Analyzing the Effect of Income Inequality on Poverty

Abigail Burke, Kaylin Berinhout, Patrick Bonnie

Abstract: This paper seeks to add to the body of work surrounding the relationship between income inequality and poverty. In this research, we hope to demonstrate how the percentage of people living below the poverty line is related to the GINI coefficient, change in GDP per capita, literacy rate, Freedom House score, infant mortality rate, and income level for a range of different countries.

I. Introduction

According to a 2018 report from wealth data company Wealth-X, there are now more billionaires on the planet than in any other time in history ("Global Billionaire" 2018). An increased ability to spread information globally has also ensured that people are more aware of the wealth of the ultra-rich in ways that were not possible in previous centuries. As discussions of billionaires, wealth gaps, and potential income taxes dominate the news cycle, uncovering the relationship between income equality and poverty has become more essential than ever. There is widespread concern in the economic community that economic growth, especially in emerging economies, has not been reflected in the incomes of large swaths of these countries populations. The GINI coefficient, a statistical measure intended to represent the income distribution of a nation's residents, has risen on average over the past few decades in OECD countries (OECD 2015). Our analysis of the relationship between poverty, inequality, and several other variables will hopefully reveal important, significant correlations; an understanding of this information will allow us to extrapolate potential methods by which poverty levels can be reduced in the future.

Our hypothesis for the simple regression model, which studies the effects of the GINI coefficient on the percentage of people living below the poverty line ($1.90 a day), predicted that the percentage of people below the poverty line would rise as the GINI coefficient rises. This predicted correlation follows economic intuition: if the discrepancy between the extremely rich and the poor rises, then it is likely that the number of people living below the poverty line will also rise. As for the multiple linear regression, it was predicted that the percentage of people living below the poverty line will decrease as GDP per capita rises due to the fact that people's standard of living typically rises as GDP per capita rises. We also predict that the percentage of people living below the poverty line will decrease as literacy rate increases because literacy rate is a good indicator of quality of an education system. Better education systems are known to correspond to better job opportunities and thus lower levels of poverty. We predict that a higher Freedom House score, a measurement of a country's political rights and civil liberties, will have an inverse relationship with the Poverty Headcount Rate. We believe that a higher Freedom House score will relate to a lower percentage of people living in poverty because countries that typically have freer, more open governments tend to be more developed, richer countries. We predict that a lower infant mortality rate will correlate with a lower Poverty Headcount Rate due to the fact that infant mortality can be used as a measure of the overall effectiveness of a country's healthcare system. It follows that a lower infant infant mortality rate indicates a better overall healthcare system which in turn would produce a healthier, more prosperous society. Finally, we predict that high income countries (those with a value of 1 for the variable HighIncome) will have lower Poverty Headcount Rates because these countries have more wealth than other countries in our data set.

II. Literature Review

For our research, it was necessary to meticulously review past writings about our topic. A related study on the effect of income inequality on poverty by Fosu (2010) uses the headcount level of poverty as the dependent variable and the GINI coefficient and PPP (purchasing power parity) adjusted mean income as independent variables. The article seeks to fill a gap by providing research that demonstrates exactly how poverty is affected by income inequality. Poverty in this study is defined as the headcount ratio of people living below $1 per day. Fosu draws two important conclusions from his analysis. First, he concludes that increased income inequality stunts the potential of income growth as a means to reduce poverty. Second, he concludes that a rise in inequality generally results in a rise in poverty. Aside from these two main conclusions, he notes that the inequality elasticity of poverty differs across regions and across countries. These differences mean that an equal increase in inequality across two countries can lead to a different increase in poverty in each.

To further our understanding of how and why different levels of income inequality lead to distinct poverty outcomes, we turn to a study by Chambers and Dhongde (2011). The researchers came to the critical conclusion that countries with higher income inequality have lower levels of growth elasticity of poverty (GEP). In the study, GEP is defined as “the extent to which poverty declines if income increases by 1 percent, for a given level of inequality.” While the findings of this study do affirm our hypothesis that higher income inequality contributes to poverty, the researchers also advise caution when testing these parametric variables through the use of a linear regression model . Our model attempts to explain the effect of our independent variables such as the GINI Index and average income (both of which have a decidedly non-linear impact) on our dependent variable: poverty. That being said, the precise parametric relationships between these variables remain unspecified, which implies the importance of using our simplified linear model to gain further insight and estimates.

