$

$$\mathrm{GINI} = \alpha_0 + \alpha_1 \ln \mathrm{GDP} + \alpha_2 \ln \mathrm{GDP}^2 + \varepsilon \quad (2)$$

$$\ln \mathrm{GINI} = \alpha_0 + \alpha_1 \ln \mathrm{GDP} + \alpha_2 \ln \mathrm{GDP}^2 + \varepsilon \quad (3)$$

$$\mathrm{GINI} = \alpha_0 + \alpha_1 \mathrm{GDP} + \alpha_2 \mathrm{GDP}^2 + \alpha_3 \mathrm{GDP}^3 + \varepsilon \quad (4)$$

$$GINI = \alpha_0 + \alpha_1 \; ln\, GDP + \alpha_2 \, ln\, GDP^2 + \alpha_3 \, ln\, GDP^3 + \varepsilon \qquad (5)$$

Although variations of the quadratic function are used in most studies, cubic functions are also used in some studies. When a cubic function is used, Kuznets curve observed like a horizontal-s instead of an inverted-u curve. In order to support the Kuznets inverted-u curve, the coefficient of $ln\, GDP^2$ and $ln\, GDP$ must be $\alpha_2 < 0$ $\alpha_1 > 0$, respectively. The second derivative shows concavity and convexity. The second derivative must be lower than zero because Kuznets curve is concave. The first derivative gives us critical point of the function. Besides, turning points of the curve is calculated by $-{}^{\alpha_1}/_{2\alpha_2}$ formula.

1.2.3. Economic Development Measures

Economic development in regressions is measured by Gross Domestic Product per capita and also by Human Development Index in literature. Gross Domestic Product and Human Development Index are explained in this section.

Gross Domestic Product: It is measured based on different approach, for instance, production, expenditure, and income. According to production approach, a monetary value of produced final good and services within the country in the specific time.

Human Development Index: It is used to measure development. HDI is ‘‘a composite index measuring human development-a long and healthy life, knowledge and a decent standard of living (Human Development Report 2016: Human Development for Everyone). Components of HDI shows below table 1.1.

Table 1.1.Human Development Index Components

The Health Index	The Education Index	The Income Index
Life expectancy at birth	Mean years of schooling	Gross national income per capita
	Expected years of schooling	

Human Development Index is equal to geometric mean of the three indices. The formula of indices as shown below.

$$\boldsymbol{Dimension\ index} = \frac{\boldsymbol{actual\ value - minimum\ value}}{\boldsymbol{maximum\ value - minimum\ value}}$$

The minimum and maximum values defined in (Human Development Report 2016: Human Development for Everyone) .These values shown in below table 1.2

Table 1.2.HDI Components Values

Dimension	Indicator	Minimum	Maximum
Health	Life expectancy(years)	20	80
Education	Expected years of schooling(years)	0	18
	Mean years of schooling	0	15
Standard of living	Gross national income per capita	100	75000

Source: Human Development Report 2016: Human Development for Everyone

Education index is calculated by arithmetic means of two indices which are years of schooling index and expected years of schooling index. Also, logarithms of the values are used for the income index.

1.2.4. The Difference between Gross Domestic Product and Human Development Index

When it looked at the literature, there are many studied about Kuznets hypothesis and many of these studies measure economic development by gross domestic product. In this study, economic development is measured by Gross Domestic Product per capita and also economic development is measured by Human Development Index.

Economic growth or gross domestic product is not only the indicator of economic development because economic development has a multidimensional process. It should include components such as welfare, education, health, employment, political voice and governance, social relationship and environment (Stiglitz, et al.). These components related to people's quality of life. One of the most common measures of economic development is Human Development Index (HDI). Human development index includes the education index, the health index, and income index. Therefore, HDI covers both GDP and other two indexes which related to human well-being. Human Development Index developed by Amartya Sen and Mahbub ul Haq in 1990.

Sen (1992) focus on two concepts which are functioning and capability. According to Sen's capability approach, individual well-being depends on functioning and capabilities. Functioning is that human life is what people value of 'doings' or 'beings'. Capability is the freedom of people to prefer between different functioning (Sen, 1997, p. 388). Some examples of functioning are: being healthy and well-nourished, being literate, being a participation of community, being respected, being happy and so on (Sen, 1992). Functioning depends on some variations such as illness, disability, age, climate, pollution, public health care, education and so on (Sen, 1997, p. 385) Therefore, Sen (1992) have emphasized the importance of individual well-being and freedom.

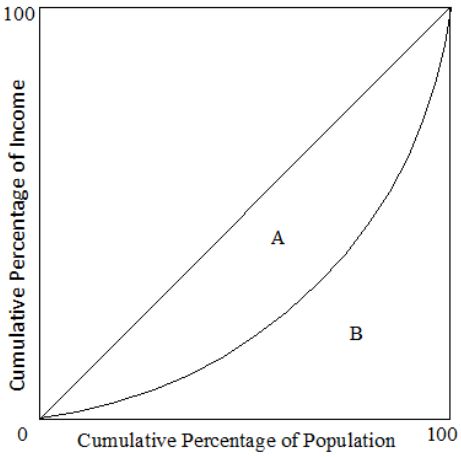
1.2.5. Income Inequality Measures

Income inequality is measured by many different methods such as the Gini coefficient, Lorenz curve, the Theil index, Atkinson index, Decile ratios, Proportion of total income earned, and Robin Hood index. Gini coefficient, Lorenz curve and Decile Dispersion ratio are explained in this section.

Decile Dispersion Ratio: Society is divided into groups according to percentiles then look at the groups' share of total income. For example, the 10^{th} percentile represents the poorest 10 percent of the population and the 90^{th} percentile shows the richest 10 percent. Decile dispersion ratio is the ratio of average income of the richest 10 percent divided by the poorest 10 percent. ‘‘The decile dispersion is readily interpretable, by expressing the income of the top percent (the rich) as a multiple of that of that poorest decile (the poor)’’ (Jonathan & Shahidur, 2009, p. 4).

The Gini coefficient: It which based on Lorenz curve is one of the most common methods to measure income inequality. Lorenz curve shows a distribution of income, graphically. The Lorenz curve is shown in figure 1.2.

Figure 1.2. Lorenz Curve



Source: Weil D. (2013) Economic Growth:Third Edition, pp.387

The x-axis represents cumulative percentage of the population (from poorest to richest) and y-axis demonstrates the cumulative percentage of income. Diagonal line demonstrates absolute equality. Percentage of population is equal to percentage of income corresponding to this population at 45-degree line. The Gini coefficient is equal to (area A/area A+area B). As area A decreases, income inequality decreases. The Gini coefficient takes the value between 0 and 1. While 0 shows perfect equality point, 1 represents perfect inequality which means income received by one person.

