The impact of Income Inequality on Human Development

Lina Arustamyan

May 2020

American University of Armenia

College of Business and Economics

Instructor: Hayk Kamalyan

Course: Econometrics

Abstract

If we look to the history, we can find that there has always been overwhelming evidence that high levels of economic inequality leads to low levels of human development. The main purpose of our research paper was to observe the effect of income inequality on human development levels by compiling GINI coefficient and Human Development Index statistics on 71 randomly selected developed and developing countries from across the world for year 2018.

During our model specification procedure we used a number of OLS and 2SLS models and concluded that GINI was indeed an endogenous variable affecting by several factors.As a result,we found that the GINI index has a negative impact on the human development index, but more so in developing countries than in developed ones. Furthermore,after running additional variables, including urban population,economic freedom index and variables showing the level of health, education and income level in the countries, we found that there are differences in significance with these variables with respect to their impacts on human development levels.

Keywords: Human Development, Income Inequality, Cross country analysis, 2SLS models

Contents

1 Introduction

There are many problems related to the use of GDP as a measure of social well-being and the use of consumption expenditure as a measure of individual well-being. The problems associated with the use of GDP are clear: only market transactions are computed; quality is not considered; distributive aspects are omitted, and stocks or durable goods and infrastructures are almost out of the picture. GDP captures an important part of the economic performance of society, but it is far from being a complete measure of economic development and surely further from being a competent measure of human development and social well-being.

The criticisms of GDP as a measure of human development have led individuals and international organizations to develop alternative measures of well-being that try to more accurately measure human development, the most well-known of which is the Human Development Index (HDI) (Sen, 1984; Alkire, 2007). It was developed and launched by Pakistani economist Mahbub-ul-Haq, followed by Amartya Sen, an Indian economist, in 1990, who developed the Capabilities Approach, an economic theory positing that freedom to achieve well-being is a matter of what people can do (Sen, 1984; Alkire, 2007).

With guidance from Sen, the UNDP published the Human Development Index (HDI) in 1990, the first attempt to develop a comprehensive indicator to measure a country's progress in human development (Sakiko, 2003). HDI consists of three aspects, which are health, income and education. Essentially four parameters are used in the calculation of HDI for measuring and ranking countries according to their social and economic development which includes the Life Expectancy at Birth, Expected Years of Schooling, Mean Years of Schooling and Gross National Income per Capita.

Even though HDI is transparent and simple to calculate and interpret, it has often been criticized for ignoring inequality in the distribution of human development across countries. Inequalities in health, education, and income are the key components of human development,that deeply impact progress towards increasing human development. "The world's richest 1 percent of people receive as much income as the poorest 57 percent. The income of the world's richest 5 percent is 114 times that of the poorest 5 percent."(2001 HDR, page 19). Our paper will mostly dwell on the impact of income inequality on human development.

All over the world, economic inequality has consistently been a hot topic of debate for governments and politicians. Besides, every country has always wanted to be able to reach high levels of human development. For instance, governments in the past such as in the Soviet Union have had complete control over their countries' economies and politics, leaving the middle and lower classes to suffer significantly in the wake of the rich's luxury.

We know that people and, therefore, countries are not identical in more aspects than just income, such as culture and genetic talents, but it is difficult to determine how these factors affect income production, life expectancy and education levels. This leaves us with just income inequality as the simplest quantifiable factor that can be included in the analysis. So the question becomes: Does income inequality play a role in producing low human development? And if so, is income inequality the main reason for the low level of human development, or are there other, more significant factors?

To analyse how the income inequality affects human development, we have used the GINI index as a proxy for income inequality and HDI as the main measure of human development. Moreover, for the multiple regression analysis, we wanted to generally introduce a list of other important social factors, along with the proxy variables for main components of HDI, that we believed could potentially affect the human development index. Furthermore, to solve the endogeneity problem coming from the GINI index we have also used some instrumental variables that explain the GINI index. We also hope to illustrate the differences between developed and developing nations.

2 Literature Review

A number of studies concentrated on the exploring the factors affecting HDI in the literature.

Subramanian(2003) in his paper analysed the important determining factors of human development to evaluate their relative impact on the human development and suggest measures to increase human development in the countries with high, medium and low HDI. Life expectancy at birth has been the most effective variable in determining the value of HDI in countries with high HDI and countries with low HDI, combined enrolment ratio is considered to be the dominant variable in the determination of HDI in countries with medium HDI. Life expectancy at birth, adult literacy rate and combined enrolment ratio had the greatest impact on HDI in countries with low human development, per capita real GDP had the greatest influence on HDI in countries with medium HDI for the whole world.

Grimm, Harttgen, Klasen, and Misselhorn (2006, 2008) developed a separate HDI for each income quintile. The results showed that in all countries inequality in human development by income quantile was very high, and was especially high in developing countries. In this research, inequality was calculated by the ratio between the HDI for the richest quintile and the poorest quintile (80/20).

Brueckner and Lederman (2018) tried to justify the predominant economic theory developed by Galor and Zeira (1993). The predominant theory is that poorer countries will have a positive correlation between inequality and aggregate output, while richer countries will have a negative correlation between inequality and aggregate output. The instrumental variables showed that long term GDPPP growth and overall country growth are negatively impacted by inequality for countries with a GDPPP of 10,000 USD. Their model estimates demonstrated that poor countries (nations will low initial GDP per capita) have a small to no relationship between inequality and GDPPP. The main conclusion from this paper would be that income inequality helps growth for poor nations, but hurts growth for richer nations.