Another explanatory variable we are focusing on is literacy rate. Literacy rate is a common measure of the strength of a country’s educational system. The effect of literacy rate on poverty was the focus of a study by Ahmad (2019) which sought to uncover the effects of literacy rate on poverty in Pakistan. This work incorporated previous studies on the effect of literacy rate on poverty in specific countries. For example, one such study cited by Ahmad found that poverty and literacy rate were inversely related in India. Ahmad’s research in Pakistan found that, although there was no short run relationship between poverty and literacy rate, an increased literacy rate resulted in a decreased poverty rate in the long run.

A paper by Arndt, McKay, and Tarp (2016) discussed the relationship between gross domestic product per capita (GDP per capita) and poverty. Focusing on Sub-Saharan Africa, the paper demonstrated that growth in per capita GDP over the past twenty years has only slightly changed the level of poverty in the region. The growth elasticity of poverty (the rate by which poverty declines for each percent of GDP per capita growth) was found to be just .54 in countries like Burkina Faso. The authors argue that perceptions that rapid growth in GDP per capita greatly affects poverty levels are mistaken; actual data demonstrates that growth in GDP per capita has a smaller effect on poverty than hypothesized. This fascinating study led us to include GDP per capita in our research because we are curious to see if these results in Sub-Saharan Africa remain consistent globally.

As demonstrated, there exists a large body of work concerning the effect of GDP, GINI coefficient, and literacy rate on poverty. Our research contributes to the overall economic literature in three ways. First, our research will provide a test to the conclusions reached by previous research, confirming results that match our own or opening the door for further analysis on results that contradict our findings. Second, we are contributing to previous research by adding a large analysis of cross-country data to the traditional analysis of poverty. As we reviewed various literature, we found that many analyses on poverty levels have only been conducted for individual countries or specific regions. Third, we are including Freedom House scores as a variable in our data. The Freedom House score is absent from all of the previous research on the subject that we analysed. We believe that Freedom House score, which measures how politically and civilly free a country is, will have an inverse relationship with Poverty Headcount Rate. By including this variable in our regression, we contribute to current literature by expanding the types of variables typically discussed in poverty analysis. In our later models, we include the variables country income level and infant mortality rate. Adding infant mortality rate may help make important inferences into where money allocated to poverty reduction will be best spent. Differentiating countries based on whether or not they are considered "high income" by the World Bank standards serves to make our prior research contributions more relevant and specific.

III. Data:

Poverty levels of a country are determined by a combination of factors that are rooted in each country's unique history and development. Countless qualitative factors might make one country more equal than another, including the colonial hierarchical influences within society or laws regarding intake and assimilation of refugees. Keeping these outside variables in mind, it becomes important to evaluate the factors impacting poverty levels that can be measured on a statistical level. Because of the large body of literature that attempts to understand why rises in total wealth of developing countries have not resulted in subsequent increases in income for most of the population, we decided to focus our simple regression model on the impact of a country's income inequality on its levels of poverty. Our dependent variable is poverty, measured through the use of the Poverty Headcount Ratio which represents the percentage of a country's population living under $1.90 per day at the 2011 international price level. The independent variable is income inequality, measured using the GINI Coefficient which shows the extent to which family income deviates from income in an economy with perfectly equitable distribution. The more equitable a country's income distribution, the closer the GINI Coefficient is to zero. We hypothesized that a higher GINI coefficient will correlate to higher poverty levels. In the multiple regression model, we included the independent variables literacy rate, log(GDP per capita), and Freedom House Score, in addition to the GINI coefficient. The descriptive statistics of these variables in relation to our model are depicted in the table below.