1.2.6. Empirical Studies on Kuznets Hypothesis

In the literature while several studies such as Ahluwalia (1976), Ram (1989), Zang (1998), Chen (2003), Huang (2004), Dawson (1997), Thornton (2001) and Oskooee & A.Gelan (2008) have supported the Kuznets hypothesis, some studies such as Ram (1997) have refuted it. Moreover, S-curve relationship has found between economic development and income inequality in some studies for instance List & Gallet (1999), Tribble (1999) and Theyson & Heller (2015). These different findings are a result of data quality and statistical methodology. There are many criticisms on the Kuznets hypothesis. Since data set is not large enough, cross-sectional analysis has to be done. These studies have shown on table 1.3. Different types of equations and analysis have used to test Kuznets hypothesis.

Table 1.3.Empirical Studies on Kuznets Hypothesis

Study	**Date/Country**	**Regression/ Variables**	**Explanatory Variables**	**Method**	**Result**
Ahluwaia (1976)	62 countries (14 developed countries, 42 developing countries and 6 socialist countries)	Quadratic/ percentage income shares, GNP per capita	Education, growth rate of population, share of agriculture in GDP, share of urban population , growth rate of GDP	Cross-section	Inverted-u curve
Ram (1989)	1960-1980/115 countries	Quadratic/ Theil's population index , GDP per capita		Time-series	Inverted-u curve

Study	Date/Country	Regression/ Variables	Explanatory Variables	Method	Result
Chen(2003)	43 and 53 countries	Barro type/ Gini index,GDP per capita	Pyhsical capital ,human capital ,instituonal policy’s variables ,share of government consumption in GDP and inflation rate	Cross-section	Inverted-u curve
Huang(2004)	75 countries	Quadratic/ Gini index,GDP per capita	State sector employee,social transfer expenses	Cross-section	Inverted-u curve
Dawson(1997)	36 less developed countries	Quadratic/ Gini index,GDP per capita		Cross-section	Inverted-u curve
Matyas(1998)	1970-1992/ 47 and 62 countries	Different type of models/ Gini index,GDP per capita		Panel data	Not accepted for all model

Study	Date/Country	Regression/ Variables	Explanatory Variables	Method	Result
Kuştepeli (2006)	1951-1998 European Union countries	2 diffirent type of models/ Gini index,GDP per capita	Nominal exchange rate,private investment growth,government spendng growth	Panel data	Could not be confirmed
Thornton (2001)	96 countries	Quadratic/ Gini index or income share of bottom 40 percent and real GDP per capita		Panel data	Inverted-u curve
Bahmani-Oskooee (2008)	1957-2002 USA	Linear/ Gini index,GDP per capita	population, exchange value of dollars	Time series	Inverted-u curve
List and Gallet (1999)	1961-1992 71 countries	Cubic	Gini index, GDP per capita	Panel data	S curve
Barro(2000)	1960-1990/100 countries	Quadratic/ Gini index,GDP	inflation, democracy index years of schooling	Panel data	Inverted-u curve

Study	Date/Country	Regression / Variables	Explanatory Variables	Method	Result
Zang(1998)	1967-1986 60 countries	Quadratic Gini,income share of poorest 40% of households ,GDP		Cross-section	Inverted-u curve
Gelan,Price(2003)	1985/73 countries(21 Industrial countries,18 sub-african countries)	Standard neoclassical growth model/ Gini,İncome	Open,tropic,countries geographic locations,ethnicity, a former French colony,a former UK colony	Cross-section	Inverted-u curve
Tribble(1999)	1947-1990 USA	Cubic/ Gini, GNP per capita,		Time-series	S curve
Theyson,Heller(2015)	1992-2007 147 countries	Cubic/ Gini index,HDI		Panel data	Inverted-u curve
Ram (1997)	Postwar years/19 developed countries	Quadratic/ Gini index, ,GDP per capita	Country-specific dummy variables	Fixed-effect panel data	U curve

Ahluwalia (1976) employed a sample of 62 countries which are different levels of income. This sample includes 14 developed countries, 42 developing countries, and 6 socialist countries. Quadratic regression is used to explain the relationship between economic growth and income inequality. Beside regression include some explanatory variables which are education, literacy rate, the share of agriculture in GDP, the share of urban population and population rate. The finding of this study has supported the Kuznets hypothesis.

Ram (1989) use a large sample of 115 countries for a period from 1960 to 1980 and accepted the Kuznets hypothesis. Theil's population-weighted index is used to measure income inequality. Also, Kuznets-type quadratic form is used as regression. Thereafter, Ram (1997) analyzed observations for 19 developed countries which cover various postwar years. Income distribution data is from Deininger-Squire dataset which is a high-quality compilation. Kuznets-type quadratic regression is used also country-specific dummy variables are added. There are many variables to affect income inequality such as public policies, historical and institutional factor. Dummy variable provides to examine all these effects on inequality. Gini index and also the income-share of each quintile have been used for a measure of inequality. The uninverted-U curve was found.

Barro (2000) examined 100 countries from 1960 to 1990 with using the quadratic functional form to explain Kuznets hypothesis. Some explanatory variables which are the ratio of government consumption to GDP, democracy index, inflation, years of schooling, fertility rate and growth rate in terms of trade are added to function. He found inequality has a negative impact on growth in poor countries but positive impact in rich countries. However, the weak relationship between inequality and growth across the entire sample.

Deininger & Squire (1996) compiled a high-quality data set which related to inequality of income must satisfy three conditions. The first data set of inequality based on household survey. National accounts data are not used. Second data cover whole population rather than only wage earner and urban or rural population. Third data cover all source of income such as monetary and non-monetary income.

Chen (2003) has used 43 and 54 countries in his study. Initial GDP, physical capital, human capital, intuitional policy’s variables, a share of government consumption in GDP and inflation rate are included in the regression as control variables. Besides, regional dummies (Latin, Africa, Asia) are added as control variables. He found interved-u relationship between economic development and income inequality.

Huang (2004) has developed a new approach, unlike traditional approaches. Quadratic regression is commonly used for examining Kuznets hypothesis. In contrast, flexible nonlinear inference approach is used to explain Kuznets hypothesis. The sample of 75 countries data is used and Kuznets hypothesis is confirmed.

Dawson (1997) employed a sample of 36 less developed countries and used quadratic regression model. Thus Kuznets inverted u curve was found statistically significant.

Matyas, et al. (1998) have used two data sets which are 47 and 62 countries from 1970 to 1992. Different type of models are used and results show that Kuznets hypothesis not accepted for all model. Also, income inequality depends on country-specific factors such as social structure, institutions, political system, more than level of development.