Gould (2007) argue that workers have higher wages in cities than in rural areas. This could come as cities make workers more productive or select workers with certain skills and abilities. The model in the study works by choosing between urban or rural areas by career choices over time. The researcher controlled for all sources of selection and endogeneity, the results show that for white-collar jobs, the city pay better. Cities don't pay better for blue-collar workers. Because of this, the research suggests that people move to cities not only because they like the location, but also to have a higher wage.

Srinivasan (1994) criticized the index for its unequal, arbitrary weighting strategy. He wasn't agree with giving income, education and life expectancy the same onethird weight. He, finally, failed to see how income impacts a country's human development as much as education.

Cysne Turchick (2012) present the strong positive correlation between unemployment and income inequality in a diversified range of economies across the globe. However, the article makes clear that this relationship is only observed in situations where the unemployment rate is no larger than 15 percent. Moreover, the article explains how unemployment is relatively higher among low-skilled workers who often suffer from longer periods of unemployment due to technological progress, etc.

Lam(1997) explores the relationship between demographic variables and income distribution within the United States. The study investigates how a changing population composition may adjust income inequality, focusing particularly on age distribution, fertility, marriage, migration, and mortality. Some analyses state that the age distribution of a population may change the overall levels of income inequality within a country, without changing the levels of income inequality between age groups. However, further analysis find out that in some cases, a larger, younger workforce may actually decrease wages.

From our study of literature, it becomes clear that many scientists have examined the factors affecting HDI, also by criticizing the ignorance of income inequality in the distribution of human development across countries. Some of them concluded that income inequality helps growth for poor nations, but hurts growth for richer nations(Brueckner and Lederman (2018)). Srinivasan criticized the fact that all main components of HDI have the same weight in the calculation of HDI, as they may differently impact the HDI. Moreover, for solving endogeneity issue coming from GINI index we have also studied the literature regarding the main factors that influence income inequality to get a better model.

Based on these studies as the basis for our research, we hope to find out the possible negative effect of income inequality on human development level in different countries, as well as how the health, income and education aspects of HDI really influence human development.

3 Data and Methodology

The data used in this analysis was obtained from various sources, including the World Bank, the United Nations Development Programme, the Human Development Report 2019 and the Heritage Fund. These data were collected for 71 developing and developed countries of the world for 2018. List of countries can be found in Appendix I(Table 4). The data sources for the variables used in our analysis are listed below in **Table 1**.

Variables	Description	Abbreviation	Source
HDI	Human Development Index for year 2018	*HDI*	United Nations
Fertility rate	Births per woman as of 2018	*Fertilityrate*	HDR 2019
CPI	Consumer Price Index(2010 =100) for year 2018	*CPI*	World Bank
Population with at least secondary education	(% ages 25 and older) for year 2018	*seced*	HDR 2019
Government expenditure on education	(as % of GDP) for year 2018	*exped*	HDR 2019
GINI	Income Inequality by country for year 2018	*GINI*	United Nations
Index of Economic Freedom	Index of Economic Freedom by country for year 2018	*EconFreedom*	Heritage Foundation
Urban Population	% of total population in cities for year 2018	*Urbanpopoftotal*	World Bank
Developing/Developed	Developing or developed countries based on income level	*dev*	World Bank

Table 1: Variables and sources of data used in regression analysis

The Human Development Index is a composite measure of development in a country. The factors that are included in the calculation of HDI are as follows: GNI per capita, mean and expected years of schooling and life expectancy at birth. The geometric mean of the three categories is taken to create the index score. However, it was illogical to use the main components as independent variables in the regression, as all the main components would explain almost 100 percent variation in HDI ,which could diminish the effect of other variables used in the regression.Instead, we used the proxy variables explaining health, income and education components of HDI.The proxy variables were determined by finding correlations between that particular HDI component and a variety of other variables related to that component. For income component we took the Consumer Price Index, for education component we chose the Government expenditure on education and Population with at least secondary education, and, finally, for the health component we took Fertility rate.

The GINI coefficient, which is our main independent variable, is the measure of income inequality. A GINI coefficient of 1 means that there is a perfect inequality where one person has control of all income, 0 is when a country has perfect income equality. We transformed the scale of GINI by keeping it in a range from 0 to 100. The next variable that we took is Index of Economic Freedom. Index of Economic Freedom is an index created by the Heritage Foundation to measure the extent to which people can pursue goals and projects without government intervention policy. Higher Index means less government intervention, and therefore higher overall well-being, which leads to an increase in income component of the HDI. We also decided to use the urban population as an additional variable because of its relevance to urban economic theory, which states that large urban population sizes lead to positive effects on employment, consumption, overall GDP, which leads to higher levels of overall income across these populations, which, in turn, leads to higher HDI levels. To get a more detailed understanding of human development, the dummy variable was created and data was separated based on the level on economic development, whether a country was developed or developing. Using information set by the World Bank, countries that fall into the upper middle income category or higher are considered developed, the rest are considered developing.

Furthermore, as we also used 2sls model for solving the endogeneity issue coming from income inequality , after studying the literature we took some instrumental variables explaining the GINI index.They are Median age, Unemployment rate, Life expectancy at birth and GDP per capita. We expected life expectancy to have a negative effect on income inequality, since longer life time implies a possible increase of welfare programs within a country. In addition, we chose the median age and unemployment rate to accent the relationship between age distribution, workforce demographics and income inequality. As was previously indicated in our literature review, in some cases a large, young workforce decreases wages and strengthens income inequality in the country, so we expected median age to have a negative impact on GINI. We also hypothesized unemployment would have a positive relationship, which would mean the more people are without work, the higher the income inequality in the country. Finally, we supposed that as income levels (in our case GDP per capita) fall, the income inequality (GINI) will rise. To avoid inappropriate scaling/huge numbers we used the natural logarithm of

GDP per capita in our analysis.