Variable Descriptions			
Variable	Description	Source	Year
Poverty (dep.)	Percentage of people living below the poverty line (PPP adjusted $1.90/day)	World Bank	2016, 2017, 2018
GINI (ind.)	Statistical measure that represents the income distribution of a nation's residents	World Bank	2016, 2017, 2018
log(GDPperCapita) (ind.)	The natural log of GDP per Capita (defined as the total monetary value of all final goods and services produced and sold on the market within a country during a certain year, divided by a country's total population), which shows the percent change in GDP per Capita over time	World Bank	2017, 2018
FreedomScore (ind.)	A measurement derived from the United Nations' Universal Declaration of Human Rights (on a scale from 0-100) of how politically and civilly free a country is	Freedom House	2018
LiteracyRate (ind.)	The percentage of adults ages 15 and over who can read and write with understanding a short simple statement about their everyday life	World Bank	2018
InfantMortality (ind.)	The number of deaths under one year of age occurring among the live births in a country during a given year (per 1,000 live births occurring among the population)	World Bank	2018
HighIncome (ind.)	High income countries are those with a GNI per capita of $12,376 or more (adjusted using the World Bank Atlas Method)	World Bank	2018

Summary Statistics					
Variable	Obs.	Mean	Std. Dev.	Min	Max
Poverty	80	6.38	13.58	0	70.3
GINI	80	37.20	7.22	25.4	59.1
LiteracyRate	80	90.46	13.71	42	100
GDPperCapita	80	19070.06	23194.12	389.4	114340.5
log(GDPperCapita)	80	9.10	1.35	5.96	11.65
FreedomScore	80	68.01	26.87	12	100
InfantMortality	80	13.99	14.16	1	61
HighIncome	80	.40	.49	0	1

In examining poverty levels, literacy rate was an important variable because it evaluates the ability of a country's workforce to improve their own intellectual growth and economic opportunities as well as the ability of a country's education system to provide the necessary skills for communication and job advancement. Adult literacy rate is an indication of the percentage of a country's population that is educated. Education allows someone to not only work more complex jobs, but to advocate for themselves and their families financially and legally. This ability allows populations to lift themselves out of the cycle of poverty. Considering this context, we assumed that countries with higher literacy rates would have lower levels of poverty..

We decided to include the natural logarithm of Gross Domestic Product (GDP) per capita as an independent variable in our regression upon the assumption that the percent change of the size of the economy of a country adjusted to its population size might be correlated with its poverty levels. We hypothesized a negative relationship between the two variables, assuming that increased change in wealth per person leads to decreased percentage of the population living on less than $1.90 a day.

The Freedom House Score attempts to assess freedom on an individual level instead of a governmental level. In our search for variables that impact poverty levels, we chose to include Freedom House Scores in attempts to find statistical evidence that proved our assumption that individual freedom reduces the percentage of a population living below the poverty levels.

We also added two more variables, HighIncome and InfantMortality. Inclusion of HighIncome, a dummy variable, allows us to differentiate the impact of our independent variables on the Poverty

Headcount Rate based on a country's average income level. This addition allows us to make important inferences into which explanatory variables become more important as countries become wealthier. As a result, important policy implications can be formed for countries at different stages of development.

We included InfantMortality because we predicted that an indicator of the overall quality and effectiveness of a country's health system would affect poverty levels. We hypothesized that healthy populations would likely save more and be more productive. Omitting this variable would leave out an important part of the picture when evaluating the level of development of a country. Groups like the World Health Organization have argued that improved access to healthcare (represented in our data by lower values of InfantMortality) is linked to poverty reduction ("Health" 2010).

We obtained data for Poverty Headcount Ratio, GINI coefficient, GDP per capita, infant mortality rate, and adult literacy rate from the World Bank's Development Research Group (DRG). These researchers secured their data through the use of primary household survey data obtained from government statistical agencies and World Bank country departments. This data may be problematic in that it includes self-reported data. Corrupt governments may skew their results in order to give off the impression that their country is more developed than it actually is.

Freedom House Scores are derived from an annual global report entitled *Freedom in the World.* This report uses methods from the United Nations' Universal Declaration of Human Rights. Scores are compiled by Freedom House analysts, using news sources, academic analyses and reports from nongovernmental organizations. Although scores are measured through a rating process which emphasizes unbiased methods and consistency, it is unavoidable that the analysts, advisers, and staff collecting the data bring their own subjectivity into the process, creating an imperfect data set.