Kuştepeli (2006) use different groups of countries such as the European Union before the latest enlargement, the European Union after latest enlargement and the European Union after participation the three candidates to analyze Kuznets hpothesis.2 different type of models are used and could not any significant results about Kuznets hypothesis.

Thornton (2001) has employed 96 countries for testing Kuznets hypothesis. Gini index or income share of bottom 40 percent and real GDP per capita is used to explain relationship the between income inequality and growth. Gini index obtains from the high-quality compilation by Deininger & Squire (1996). Results support Kuznets hypothesis.

Zang (1998) use sample of 60 countries for a period from 1967 to 1986 and accepted the Kuznets hypothesis. Quadratic regression is used also time add as an explanatory variable. Gini coefficient, the income share of poorest 40% of households, and GDP per capita according to PPP are used.

Oskooee & A.Gelan (2008) examine the relationship between economic growth and income inequality and also test impact of currency depreciation on income distribution in the short-run and in the long-run. Data obtained from time series which are a period from 1957 to 2002. Error-correction model are used. The results support Kuznets hypothesis. Depreciation of dollars affects adversely income inequality in the short-run but lose its effect over time.

(List & Gallet, 1999) have examined Kuznets curve for using 71 countries which include lower-developed and higher-developed countries from 1961 to 1992. Cubic regression is used also random-effects panel data model is more fitted according to Hausman statistics. There are 2 thresholds which are 1487 Dollars and 12115 Dollars. Kuznets curve is valid from lower-developed countries to middle-developed countries but there is a positive relationship between income inequality and growth in the advanced economies.

Price (2003) have employed standard neoclassical growth model that put into the Kuznets equation. Cobb-Douglas productions function and also labour which consisted of skilled and unskilled labour are used in the equation. Industrial countries and sub-Saharan African countries have been tested separately and in all countries. Physical capital variable is positive in the sub-Saharan African countries in Kuznets equation with illiteracy rate and without illiteracy rate, without technology augmented. This implies dualistic economy. Kuznets curve does not exist in the Industrial countries but valid in the sub-African countries.

Theyson & Heller (2015) use a large sample of 147country for a period from 1992 to 2007.HDI which includes GDP per capita population health and education is used for a measure of economic development. List & Gallet (1999) cubic functional form is used but HDI has used instead of GDP per capita. When log GDP is used as a measurement of development, traditional Kuznets curve is founded. When HDI is used instead of log GDP, income inequality first rise then decrease significantly. The relationship between HDI components such as education index, life expectancy index and income inequality like this S-curve form. HDI components are a reason of this significant decreasing. Also this S-curve form different from previous literature, inequality firstly decrease at the beginning of the development and increasing then falling again at advanced level of development.

CHAPER 2

THE RELATIONSHIP BETWEEN ECONOMIC DEVELOPMENT AND INCOME INEQUALITY: EMPIRICAL ANALYSIS

2.1. METHOD of THIS STUDY: PANEL DATA ANALYSIS

As it is mentioned before, this study investigates the validity of the Kuznets hypothesis between the years 1945 and 2015 in 45 countries was examined using panel data analysis.

2.1.1. Definitions

There are 3 econometric data types which are time series, cross section data and panel data according to data structure. Time series data, cross sectional data and panel data are explained in this section.

Time series: Values of a variable at periodic time such as yearly, quarterly, weekly (Gujarati, 2004, p. 25). For example, USA unemployment rate between 1950 and 2000.

Cross-section: Values of variables at a certain time (Gujarati, 2004, p. 26). For example, unemployment rates of different countries in 2000.

Panel data (Longitudinal data): It is consist of time series and cross section data (Gujarati, 2004, p. 636). Panel data set contain a number of individuals, households or firms at different point of time (Greene, 2002, p. 283). For example, unemployment rates of OECD countries from 1950 to 2000.

National Longitudinal Survey of Labor Market Experience (NLS) and University of Michigan's Panel Study of Income Dynamics (PSID) are the most important data set in US. (Baltagi, 2005, p. 1). While panel data collection in USA began in 1960s, the European panels started in the 1980s. The Statistical Office of the European Communities (EuroStat) centrally collected and coordinated the European Community Household Panel (Baltagi, 2005, p. 3).

According to observation, panel data is splitted into two types which are balanced panel and unbalanced panel. Balanced panel is that every variable have observed for all sample time. If every variable have not observed for all sample time .It means that we have missing observations. This name is unbalanced panels (Baltagi, 2005, p. 165).

According to Baltagi (2005), panel data regression is shown below form (Baltagi, 2005, p. 11).

$$\boldsymbol{y_{it} = \alpha + X'_{it}\beta + u_{it} \qquad i = 1, \dots, N; t = 1, \dots, T}$$

where i shows cross-section dimension such as households, individuals, firms. Also t shows time-series dimension. $\boldsymbol{\beta}$ is Kx1 and $\boldsymbol{X_{it}}$ is the itth observation on K explanatory variables. $\boldsymbol{u_{it}}$ is the error term. Error term assumed to be with zero mean and constant variance. Also error term is uncorrelated with other variables.

2.1.2. Panel Data Benefits and Disadvantages

- Panel data enable researchers to do use both cross section data and time series data at the same time. Besides, individual heterogeneity can be controlled in panel data although cannot be controlled in time-series and cross-section analysis. Individuals and countries are heterogeneous because of some factors a like cultural factors or government regulation. State-invariant variables and time-invariant variables can be controlled in panel data (Baltagi, 2005, p. 4). Individuals or countries behavior can be modelled with panel data because individual heterogeneity may be included in model. (Greene, 2002, p. 284)
- Panel data include more observation thus collinearity decreases and degrees of freedom increases. (Baltagi, 2005, p. 5)
- Panel data is suitable to study for variables that change over time such as job turnover, income mobility. (Baltagi, 2005, p. 6)
- Time-invariant variables may be included in panel data thus panel data are better explain than cross section and time series data. (Baltagi, 2005, p. 6)
- Problems may arise in collection and organization of panel data. (Baltagi, 2005, p. 7)

2.1.3. Panel Data Models

When estimating the panel data model, three approaches which are the common constant model, fixed-effects model and random-effects model are used.

2.1.3.1. The Common Constant Model

It is also known the pooled OLS method. Constant term is same for all countries. Data set consists of countries with the same characteristics such as only middle-income countries (Asteriou & Hall, 2007, p. 345).

2.1.3.2. The Fixed Effects Model

If data set are selected from specific set of N firms such as N European countries, N high-developed countries, fixed effects model is more appropriate.