Variables	Obs.	Mean	Std.Dev.	Min	Max
HDI	71	77.7662	15.14178	37.65911	95.36883
Fertilityrate	71	2.296501	1.174677	1.052	7.001
CPI	71	144.2546	56.47253	99.18695	382.5008
seced	71	73.26093	27.0043	6.62	100
expeded	71	4.873239	1.416329	1.3	7.7
GINI	71	35.12958	7.40977	25	63
EconFreedom	71	65.19155	9.323178	44	81.7
Urbanpopoftotal	71	63.57001	22.35682	16.425	98.001
dev	71	0.5915493	0.4950459	0	1

Table 2: Descriptive Statistics on HDI and factors affecting it

Table 2 contains the descriptive statistics on HDI and independent variables used in our regression models to explain human development level in 71 countries that we chose. As we can observe, our dependent variable HDI illustrates significant variation ,ranging from 37.66 to 95.37, which indicates the fact that it is meaningful to observe the factors affecting HDI. Put simply, as more than half of the countries in our dataset are developed countries, such a wide variation in HDI shows that the same level of industrialization don't fully explain the differences in development level in different countries. The variation in the explanatory variables is also large, which will help us in explaining the variation in HDI. Appendix(Figure 4) contains the descriptive statistics of instrumental variables.

4 Model Selection Procedure and Results

We estimated a several models via OLS and 2SLS to obtain relatively the best model which will explain the impact of income inequality on human development. The results of the models made with OLS are presented in Table 5. The STATA outputs of all the estimated models are located in Appendix (Figure 5-8)

Independent variables	Model 1	Model 2	Model 3	Model 4
Constant	110.7633*** *(se = 7.841578)*	57.54408*** *(se = 7.435509)*	60.62984*** *(se =5.761502)*	78.52018*** *(se =6.569146)*
GINI	-0.9392973*** *(se =0.2184785)*	-0.118607 *(se =0.0894995)*	-0.1029341 *(se =.0880855)*	-0.0702949 *(se =0. 0784354)*
CPI		-0.0256181** *(se = 0.0121756)*	-0.023689* *(se =0.0119675)*	-0.2211046*** *(se =0.0468176)*
seced		0.1961761*** *(se =0.0343011)*	0.1988933*** *(se = 0.033671)*	0.203691*** *(se =0.0298639)*
expeded		-0.2496799 *(se = 0.4332324)*		
Fertilityrate		-3.18868*** *(se =0 .7519122)*	-3.332644*** *(se =0.7309925)*	-3.393479*** *(se =0.6480479)*
EconFreedom		0.0660635 *(se =0.0627802)*		
Urbanpopoftotal		0.282287*** *(se =0.0336501)*	0.2713849*** *(se = 0.032315)*	0.2467627*** *(se =0.0894995)*
CPI2(CPI^2)				0.0004506*** *(se =0 .0292007)*
R^2	0.2113	0.9164	0.9144	0.9338
N	71	71	71	71

* p<0.10, ** p<0.05, *** p<0.01

Table 3: Models estimated via OLS

Model 1 is the simple linear regression of the HDI on the GINI coefficient.

HDI=110.76-0.93GINI

As we can observe, there is a negative relationship between GINI and HDI in Model 1. 1 unit increase in the GINI coefficient leads to about 0.94

unit decrease in HDI. The coefficient of GINI is significant at 1% percent significance level and GINI singly explains 21% of the variation in HDI. These results indicates that GINI plays a significant role alone in explaining HDI ,but other factors are also needed to explain the remaining 79% of the variation in HDI. The positive sign of the constant term is also logic, which means that the perfect income equality will result in the higher level of human development. However, the constant exceeds permissible interval of HDI, which could possibly the result of just mathematical calculation and not economic relation or there are factors that better explain HDI, that should be considered in the regression analysis.

In **Model 2** we have already added other factors explaining HDI, including health, education, income components and some social factors.

HDI=57.54-0.11GINI - 0.02CPI+ 0.19seced-3.18Fertilityrate+0.28Urbanpopoftotal-0.24expeded +0.06EconFreedom

The significance of the most of the independent variables included in this model and high level of R squared indicates the relevance of that factors in affecting HDI. Moreover, the correlation matrix shows that there isn't a perfect collinearity between explanatory variables.