Before analyzing the validity of our hypothesis through regression analysis, it is necessary to check to make sure that our data and variables fit with the Gauss-Markov assumptions. Before evaluating these assumptions on the basis of multiple linear regression, we will assess their equivalents in terms of simple linear regression (SLR).

Assumption SLR.1 states that the model is linear in parameters. This means that the simple linear regression equation must be written as follows:

$$y = \beta_0 + \beta_1 x + u$$

This condition is satisfied by our model because our independent variable, income inequality, and our dependent variable, poverty, are linearly related in our estimated simple regression equation.

Assumption SLR.2 requires that the sample of data we use is randomly drawn from the population. This assumption holds true for our data because the World Bank and Freedom House use controlled data acquisition methods which ensure unbiased sampling from a wide and representative set of data.

Assumption SLR.3 states that there is sample variation in the explanatory variable, meaning that there are different sample outcomes for each instance of x. Our data fits this assumption because there is plenty of variation in the GINI coefficients across the set of countries that we examined.

Assumption SLR.4 holds that the data must have a zero conditional mean. This means that u, the error term, will have an expected value of zero given any value of the independent variable. This can be shown in the following equation:

$$E(u|x) = 0$$

This assumption is not satisfied by our model. The u term, or the unobserved term, includes variables which are conditional on our independent variable: income inequality. The u term may contain information such as the qualitative variables mentioned above (colonial legacy, and structure regarding intake of refugees), however, it likely also contains data concerning factors such as the ethnic homogeneity of a country's population or the strength of laws regarding patents for technological innovation. The latter two factors are likely to be related to our independent variable: income inequality, as well as the variable it attempts to describe: poverty. This means that $E(u|x) \neq 0$, and assumption SLR.4 is violated. This violation means that we are uncertain that our estimators are unbiased. This will lead to either and underestimation or an overestimation of the coefficients on our variables in our simple regression model. Regardless of this, we still gain valuable information on the general relationship between income inequality and poverty by running a simple regression model.

Assumption SLR.5 brings homoskedasticity into the equation. This means that the error u has the same variance for any value of the independent variable, represented below.

$$Var(u|x) = \sigma^2$$

For our model, we are able to assume that this assumption holds true. It is likely that the value of the variance of the residuals is constant for each country's GINI coefficient, meaning that the explanatory variable's value are unrelated to that of the unobserved factors.

The Gauss Markov assumptions for multiple linear regression (MLR) paint the same picture as those for simple linear regression for assumptions MLR.1 and MLR.2. Even with our additional independent variables, our model is still linear in its parameters, and part of a randomly drawn sample from the population.

MLR.3 states that there is no multicollinearity in the model. This means that none of the independent variables are constant, and there is no exact linear relationships among the independent variables.

We can assess the validity of this by examining the correlation coefficients between the independent variables.

MultiCollinearity Table						
	GINI	LiteracyRate	log(GDPperCapita)	FreedomScore	InfantMortality	HighIncome
GINI	1.0000					
LiteracyRate	-0.2082	1.0000				
log(GDPperCapita)	-0.3672	0.7505	1.0000			
FreedomScore	-0.1561	0.4463	0.6845	1.0000		
InfantMortality	0.3592	-0.8862	-0.8081	-0.4829	1.0000	
HighIncome	-0.4437	0.4631	0.8079	0.7392	-0.5993	1.0000

Seeing as none of the correlation coefficients are equal to one, we can confirm that there is no perfect collinearity between independent variables.

MLR.4 and SLR.4 are based on the same requirement that there is no information about the mean of the unobserved factors found in the independent variables. Given that we add the independent variables LiteracyRate, log(GDPperCapita), InfantMortality, HighIncome, and Freedom Score, there will be fewer sources of unexplained factors contained in the error term. These additions will decrease the bias in u, but only to a certain extent. MLR.4, like SLR.4, will still not be completely fulfilled.