Fixed effects model shows below form (Greene, 2002, p. 285).

$$y_{it} = X'_{it}\beta + \alpha_i + \varepsilon_{it}$$

Where α_i shows individual specific constant term.Individual heterogeneity is included model with constant term (Asteriou & Hall, 2007, p. 348).Dummy variable is defined for each country thus it allows us to observe country specific factors. Fixed effects model is called the least-squares dummy variables (LSDV) (Asteriou & Hall, 2007, p. 346). Dummy variables can be defined for both country and time in fixed effect model. This model called the two way fixed effect models (Asteriou & Hall, 2007, p. 347).

2.1.3.3. The Random Effects Model

If N individuals are randomly selected from within a group, random effects model is more appropriate (Baltagi, 2005, p. 14).Random effects model also known as the error components model.

Random effects model shows below form (Greene, 2002, p. 285).

$$y_{it} = X'_{it}\beta + \alpha + u_i + \varepsilon_{it}$$

where α which is constant term.u_i is a group specific random element. It depends on just group, not time. Individual heterogeneity is included model as a component of error term (Greene, 2002, p. 294).

2.1.3.4. The Preference between Fixed and Random Effects Models

Panel data models which are fixed and random effects models are examined in the previous section. It is important which model will be chosen? There are advantages and disadvantages to use fixed and random fixed effect models (Greene, 2002, p. 301).

Hausman's specification test is used to decide between fixed and random effects models.

2.1.3.5. Hausman's Specification Test

Hausman test has used to decide which test is more appropriate. Hausman test depends on comparing $\boldsymbol{\beta}_{GLS}$ and $\boldsymbol{\beta}_{Within}$ (Baltagi, 2005, p. 66).

E($\boldsymbol{u}_{it}$/ $\boldsymbol{X}_{it}$)=0 , this is more important assumption in error component model (Baltagi, 2005, p. 66).

Null hypothesis is that there is no relationship between error term and explanatory variables.If null hypothesis is rejected, alternative hypothesis is accepted thus fixed effect model is used. If null hypothesis can not be rejected, random effect model is used. Null hypothesis distributed according to chi-square distribution with k degrees of freedom (Baltagi, 2005, p. 67).

2.1.4. Panel Data Assumptions

After we decide to model which are fixed or random, basic assumptions which are heteroskedasticity, autocorrelation and cross section dependence must be controlled for model. The assumptions of heteroscedasticity, autocorrelation and cross section dependence are examined with some tests. These tests and basic assumptions are explained in section 2.1.1.4.1, 2.1.1.4.2 and 2.1.1.4.3.

2.1.4.1. Testing for Heteroskedasticity

Homoskedasticity is assumption of the classical linear regression model. Variance of error terms is constant for every observation in homoskedasticity. Heteroskedasticity means that variance of error term is different for every observation (Asteriou & Hall, 2007, p. 101). White's test, Modified Wald test and the Breusch-Pagan test are used to test heteroskedasticity (Greene, 2002, p. 324).Null hypothesis is that variances are constant in Wald test. When null hypothesis is rejected, heteroskedasticity problem arise (Greene, 2002, p. 323).Breusch and Pagan (1979) find a Langrange Multiplier test. Also, LM distributed according to chi-squared distribution. Null hypothesis is that variances are same in LM test (Greene, 2002, p. 224).

2.1.4.2. Testing for Autocorrelation

Autocorrelation is that correlation between error terms is zero. It means that there is no relationship between different error terms (Asteriou & Hall, 2007, p. 134).The Durbin-Watson test, the Breusch-Godfrey LM test are used to test autocorrelation. Null hypothesis is that there is no autocorrelation thus alternative hypothesis shows serial correlation (Asteriou & Hall, 2007, pp. 140-144).

2.1.4.3. Testing for Cross Sectional Dependence

Common shocks such as oil price shock, financial crises effect individuals at different levels in macro panel. This different affect shows cross sectional dependence. Breusch-Pagan(1980) Lagrange Multiplier test and Pesaran’s (2004) CD test are widely used to test cross section dependence. There is no cross-section dependence under null hypothesis. When accept the alternative hypothesis, cross section dependence is observed (Sarafidis, et al., 2008). Breusch and Pagan test is used for T>N. When N>T, Pesaran test is used. N shows country and T shows time.

2.2. EMPIRICAL ANALYSIS

As it is mentioned before, this study investigates the relationship between income inequality and economic development in 45 countries from 1945 to 2015.Income inequality is measured by Gini coefficient. Also economic development is measured by both GDP per capita which is more traditional way and Human Development Index in separate equations. A limited number of countries were used in this study because the data used in study is problematic especially to compile with the Gini coefficient.

In order to examine Kuznets hypothesis empirically, following quadratic form is used:

$$\boldsymbol{GINI_{it} = \alpha_0 + \alpha_1 \ln GDP_{it} + \alpha_2 \ln GDP^2_{it} + \varepsilon_{it}} \quad (1)$$

where $\boldsymbol{GINI_{it}}$ indicates Gini coefficient for country i in year t, lnGDP denotes logarithm of per capita GDP, $\boldsymbol{\ln GDP^2}$ refers logarithm of GDP per capita squared and $\boldsymbol{\varepsilon}$ is error term. Evidence for Kuznets curve, coefficients of $\boldsymbol{\ln GDP}$ and $\boldsymbol{\ln GDP^2}$ must be $\boldsymbol{\alpha_1}$>0 and $\boldsymbol{\alpha_2}$<0 respectively.

Firstly, economic development is measured by GDP per capita in model (1). Secondly, HDI is used instead of GDP per capita for measure of economic development in this study. This has shown below in model (2).

$$\boldsymbol{GINI_{it} = \alpha_0 + \alpha_1 HDI_{it} + \alpha_2 \mathrm{HDI}^2_{\mathrm{it}} + \varepsilon_{it}} \quad (2)$$

where $\boldsymbol{GINI_{it}}$ indicates Gini coefficient for country i in year t, $\boldsymbol{HDI_{it}}$ refers human development index for i country in year t and $\boldsymbol{\mathrm{HDI}^2_{\mathrm{it}}}$ indicates human development index squared for i country in year t.

Model (1) and model (2) are employed to analyze Kuznets hypothesis, empirically.

2.2.1. Data

To examine the Kuznets hypothesis, a sample of 45 countries are employed which are different levels of income for a period from 1995 to 2014.These sample includes 30 high income countries, 11 upper-middle income countries and 4 lower-middle income countries. These countries are shown according to income level in table 2.1.