	GINI	Urbanp~l	EconFr~m	Fertil~e	expeded	seced	CPI
GINI	1.0000						
Urbanpopof~l	-0.2454	1.0000					
EconFreedom	0.0380	-0.0936	1.0000				
Fertilityr~e	0.4005	-0.5907	-0.1588	1.0000			
expeded	-0.2009	0.3348	-0.0351	-0.3577	1.0000		
seced	-0.5276	0.5921	0.1082	-0.7159	0.3764	1.0000	
CPI	0.2292	-0.4328	-0.0197	0.5494	-0.3478	-0.5081	1.0000

Figure 1: Correlation matrix of explanatory variables

The signs of the significant variables are meaningful and consistent with our expectations. For example, the increase in inflation by 1 unit (percentage point) will lead to overall decrease in national output(GNI) which in turn will decrease the human development level by about 0.02 unit .1 percentage point increase in population with at least secondary education and urban population will lead to an increase in the HDI by 0.19 and 0.28 unit respectively. In Model 2 the highest predictive power has the Fertility rate. The negative effect of Fertility rate on human development may have 2 reasons. First, according to World Health Organization "about 295 000 women died during and following pregnancy and childbirth in 2017", which means increasing fertility rate can negatively affect life expectancy.The second reason is that expenses on one additional child may decrease the healthcare expenses needed for the parents. The sign of our main variable's coefficient is still meaningful, that is 1 unit increase in GINI will decrease HDI by about 0.12 unit. However, when accounting for more factors, GINI coefficient became insignificant, which may be because of potential endogeneyity coming from GINI which we will check in the subsequent models by 2SLS.

In **Model 3** we have already dropped the insignificant variables expended and EconFreedom after checking also their jointly significance by F test. We have got that F value is equal to about 0.76, which is smaller than F critical value (3.15), so we didn't reject the null hypothesis that expended = EconFreedom =0.

HDI=60.62-0.10GINI - 0.02CPI+ 0.19seced-3.33Fertilityrate+0.27Urbanpopoftotal

The signs and the significance levels of the remaining variables still remain the same as in Model 2. Although we have decrease in R squared value, which is logical as we reduced the number of explanatory variables, the adjusted R squared in model 3 is higher than in Model 2 which indicates the better fit. After dropping the insignificant variables in Model 3, we have checked if there is a linear relationship between HDI and all the explanatory variables used in the model. As a result , we observed that there is an exact curvilinear relationship between HDI and CPI. The STATA output of the scatter plot is located in Appendix(Figure 9) .

In **Model 4** we have also added the square term of the CPI because of the curvilinear relationship between HDI and CPI.

HDI=78.52 -0.07GINI - 0.22CPI+0.0004CPI^2+$0.20 seced - 3.39 Fertilityrate + 0.24 Urbanpopoftotal$

The coefficient of the square term of CPI is also significant even at 1% significance level. Inclusion of the square term of CPI has already changed the partial effect of consumer price index on HDI, that is 1 percentage point increase in CPI will decrease HDI by -0.2211046 + 2*0.0004506=-0.2202034.Moreover, Model 4 has the relatively highest goodness of fit measure(R^2 = 93.38).

The coefficients interpretation logic is similar with Model 2. However, GINI coefficient is still insignificant, so we decided to use 2SLS model because of the potential endogeneity problem coming from GINI.

Estimating models via 2SLS

Our main explanatory variable, GINI, is indeed a potentially endogenous variable . Thus, by studying some literature we include four instrumental variables that explain GINI and use 2SLS model using the form of Model 4 to check our hypothesis . The possible effect of instrumental variables was already introduced in Data Methods section.Appendix(Figure 10). contains the STATA output regressing GINI on instrumental variables to understand the effect of that variables on GINI.

We have got, that only Unemployment rate and Median age have a significant impact on GINI and signs of the coefficients are consistent with our expectations.

In the first step regression of 2SLS we regress GINI on the independent variables used in Model 4 along with the interaction term and four instrumental variables that we chose. After that we save the predicted value of GINI. In the second step regression we regress HDI on CPI, CPI2, seced, Fertilityrate, Urbanpopoftotal ,gdev and already predicted value of GINI.As a result our 2SLS model will be the following

HDI=95.96-0.53 $\hat{GINI}-0.19CPI+0.0003CPI^2+0.13seced-3.06Fertilityrate+0.26Urbanpopoftotal$

In this 2SLS model(**Model 5**) coefficient of GINI is already significant compared with OLS Model 4, which indicates that GINI is indeed endogenous variable.For checking endogeneity we have additionally used the Hausman test which is provided in Appendix (Figure 12). Moreover, 2SLS model has increased the predictive power of GINI, that is 1 unit increase in GINI index reduces the human development level by 0.53 unit compared with only -0.07 in Model 4. Besides GINI the second most powerful predictor is Fertility rate which shows the health component of HDI.

Finally, we used another 2SLS model(**Model 6**) with the interaction term ginihat*dev(ginidev) included to capture the difference of the GINI(already predicted value) coefficient on HDI in developed and developing countries.

HDI= 84.2-0.37$\hat{GINI}$$-0.13CPI+0.0002CPI^2+0.13seced-2.9Fertilityrate+$ $0.2Urbanpopoftotal + 0.185ginidev$

From this model we can observe, that the effect of income inequality in developing countries is by a factor of -0.37, while in developed countries by a factor of -0.37+0.185 = -0.185. Thus, we can conclude that the effect of income inequality on human development level is relatively bigger in developing countries rather than in developed ones. The effects of other variables included in the model is almost the same as in Model 4 and 5.

All the variables included in Model 5 and 6 are significant and the R squared values are 0.9476 and 0.9531 respectively.The possible reasons for such an unrealistically good results with very high R squared values are introduced in Conclusion part of our research paper.The STATA output of regression results are provided in Appendix (Figure 11 and 13).

Checking 2SLS assumptions

After estimating our Model 5 and 6 we checked 2SLS assumptions to ensure that our estimators are consistent.

2SLS.1. Zero Conditional Mean E(u$\|x_1, \ldots x_k) = 0$

From the residual histograms in Figure 2 we can visually observe that residuals are normally distributed with mean 0. In addition, we have performed the Jarque-Bera test for normality which confirmed the normality of the residuals. Test results are provided in Appendix(Figure 14)

2SLS.3. No perfect collinearity

We have performed a VIF test for testing the mullticollinearity. The test surely will show the correlation between CPI and its squared term, but no perfect collinearity assumption refers to an exact linear relationship. Thus,

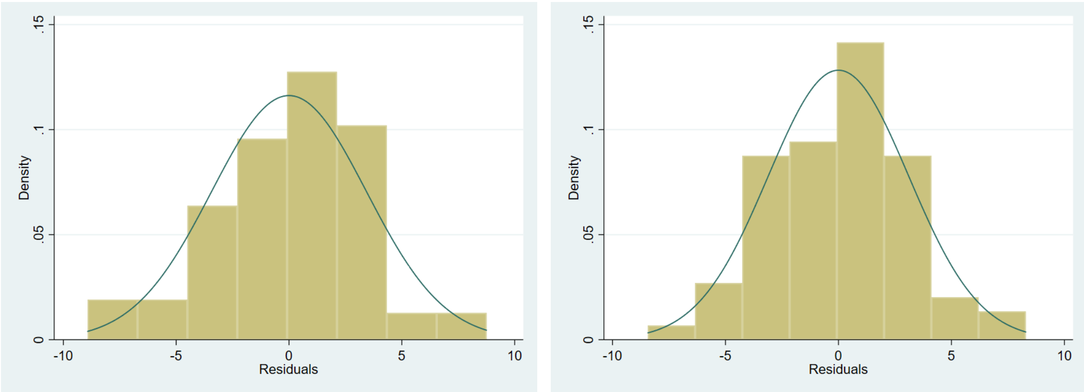

Figure 2: Residual Histograms of Model 5 and 6

based on the VIF test we conclude that we don't have mullticollinearity. Test results are provided in Appendix (Figure 15)(VIF > 10 refers to mullticollinearity).Under 2SLS.1 and 2SLS.2 our estimators are **consistent**.

2SLS.3. Homoskedasticity assumption

From the residuals plots in Figure 3 we can observe the random pattern of the residuals , which implies homoscedasticity. Adittionaly we have used Breusch-Pagan / Cook-Weisberg test for heteroskedasticity. The results of the tests are located in Appendix(Figure 16).Under 2SLS.3. our estimators are also **efficient**.

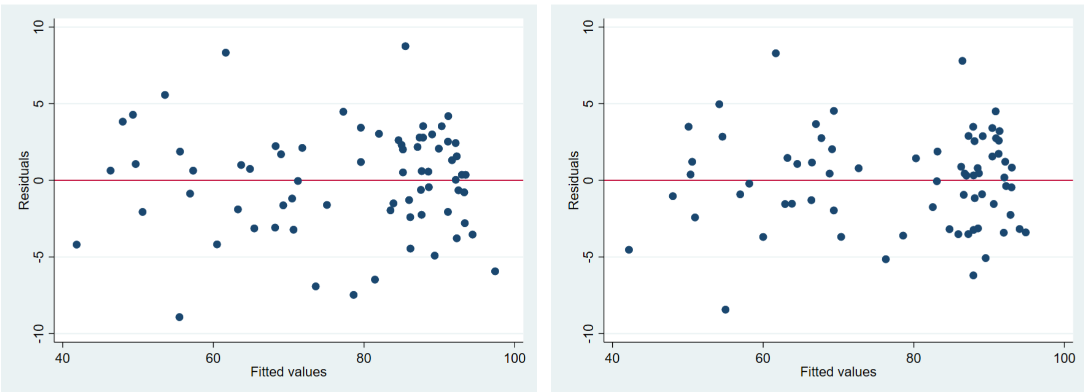

Figure 3: Residual plots of Model 5 and 6

5 Conclusion

We conclude that GINI index has a significant negative effect on Human Development Index when it is also predicted by factors such as Median age, Life Expectancy, Unemployment rate and GDP per capita. The negative effect of income inequality on human development level is explained by the fact that in countries with high level of income inequality people don't have an equal opportunities to get the same level of education or equal access to the healthcare services , which, in turn, reduces the overall human development level in the country.In addition, the negative effect of income inequality on human development level was larger in developing nations rather than in developed ones. After adding additional variables , we have found that Economic Freedom index and Government expenditure on education didn't have a significant effect on HDI. Moreover, among the significant variables representing the health, education and income aspects of human development, the highest predictive power has Fertility rate, which is negatively correlated with life expectancy.

However, all our estimated models have a suspiciously high goodness of fit measure(R squared), which may have several reasons. The first reason is that the variables that we chose for representing health , education and income levels are correlated with the main components included in the calculation of HDI, so they probably explain the big part of variation in HDI. The next reason is that there is actually a two way causality between GINI index and Human Development Index, which can also affect estimation results.

A References

Milorad Kovacevic:Human Development Research Paper 2010/35 Measurement of Inequality in Human Development – A Review

Suman Seth and Antonio Villar 2017: Measuring Human Development and Human Deprivations(OPHI WORKING PAPER NO. 110)

Determinants of the Levels of Development Based on the Human Development Index: A Comparison of Regression Models for Limited Dependent Variables: Article in Review of European Studies January 2014

EPAR Request No. 309:Review of Human Development Indices

https://databank.worldbank.org/source/world-development-indicators

http://hdr.undp.org/en/data

HDR 2019

https://www.heritage.org/index/explore

Gould, E. (2007). Cities, Workers, and Wages: A Structural Analysis of Urban Wage Premium. Oxford

University Press: The Review of Economic Studies 74(2) 477-506

Brueckner, M. and D. Lederman (2018). Inequality and Economic Growth: The Role of Initial Income.