Table 2.1.Countries Used in This Study

High Income Countries	Upper-middle Income Countries	Lower-middle Income Countries
Australia	Belarus	Armenia
Austria	Brazil	Bolivia
Belgium	Bulgaria	El Salvador
Canada	China	Georgia
Czech Republic	Colombia	
Denmark	Costa Rica	
Finland	Dominican Republic	
France	Ecuador	
Germany	Romania	
Greece	Panama	
Hungary	Paraguay	
Iceland		
Ireland		
Israel		
Italy		
Japan		
Luxembourg		
Malta		
Netherlands		
New Zealand		
Norway		
Poland		
Portugal		
Slovenia		
Spain		
Sweden		
Switzerland		
United Kingdom		
United States		

Uruguay		

The data for Gini coefficients are taken from United Nations University for World Income Inequality Database (WIID). WIID compiles data from various databases. There are many Gini coefficients for every year. In this study, data is taken according to the method used by (Theyson & Heller, 2015). HDI is taken from United Nations Development Program. GDP per capita ppp 2011 is obtained from World Bank Database. In this study unbalanced panel data set is employed because the data is missing for some years.

2.2.2 Empirical Results

Panel data analyze is applied. Results which belong to regression 2 and regression 1 are shown below, respectively.

Firstly fixed effect model and random effect model were estimated then Hausman test is used to decide which of fixed effect model or random effect model is more appropriate. Country and time effects were investigated on obtained model. Also, panel data basic assumptions were controlled. After testing basic assumptions, model was estimated again robust estimator.

Firstly, in order to examine Kuznets hypothesis empirically, model (2) is used:

$$\boldsymbol{GINI_{it} = \alpha_0 + \alpha_1 HDI_{it} + \alpha_2 HDI_{it}^2 + \varepsilon_{it}}$$

Fixed effect and random effect models results which belong to model (2) are shown in table 2.2, 2.3.

Table2.2.Fixed Effect Model

Fixed-effects (within) regression Group variable: id			Number of obs=808 Number of groups=45
R-sqr: Within=0.082 Between=0.575 Overall=0.465			Obs per group: Min=7 Avr=18 Max=20
Gini	Coef.	Std. Err.	P> | t|
Hdi	-171.506	25.450	0.000
Hdisquare	101.707	16.518	0.000
_cons	106.856	9.786	0.000
F test			F(44,761)=133.45
			Prob >F=0.0000

HDI, HDI square, and constant term are significant at 1% level.HDI and HDI square coefficients shows that there is uninverted–u curve between economic development and income inequality. It shows that income inequality decreases to a certain point and then increases with economic development. When HDI increase 1 unit the Gini decreases by 171.506 units. Coefficient of HDI square is 101.707, if HDI square increases 1%, Gini increases by 101.707 units. R-squared is 0.082 thus 8% of changes in Gini are explained by independent variables which are the HDI and HDI square. Probability of F-statistic is zero therefore all the variables jointly in the model significantly affect the Gini at 1% significance level.

Table 2.3.Random Effect Model

Random-effects GLS regression Group variable: id			Number of obs=808 Number of groups=45
R-sqr: Within=0.080 Between=0.619 Overall=0.521			Obs per group: Min=7 Avr=18 Max=20
Gini	Coef.	Std. Err.	P> | t|
Hdi	-172.189	26.041	0.000
Hdisquare	98.612	16.890	0.000
_cons	109.375	10.051	0.000

HDI, HDI square and constant term are significant at 1% level. Uninverted-u curve is observed as like fixed effect model. When HDI increase 1 unit the Gini decreases by 172.189 units. Also coefficient of HDI square is 98.612. There is positive relationship between HDI square and Gini.

Hausman test is used to decide which of fixed effect model or random effect model is more appropriate. Hausman test result is shown in table 2.4.

Table2.4.Hausman Test

	Coefficients	
	Fixed	Random
Hdi	-171.506	-172.189
Hdisquare	101.707	98.612
Chi2(2)=191.25, Prob>chi2=0.0000		

According to Hausman test, null hypothesis is rejected so random model is inconsistent and fixed effect is better for model (2).

After determining the fixed effect model, period effect and country effect in fixed effect model have been controlled for model (2). Results are shown in below tables which are 2.5 and 2.6.

Table2.5.Time-fixed Effects

Fixed-effects (within) regression Group variable: id			Number of obs=808 Number of groups=45
R-sqr: Within=0.135 Between=0.412 Overall=0.285			Obs per group: Min=7 Avr=18 Max=20
Gini	Coef.	Std. Err.	P> | t|
Hdi	-199,988	25.727	0.000
Hdisquare	140.92	18.458	0.000
T			
1996	-0.263	0.523	0.615
1997	-0.633	0.531	0.234
1998	-0.067	0.536	0.900
1999	-0.193	0.562	0.730
2000	0.822	0.560	0.883
2001	-0.319	0.582	0.583
2002	0.426	0.619	0.491
2003	-0.760	0.631	0.229
2004	-1.224	0.644	0.058
2005	-1.150	0.681	0.092
2006	-1.148	0.711	0.107
2007	-1.718	0.739	0.020
2008	-1.824	0.758	0.016
2009	-1.750	0.771	0.024
2010	-2.347	0.813	0.004
2011	-2.484	0.831	0.003

T	Coef.	Std. Err.	P> | t|
2012	-3.017	0.859	0.000
2013	-3.022	0.897	0.001
2014	-3.131	0.920	0.001
Cons	105.376	9.699	0.000
F test			F(19,742)=2.38 Prob >F=0.0008

Dummy variables are defined to observe the effect of each year separately. Null hypothesis which is the coefficients for all years are jointly equal to zero is rejected; therefore time fixed effects are needed. When looking at the table 2.5, the values after 2007 are significant therefore time effects are observed after 2007.

Country effects have been omitted because of collinearity.

After look at the time and country effects separately, both time and country effects have been considered together. These results are shown in table 2.6.

Table 2.6.Fixed Effects Model with time and country effects

Fixed-effects (within) regression Group variable: id			Number of obs=808 Number of groups=45
R-squared=0.9528 Adj R-squared=0.9487			
Gini	Coef.	Std. Err.	P> | t|
Hdi	-199.988	25.727	0.000
Hdisquare	140.92	18.458	0.000
Cons	92.639	9.781	0.000
Id			
2	3.164	0.756	0.000
3	-0.214	0.732	0.769
4	-1.110	0.848	0.191

Id	Coef.	Std. Err.	P> | t|
5	3.035	0.732	0.000
6	1.384	0.771	0.073
7	2.871	0.813	0.000
8	4.814	0.774	0.000
9	1.258	0.735	0.087
10	3.559	0.771	0.000
11	8.105	1.038	0.000
12	0.914	0.876	0.297
13	6.636	0.749	0.000
14	4.062	0.846	0.000
15	6.377	0.764	0.000
16	8.505	0.843	0.000
17	7.813	0.735	0.000
18	3.565	0.852	0.000
19	14.446	1.154	0.000
20	0.834	0.837	0.319
21	6.042	1.232	0.000
22	19.381	0.741	0.000
23	2.994	0.884	0.001
24	11.345	0.937	0.000
25	20.350	1.762	0.000
26	10.024	0.865	0.000
27	6.125	1.170	0.000
28	12.851	0.907	0.000
29	7.661	0.792	0.000
30	11.195	1.113	0.000
31	23.933	1.418	0.000
32	35.415	1.800	0.000
33	23.869	2.173	0.000
34	10.770	1.599	0.000
35	12.298	1.599	0.000

Id	Coef.	Std. Err.	P> | t|
36	34.968	1.960	0.000
37	27.406	1.613	0.000
38	28.311	1.954	0.000
39	30.314	1.855	0.000
40	33.342	1.541	0.000
41	31.689	2.129	0.000
42	12.069	1.463	0.000
43	15.828	1.894	0.000
44	32.747	2.274	0.000
45	26.999	2.169	0.000
T			
1996	-0.2632	0.5235	0.615
1997	-0.6330	0.5310	0.234
1998	-0.0672	0.5368	0.900
1999	-0.1939	0.5625	0.730
2000	0.8225	0.5602	0.883
2001	-0.3195	0.5824	0.583
2002	0.4265	0.6190	0.491
2003	-0.7604	0.6310	0.229
2004	-1.2242	0.6442	0.058
2005	-1.1502	0.6817	0.092
2006	-1.1489	0.7111	0.107
2007	-1.7184	0.7390	0.020
2008	-1.8244	0.7589	0.016
2009	-1.7500	0.7710	0.024
2010	-2.3475	0.8139	0.004
2011	-2.4841	0.8315	0.003
2012	-3.0176	0.8596	0.000
2013	-3.0221	0.8976	0.001
2014	-3.1315	0.9200	0.001
F(65.742)=230.63			Prob >F=0.000

When we look at the table 2.6, time and country effects have been tested together .The Prob>F = 0.0000 is <0.005 , so the null hypothesis that the coefficients for all country and years are jointly equal to zero is rejected thus both country and year effects are included in the model (2). Country effects testing separately, country effect omitted because of collinearity but time and country effects have been tested together, country dummy variables significant for 38 countries. Also time dummy variables are significant after 2007. Therefore, the time effects and country effects are included in model (2) and it is shown in table 2.7.

Table 2.7.Fixed Effects Model with time and country effects for model (2)

Fixed-effects (within) regression Group variable: id			Number of obs=808 Number of groups=45
R-squared=0.952 Adj R-squared=0.948			
Gini	Coef.	Std. Err.	P> | t|
Hdi	-199.9885	25.72777	0.000
Hdisquare	140.92	18.45812	0.000
Cons	92.6396	9.7813	0.000

HDI, HDI square and constant term are significant at 1% level.HDI and HDI square coefficients shows that there is uninverted–u curve between economic development and income inequality. When HDI increase 1 unit the Gini decreases by 199.988 units. Coefficient of HDI square is 140.92, if HDI square increases 1% ,Gini increases by 140.92 units. R-squared is 0.952 thus 95% of changes in Gini are explained by independent variables which are the HDI and HDI square. Probability of F-statistic is zero therefore all the variables jointly in the model significantly affect the Gini at 1% significance level.

2.2.3. Descriptive Statistics

After determining the time and country fixed effects, basic assumptions of panel data analysis must be controlled in the model (2). Cross-sectional dependence, heteroskedasticity and serial correlation are controlled. Which tests to use for basic assumptions are shown in table 2.8

Table 2.8.Basic assumptions tests

Cross-sectional Dependence	Pesaran CD Test
Heteroskedasticity Test	Modified-Wald Test
Serial Correlation Test	Wooldrige Test for Autocorrelation

Basic assumptions tests results are shown in table 2.9.

Table 2.9.Basic assumptions test results

Cross-sectional Dependence	Pesaran's test of cross sectional dependence= 1.561	Pr = 0.1186
Modified-Wald Test	Chi2(45)=8316.48	P>chi2 = 0.0000
Wooldrige Test for Autocorrelation	F(1,43)=4.746	Prob>F = 0.0349

When N>T, Pesaran test is used for test correlation between units. The null hypothesis which residuals are not correlated are not rejected thus there is no cross-sectional dependence.

Heteroskedasticity is investigated by Modified-Wald test.Prob>chi2 = 0.0000 shows presence of heteroskedasticity. The null hypothesis which is homoskedasticity is rejected.

Autocorrelation is investigated by Wooldrige test.The null that is no serial correlation is rejected and conclude the data have first-order autocorrelation.

As a result of all these tests, heteroskedasticity and autocorrelation problem are found in model(2). Model (2) is re-estimated with Driscoll-Kraay estimator thus we get rid off the problems. It is shown table 2.10.

Table 2.10.Regression with Driscoll-Kraay standard errors

Group variable: id Max lag:2			Number of obs=808 Number of groups=45
R-sqr = 0.1355			
Gini	Coef.	Std. Err.	P> | t|
Hdi	-199.9885	62.23764	0.005
Hdisquare	140.92	46.77423	0.007
_cons	105.3764	20.69298	0.000
F test			F(21,19)=3264081.76 Prob >F=0.0000

Table 2.10 shows all variables which are HDI, HDI square and constant term are significant.HDI and HDI^2 coefficients shows that there is uninverted–u curve between economic development and income inequality within countries in used from 1995 to 2014. It is mean that income inequality first decrease and after the turning point, income inequality start to increase with economic development. Turning point is 0.7095 for model (2). Income inequality has decreased until HDI reach to 0.7095 then it starts to increase.

As it is mention before, Model (1) and Model (2) are used in this study.

In order to examine Kuznets hypothesis empirically, model (1) is used:

$$\boldsymbol{GINI_{it} = \alpha_0 + \alpha_1 \ln GDP_{it} + \alpha_2 \ln GDP_{it}^2 + \varepsilon_{it}} \quad (1)$$

Firstly fixed effect model and random effect model are estimated which belongs to model (1) then Hausman test is used to decide which of fixed effect model or random effect model is more appropriate.

Fixed effect model and random effect model are shown in table 2.11 and 2.12.