Journal of Economic Growth 25(3) 341-366

Ishan Parikh, Vivek Srinivasan, Jay Patel(2018):Estimation of the effect of Income Inequality on Human Development: A Cross Sectional Study.

Srinivasan, T.N. (1994) “Human Development: A New Paradigm or Reinvention of the Wheel?”

American Economic Review, 84(2), 238243. Retrieved April 5, 2016.

Alexa Grzech, Kanhai Patel, James Walker(2016):A CrossCountry Analysis: the Effect of Income, Life Expectancy, and Education on the Human Development Index

Lam, David. "Chapter 18 Demographic Variables and Income Inequality." Handbook of Population and Family Economics Handbook of Population and Family Economics Volume 1, 1997, pp. 1015–1059. ScienceDirect, doi:10.1016/s1574-003x(97)80010-4.

Cysne, Eubens Penha, and David Turchick. "Equilibrium Unemployment-Inequality Correlation." Journal of Macroeconomics, vol. 34, no. 2, June 2012, pp. 454–469. ScienceDirect, doi: 10.1016/j.jmacro.2011.12.009.

B Appendix

Developed	Developing
Norway	Ukraine
Switzerland	Tunisia
Ireland	Mongolia
Germany	Philippines
Australia	Moldova (Republic of)
Iceland	Indonesia
Sweden	South Africa
Netherlands	Egypt
Denmark	Viet Nam
Finland	Morocco
Canada	Kyrgyzstan
United Kingdom	El Salvador
United States	Tajikistan
Belgium	Nicaragua
Japan	India
Austria	Honduras
Luxembourg	Bangladesh
Israel	Zambia
Korea (Republic of)	Nepal
Slovenia	Angola
Spain	Zimbabwe
Czechia	Pakistan
France	Rwanda
Malta	Lesotho
Italy	Sudan
Estonia	Malawi
Cyprus	Ethiopia
Greece	Gambia
Poland	Niger
Lithuania	
Slovakia	
Latvia	
Portugal	
Chile	
Hungary	
Croatia	
Argentina	
Russian Federation	
Belarus	
Kazakhstan	
Romania	
Uruguay	

Table 4: List of countries

Variable	Obs	Mean	Std. Dev.	Min	Max
Medianage	71	34.62977	9.463032	15.151	48.358
Unemployme~e	71	6.537592	4.698634	.273	26.958
Lifeexpect~h	71	75.4751	7.106438	53.705	84.47
lgdp	71	9.747077	1.174133	6.969247	11.63812

Figure 4: Descriptive Statistics on instrumental variables affecting GINI

Source	SS	df	MS			
				Number of obs	=	71
				F(1, 69)	=	18.48
Model	3390.88894	1	3390.88894	Prob > F	=	0.0001
Residual	12658.2582	69	183.453018	R-squared	=	0.2113
				Adj R-squared	=	0.1999
Total	16049.1472	70	229.273531	Root MSE	=	13.544

HDI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
GINI	-.9392973	.2184785	-4.30	0.000	-1.37515	-.5034446
_cons	110.7633	7.841578	14.13	0.000	95.1198	126.4068

Figure 5: Model 1 output

Source	SS	df	MS			
				Number of obs	=	71
				F(7, 63)	=	98.67
Model	14707.6339	7	2101.09055	Prob > F	=	0.0000
Residual	1341.51329	63	21.2938618	R-squared	=	0.9164
				Adj R-squared	=	0.9071
Total	16049.1472	70	229.273531	Root MSE	=	4.6145

HDI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Fertilityrate	-3.18868	.7519122	-4.24	0.000	-4.691257	-1.686103
CPI	-.0256181	.0121756	-2.10	0.039	-.0499491	-.0012871
seced	.1961761	.0343011	5.72	0.000	.1276309	.2647214
expeded	-.2496799	.4332324	-0.58	0.566	-1.115426	.616066
GINI	-.118607	.0894995	-1.33	0.190	-.2974576	.0602435
Urbanpopoftotal	.282287	.0336501	8.39	0.000	.2150427	.3495313
EconFreedom	.0660635	.0627802	1.05	0.297	-.0593928	.1915198
_cons	57.54408	7.435509	7.74	0.000	42.6854	72.40276

Figure 6: Model 2 output

Source	SS	df	MS			
				Number of obs	=	71
				F(5, 65)	=	138.85
Model	14675.1895	5	2935.03789	Prob > F	=	0.0000
Residual	1373.95768	65	21.1378105	R-squared	=	0.9144
				Adj R-squared	=	0.9078
Total	16049.1472	70	229.273531	Root MSE	=	4.5976

HDI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Fertilityrate	-3.332644	.7309925	-4.56	0.000	-4.792537	-1.872751
CPI	-.023689	.0119675	-1.98	0.052	-.0475897	.0002117
seced	.1988933	.033671	5.91	0.000	.1316477	.2661389
GINI	-.1029341	.0880855	-1.17	0.247	-.278853	.0729848
Urbanpopoftotal	.2713849	.032315	8.40	0.000	.2068475	.3359223
_cons	60.62984	5.761502	10.52	0.000	49.12333	72.13636