Table 2.11.Fixed Effect Model

Fixed-effects (within) regression Group variable: id			Number of obs=812 Number of groups=45
R-sqr: Within=0.0618 Between=0.5047 Overall=0.4445			Obs per group: Min=7 Avr=18 Max=20
Gini	Coef.	Std. Err.	P> | t|
lnGDP	-15.368	5.756	0.008
lnGDP square	0.663	0.310	0.033
_cons	122.487	26.549	0.000
F test			F(44,765)=157.98
			Prob >F=0.0000

LnGDP, lnGDP square and constant term are significant at 5% level. When lnGDP increase 1% unit the Gini decreases by 0,153 units. Coefficient of lnGDP square is 0.663, if lnGDP square increases 1% ,Gini increases by 0.663 units. LnGDP and LnGDP square coefficients show that there is uninverted–u curve. Probability of F-statistic is zero therefore all the variables jointly in the model significantly affect the Gini at 1% significance level.

Table 2.12.Random Effect Model

Random-effects GLS regression Group variable: id			Number of obs= 812 Number of groups= 45
R-sqr: Within=0.0598 Between=0.5275 Overall=0.4729			Obs per group: Min=7 Avr=18 Max=20
Gini	Coef.	Std. Err.	P> |z|
lnGDP	-9.780	5.653	0.084
lnGDP square	0.323	0.303	0.286
_cons	100.654	26.242	0.000

While inverse relationship (negative) is found between gini and lnGDP but there is same directional relationship between gini and lnGDP square. When lnGDP increase 1% unit the Gini decreases by 0,097 units. Coefficient of lnGDP square is 0.323, if lnGDP square increases 1%, Gini increases by 0.032 units. Uninverted-u curve is observed as like fixed effect model. R-squared is lower than fixed effect model (2).

Hausman test is used to decide which of fixed effect model or random effect model is more appropriate.

Table 2.13.Hausman Test

	Coefficients	
	Fixed	Random
LnGDP	-15.36861	-9.780621
lnGDP square	0.66322002	0.3237956
Chi2(2)=25.41, Prob>chi2=0.0000		

According to Hausman test, fixed effect used for model(1).After determining the fixed effect model, time effect and country effect in fixed effect model have been controlled. Results are shown in below tables which are 2.14 and 2.15.

Table 2.14.Time-fixed Effects

F(19,746)=1.32	Prob>F =0.1648

Null hypothesis which is the coefficients for all years are jointly equal to zero is accepted; therefore time fixed effects are not needed for model (1).

Table 2.15.Country-fixed Effects

Country effects in model (1) omitted because of collinearity

After look at the time and country effects separately, both time and country effects have been considered together. These results are shown in table 2.16.

Table 2.16.Fixed effect model with time and country effects

Fixed-effects (within) regression Group variable: id			Number of obs=812 Number of groups=45
R-squared=0.9504 Adj R-squared=0.9461			
Gini	Coef.	Std. Err.	P> | t|
LnGDP	-15.922	6.649	0.017
lnGDPsquare	0.728	0.379	0.055
Cons	111.795	28.694	0.000
Id			
2	2.493	0.759	0.001
3	-0.020	0.763	0.979
4	1.042	1.050	0.322
5	2.980	0.750	0.000

Id	Coef.	Std. Err.	P> | t|
6	0.300	0.756	0.691
7	1.490	0.760	0.050
8	3.679	0.764	0.000
9	1.648	0.761	0.031
10	2.643	1.599	0.099
11	7.197	1.060	0.000
12	0.777	0.894	0.385
13	6.278	0.768	0.000
14	5.380	0.813	0.000
15	6.311	0.782	0.000
16	6.805	0.771	0.000
17	7.420	0.772	0.000
18	4.733	0.974	0.000
19	10.809	0.983	0.000
20	-1.189	0.978	0.224
21	2.476	1.092	0.024
22	19.983	0.811	0.000
23	0.556	1.001	0.579
24	8.751	0.953	0.000
25	11.244	2.109	0.000
26	7.959	0.833	0.000
27	2.099	1.172	0.074
28	11.503	1.044	0.000
29	7.365	0.895	0.000
30	7.107	1.302	0.000
31	18.584	1.469	0.000
32	29.579	1.573	0.000
33	16.666	2.039	0.000
34	4.690	1.658	0.005
35	6.573	1.596	0.000
36	28.665	1.757	0.000

37	21.303	1.651	0.000
38	21.722	1.802	0.000
39	23.457	1.841	0.000
40	27.554	1.553	0.000
41	24.534	1.994	0.000
42	6.762	1.464	0.000
43	7.195	2.106	0.001
44	25.196	2.135	0.000
45	20.258	1.966	0.000
T			
1996	-0.319	0.531	0.548
1997	-0.411	0.534	0.441
1998	0.181	0.534	0.441
1999	0.220	0.551	0.689
2000	0.658	0.547	0.229
2001	0.431	0.553	0.436
2002	1.253	0.580	0.031
2003	0.281	0.574	0.624
2004	-0.055	0.583	0.925
2005	0.234	0.601	0.697
2006	0.415	0.624	0.506
2007	0.039	0.641	0.951
2008	0.061	0.643	0.924
2009	0.120	0.618	0.846
2010	-0.232	0.641	0.717
2011	-0.275	0.648	0.671
2012	-0.711	0.656	0.279
2013	-0.540	0.665	0.418
2014	-0.498	0.680	0.464
F test			F(65.746)=219.81
			Prob >F=0.000

When we look at the table 2.16, time and country effects have been tested together. Country effects testing separately, country effect omitted because of collinearity but time and country effects have been tested together country dummy variables significant in some countries even though time effect is not significant for all countries. Also coefficient of LnGDP square is insignificant in table 2.16. Therefore, time and country effects are not included in model (1), fixed effect model is used for model (1).

After determining fixed effects model, basic assumptions of panel data analysis must be controlled in the model (1). Cross-sectional dependence, heteroskedasticity and serial correlation are controlled. Its results are shown in table 2.17.

Table 2.17.Basic assumptions test results

Cross-sectional Dependence	Pesaran CD test= 0.979	Pr = 0.3275
Modified-Wald Test	Chi2(45)=8876.00	P>chi2 = 0.0000
Wooldrige Test for Autocorrelation	F(1,43)=4.305	Prob>F = 0.0440

When N>T, Pesaran test is used for test correlation between units. Null hypothesis which is residuals are not correlated is accepted, thus we have not observe cross-sectional dependence in model (1).

Heteroskedasticity is investigated by Modified-Wald test.Prob>chi2 = 0.0000 shows presence of heteroskedasticity. The null hypothesis which is homoskedasticity is rejected.

Autocorrelation is investigated by Wooldrige test.The null that is no serial correlation is rejected and concludes the data have first-order autocorrelation.

As a result of all these tests, hetereskedasticity and autocorrelation problem are found in model (1).

Model (1) is re-estimated with Driscoll-Kraay estimator thus we get rid off the problems. It is shown table 2.18.