Figure 7: Model 3 output

Source	SS	df	MS			
				Number of obs	=	71
				F(6, 64)	=	150.42
Model	14986.4143	6	2497.73572	Prob > F	=	0.0000
Residual	1062.7328	64	16.6052001	R-squared	=	0.9338
				Adj R-squared	=	0.9276
Total	16049.1472	70	229.273531	Root MSE	=	4.0749

HDI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Fertilityrate	-3.393479	.6480479	-5.24	0.000	-4.688104	-2.098855
CPI	-.2211046	.0468176	-4.72	0.000	-.3146335	-.1275758
seced	.203691	.0298639	6.82	0.000	.144031	.2633511
GINI	-.0702949	.0784354	-0.90	0.373	-.2269876	.0863978
CPI2	.0004506	.0001041	4.33	0.000	.0002427	.0006586
Urbanpopoftotal	.2467627	.0292007	8.45	0.000	.1884276	.3050978
_cons	78.52018	6.569146	11.95	0.000	65.3968	91.64356

Figure 8: Model 4 output

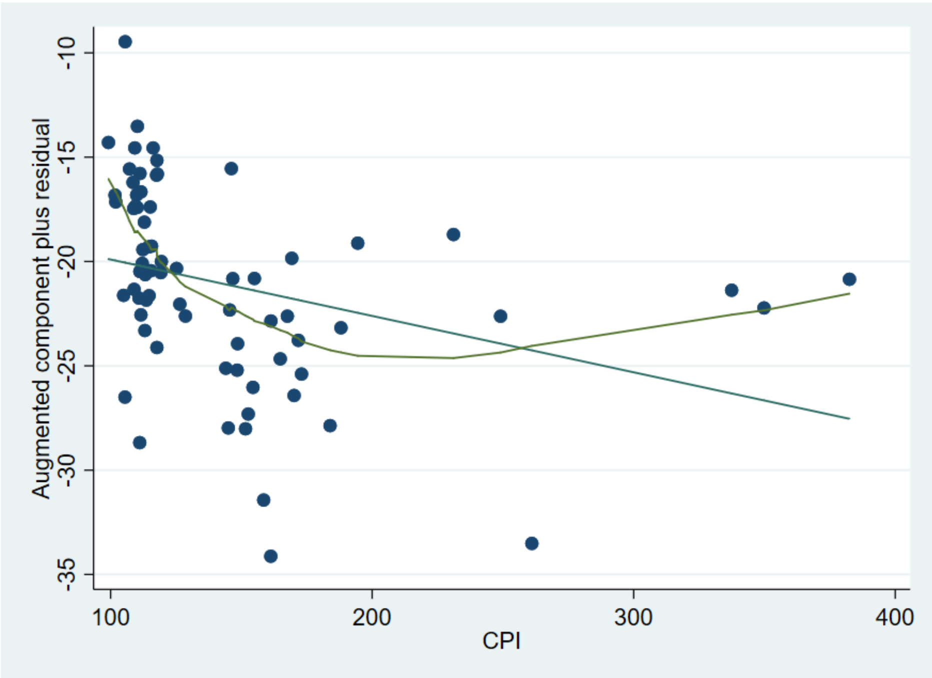

Figure 9: Model fitting of CPI

Source	SS	df	MS			
				Number of obs	=	71
				F(4, 66)	=	11.80
Model	1602.40334	4	400.600834	Prob > F	=	0.0000
Residual	2240.92455	66	33.9534023	R-squared	=	0.4169
				Adj R-squared	=	0.3816
Total	3843.32789	70	54.9046841	Root MSE	=	5.827

GINI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Medianage	-.4044252	.1425031	-2.84	0.006	-.6889418	-.1199087
Unemploymentrate	.5938978	.1573645	3.77	0.000	.2797095	.9080861
Lifeexpectancyatbirth	-.0012206	.1923633	-0.01	0.995	-.3852863	.3828451
lgdp	.1765341	.6026093	0.29	0.770	-1.026614	1.379682
_cons	43.6235	12.51926	3.48	0.001	18.62799	68.61902

Figure 10: Regression on instrumental variables

```
. reg GINI Medianage Unemploymentrate Lifeexpectancyatbirth lgdp Fertilityrate seced CPI CPI2 Urbanpopoft
```

Source	SS	df	MS			
				Number of obs	=	71
				F(9, 61)	=	7.57
Model	2027.72196	9	225.30244	Prob > F	=	0.0000
Residual	1815.60592	61	29.7640315	R-squared	=	0.5276
				Adj R-squared	=	0.4579
Total	3843.32789	70	54.9046841	Root MSE	=	5.4556

GINI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Medianage	-.5140711	.1958208	-2.63	0.011	-.9056389	-.1225032
Unemploymentrate	.4357951	.1566654	2.78	0.007	.1225232	.7490671
Lifeexpectancyatbirth	-.4064868	.2231919	-1.82	0.073	-.8527865	.039813
lgdp	.1807843	.5902622	0.31	0.760	-.9995178	1.361086
Fertilityrate	-2.457642	1.301804	-1.89	0.064	-5.060761	.1454771
seced	-.0938703	.0400912	-2.34	0.022	-.1740377	-.0137029
CPI	-.1079925	.0761171	-1.42	0.161	-.260198	.044213
CPI2	.0001805	.0001654	1.09	0.279	-.0001502	.0005113
Urbanpopoftotal	.1077983	.0470605	2.29	0.025	.013695	.2019016
_cons	95.92246	23.48706	4.08	0.000	48.95717	142.8877

```
. predict ginihat
(option xb assumed; fitted values)

. reg  HDI Fertilityrate CPI seced  ginihat CPI2  Urbanpopoftotal
```

Source	SS	df	MS			
				Number of obs	=	71
				F(6, 64)	=	196.79
Model	15223.9711	6	2537.32852	Prob > F	=	0.0000
Residual	825.176031	64	12.8933755	R-squared	=	0.9486
				Adj R-squared	=	0.9438
Total	16049.1472	70	229.273531	Root MSE	=	3.5907

HDI	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Fertilityrate	-3.067362	.5752982	-5.33	0.000	-4.216652	-1.918072
CPI	-.1987858	.0415304	-4.79	0.000	-.2817523	-.1158192
seced	.1339706	.0302569	4.43	0.000	.0735256	.1944156
ginihat	-.5328952	.1208036	-4.41	0.000	-.7742282	-.2915623
CPI2	.0003916	.0000926	4.23	0.000	.0002067	.0005766
Urbanpopoftotal	.268132	.0261347	10.26	0.000	.2159218	.3203421
_cons	95.96546	6.889704	13.93	0.000	82.20169	109.7292

Figure 11: 2SLS Model 5

```
. ivregress 2sls HDI   seced  Fertilityrate CPI  Urbanpopoftotal  CPI2  ( GINI = Medianage  Unemploymentra

Instrumental variables (2SLS) regression          Number of obs     =         71
                                                  Wald chi2(6)      =     658.95
                                                  Prob > chi2       =     0.0000
                                                  R-squared         =     0.8978
                                                  Root MSE          =     4.8066

------------------------------------------------------------------------------
            HDI |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
           GINI |  -.5328952    .161709    -3.30   0.001    -.8498391   -.2159514
          seced |   .1339706   .0405022     3.31   0.001     .0545878    .2133534
  Fertilityrate |  -3.067362   .7701004    -3.98   0.000    -4.576731   -1.557993
            CPI |  -.1987858   .0555931    -3.58   0.000    -.3077462   -.0898254
Urbanpopoftotal |    .268132   .0349843     7.66   0.000     .1995641    .3366998
           CPI2 |   .0003916   .0001239     3.16   0.002     .0001487    .0006346
          _cons |   95.96546   9.222632    10.41   0.000     77.88943    114.0415
------------------------------------------------------------------------------
Instrumented:  GINI
Instruments:   seced Fertilityrate CPI Urbanpopoftotal CPI2 Medianage
               Unemploymentrate Lifeexpectancyatbirth lgdp

. estat endogenous GINI

  Tests of endogeneity
  Ho: variables are exogenous

  Durbin (score) chi2(1)          =  18.7781  (p = 0.0000)
  Wu-Hausman F(1,63)              =  22.6537  (p = 0.0000)
```

Figure 12: Hausman test

```
. reg  HDI Fertilityrate CPI seced  ginihat ginidev CPI2  Urbanpopoftotal

      Source |       SS           df       MS      Number of obs   =        71
-------------+----------------------------------   F(7, 63)        =    204.27
       Model |  15371.8857         7  2195.98367   Prob > F        =    0.0000
    Residual |  677.261434        63  10.7501815   R-squared       =    0.9578
-------------+----------------------------------   Adj R-squared   =    0.9531
       Total |  16049.1472        70  229.273531   Root MSE        =    3.2787

---------------------------------------------------------------------------------
            HDI |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+----------------------------------------------------------------
  Fertilityrate |  -2.900388   .5272375    -5.50   0.000    -3.953988   -1.846788
            CPI |  -.1334155   .0418169    -3.19   0.002    -.2169798   -.0498512
          seced |   .1335729   .0276281     4.83   0.000     .0783625    .1887833
        ginihat |  -.3706874   .1186591    -3.12   0.003    -.6078087   -.1335661
        ginidev |   .1854103   .0499846     3.71   0.000      .085524    .2852965
           CPI2 |   .0002626   .0000914     2.87   0.006     .0000799    .0004453
Urbanpopoftotal |   .2017068   .0298357     6.76   0.000     .1420849    .2613287
          _cons |   84.26096   7.038058    11.97   0.000     70.19652     98.3254
---------------------------------------------------------------------------------
```

Figure 13: 2SLS Model 6

```
. jb res5
Jarque-Bera normality test:  .5533 Chi(2)  .7583
Jarque-Bera test for Ho: normality:

. jb res6
Jarque-Bera normality test:  .1911 Chi(2)  .9089
Jarque-Bera test for Ho: normality:
```

Figure 14: Jarque Bera test

. vif

Variable	VIF	1/VIF
CPI	29.86	0.033486
CPI2	28.61	0.034957
seced	3.62	0.275902
Fertilityr~e	2.48	0.403316
ginihat	2.30	0.435711
Urbanpopof~l	1.85	0.539525
Mean VIF	11.45	

. vif

Variable	VIF	1/VIF
CPI	36.31	0.027539
CPI2	33.45	0.029894
ginidev	4.22	0.236797
seced	3.62	0.275898
Urbanpopof~l	2.90	0.345164
ginihat	2.66	0.376535
Fertilityr~e	2.50	0.400376
Mean VIF	12.24	

Figure 15: VIF test

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of HDI

         chi2(1)      =     1.54
         Prob > chi2  =   0.2142

.
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of HDI

         chi2(1)      =     1.63
         Prob > chi2  =   0.2012

.
```

Figure 16: Heteroscedasticity test