Table 2.18.Regression with Driscoll-Kraay standard errors

Group variable: id Max lag:2			Number of obs=808 Number of groups=45
R-sqr = 0.0618			
Gini	Coef.	Std. Err.	P> | **t**|
LnGDP	-15.368	8.550	0.088
LnGDPsquare	0.663	0.439	0.148
_CONS	122.48	41.538	0.008

When re-estimate the model (1), lnGDP square coefficient is insignificant. Also , R square is low.Therefore, while robustance test solve the heteroskedasticity and autocorrelation problem, on the other side it made coefficient of lnGDP square insignificant.

CONCLUSION

In this study, Kuznets hypothesis is investigated for 45 countries are employed which are at different levels of income for a period from 1995 to 2014.This sample includes 30 high income countries, 11 upper-middle income countries, and 4 lower-middle income countries.

Two quadratic functions are used in order to test Kuznets hypothesis. In the first model, income inequality as measured by the Gini coefficient and economic development as measured by GDP per capita. In the second model, Income inequality as measured by the Gini coefficient and economic development as measured by HDI. In the literature, economic development is usually measured by GDP but HDI has gain importance in recent years with Sen's capability approach. HDI consists of three components which are the education index, the health index, and the income index. Education index is calculated by arithmetic means of two indices which are years of schooling index and expected years of schooling index. Health indicator is life expectancy at birth also the standard of living is measured by gross national income per capita. It is shown HDI included both GDP index and other two indexes which related to human well-being.

As a result, when testing basic assumptions in the model which economic development measured by GDP, heteroscedasticity and autocorrelation problem arised. Therefore, the model is estimated again with robust estimator to remove the deviation from the assumptions. After reestimation, lnGDP and lnGDP square coefficients become insignificant so that there is no significant relationship between economic development and income inequality in this model. HDI better explained the relationship between economic development and income inequality. HDI components have been increased the explanatory power of the model.

U relationship has been found between economic development and income inequality for 45 countries surveyed from 1995 to 2014. It means that income inequality decrease to a certain point in the early phases of economic growth then starts to increase with economic development.

BIBLIOGRAPHY

Ahluwalia, M. S., 1976. Income Distribution and Development:Some Stylized Facts. *The American Economic Review.*

Anon., 2016. *Human Development Report 2016:Human Development for Everyone,* s.l.: United Nations Development Programme.

Asteriou, D. & Hall, S. G., 2007. *Applied Econometrics a Modern Approach.* revised edition ed. New York: Palgrave Macmillan.

Baltagi, B. H., 2005. *Econmetric Analysis of Panel Data.* 3rd ed. s.l.:John Wiley& Sons, Ltd.

Barro, R. J., 2000. Inequality and Growth in a Panel of Countries. *Journal of Economic Growth,* pp. 5-32.

Chen, B.-L., 2003. An Inverted-u relationship between inequality and long-run growth. *Economics Letters 78,* pp. 205-212.

Dawson, P., 1997. On testing Kuznets' economic growth hpothesis. *Applied Economics Letters,* pp. 409-410.

Deininger, K. & Squire, L., 1996. A Data Set Measuring Income Inequalty. *The World Bank Economic Revieew,* vol. 10(3), pp. 565-91.

Deininger, K. & Squire, L., 1998. New ways of looking at old issues :inequality and growth. *Journal of Development Economics*, Volume 57, pp. 259-287.

Greene, W. H., 2002. *Econometrics Analysis.* 5th edition ed. New Jersey: Prentice Hall.

Gujarati, D. N., 2004. *Basic Econometrics.* s.l.:The McGraw-Hill .

Hoechle, D., 2007. Robust standart errors for panel regressions with cross-sectional dependence. *The Stata Journal*, 7(3), pp. 281-312.

Hsiao, C., 2014. *Analysis of Panel Data.* s.l.:Cambridge Univesity Press.

Huang, H.-C. R., 2004. A flexible nonlinear inference to the Kuznets hypothesis. *Economics Letters 84,* pp. 289-296.

Jonathan, H. & Shahidur, K., 2009. *Handbook on poverty and inequality.* Washington: World Bank.

Kuştepeli, Y., 2006. Income Inequality, Growth, and the Enlargement of the European Union. *Emerging Markets Finance and Trade,* 42(6), pp. 77-88.

Kuznets, S., 1955. Economic growth and income inequality. *The American Economic Review,* pp. 1-28.

List, J. A. & Gallet, C. A., 1999. The Kuznets curve :What Happens After the Inverted-U. *Review of Development Economics,* 3(2), pp. 200-206.

List, J. A. & Gallet, C. A., 1999. The Kuznets Curve :What Happens After the Inverted-U. *Review of Development Economics,* 3(2), pp. 200-206.

Matyas, L., Konya, L. & Macquarie, L., 1998. The Kuznets U-curve hypothesis :some panel data evidence. *Applied Economics Letters,* pp. 693-697.

Oskooee, M. B. & A.Gelan, 2008. Kuznets inverted-u hpothesis revisited: a time-series approach using USA data. *Applied Economics Letters,* pp. 677-681.

Price, A. G. G., 2003. African Economics and the Kuznets curve: an exploratory investigation. *Appled Economics Letters,* pp. 747-751.

Ram, R., 1989. Level of Development and Income Inequality:An extension of Kuznets-Hpothesis to the World Economy. *KYKLOS,* Issue Vol. 42, pp. 73-88.

Ram, R., 1997. Level of Economic Development and Income inequality :Evidence from the Postwar Developed World. *Southern Economic Journal,* 64(2), pp. 576-583.

Sarafidis, V., Yamagata, T. & Robertson, D., 2008. A Test of Cros Section Dependence for a Linear Dynamic Panel Model with Regressors. *Journal of Econometrics.*

Sen, A., 1992. Inequality Reexamined. *Harvard University Press.*

Sen, A., 1997. From income inequality to ecenomic inequality. *Southern Economic Journal,* 64(2).

Stiglitz, J. E., Sen, A. & Fitoussi, J.-P., n.d. *Report by the Commision on the Measurement of Economic Performance and Social Progress,* s.l.: s.n.

Theyson, K. C. & Heller, L. R., 2015. Development and Income Inequality : A New Specification of the Kuznets Hpothesis. *The Journal of Developing Areas,* 49(3), pp. 103-118.

Thornton, J., 2001. The Kuznets inverted-u hpothesis:panel data evidence from 96 countries. *Applied Economics Letters,* pp. 15-16.

Tribble, R., 1999. A Restatement of the S-curve Hpothesis. *Review of Development Economics,* 3(2), pp. 207-214.

Weil,D., 2014.Economic Growth.3rd edition. Pearson

Zang, H., 1998. The Stability of Kuznets Curve: some further evidence. *Applied Economics Letters,* pp. 131-133.