MLR.5 is no different from SLR.5 except for the addition of new independent variables into the assumption that the value of the explanatory variables are unrelated to that of the unobserved factors. We can still assume that this is a true statement for our model.

IV. Results

The following table shows the results for the simple regression of the GINI coefficient on poverty in our model. We regress the GINI coefficient on poverty as a means to demonstrate the effect of inequality on poverty.

Simple Linear Regression 1				
Variable	**Coefficient**	**Std. Err.**	**t-score**	**P > |t|**
GINI	0.620799	.2009792	3.09	0.003

Intercept	-16.7198	7.615093	-2.20	0.031
R^2	0.1090			

The traditional regression equation is:

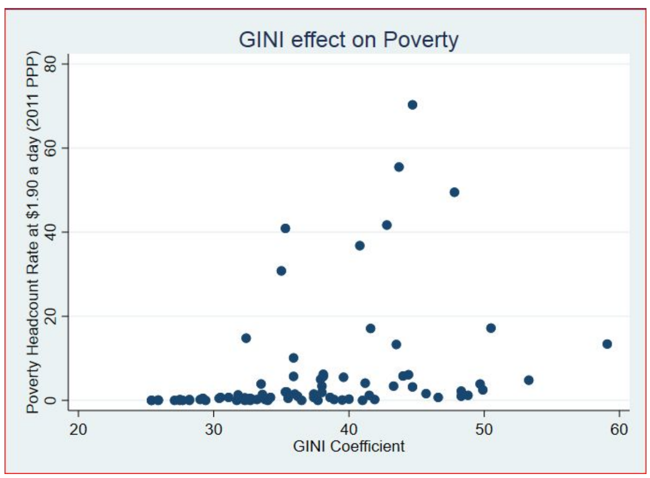
$$y = \beta_0 + \beta_1 x + u$$

Our regression yielded the following equation:

$$Poverty = -13.79 + 0.62 GINI + u$$

For our equation, poverty is a variable name for Poverty Headcount Rate and GINI is a variable name for the GINI Coefficient. Therefore, y = Poverty and x = GINI.

The normal simple regression model shows that there is a positive correlation between the GINI coefficient and the Poverty Headcount Rate meaning that as the GINI coefficient rises, the Poverty Headcount Rate will also rise. More specifically, the correlation coefficient on GINI is 0.62. This reveals that an increase in the GINI coefficient by 1 will result in an increase in poverty headcount rate by 0.62 percent. This fairly strong positive relationship can be seen in the following scatter plot.



However, we know that the simple regression model will not yield a ceteris paribus effect of GINI coefficient on poverty unless there are no other variables which have an effect on poverty. Therefore, it is necessary to use a multiple linear regression to attempt to uncover a ceteris paribus effect on poverty.

For our multiple linear regression, we continue to use Poverty as the dependent variable and now use GINI, LiteracyRate, log(GDPperCapita), FreedomScore, InfantMortality, and HighIncome as explanatory variables. The results of our first multiple regression model are as follows:

Multiple Linear Regression I (MLR I)				
Variable	**Coefficient**	**Std. Err.**	**t-score**	**P > |t|**
GINI	0.2099268	0.1599463	1.31	0.193
LiteracyRate	-0.4718798	0.168305	-2.80	0.006
log(GDPperCapita)	-5.318058	1.604515	-3.31	0.001
FreedomScore	0.1381265	0.0530972	2.60	0.011
InfantMortality	-0.0864877	0.1826027	-0.47	0.637
Intercept	81.45898	20.58873	3.96	0.000
R^2	0.5861			

The traditional formula for a multiple regression is:

$$y = \square_0 + \square_1 x_1 + \square_2 x_2 + \ldots + \square_n x_n + u$$

Our multiple linear regression yielded the following equation:

$$Poverty = 81.46 + 0.21GINI - 0.47LiteracyRate - 5.32\log(\text{GDPperCapita}) + .14FreedomScore - 0.086InfantMortality + u$$

The coefficients on the explanatory variables reveal a lot of information about each variable's relationship to the dependent variable, Poverty. $\square_1$, the coefficient on GINI, is equal to 0.21. This tells us that for an increase in the GINI coefficient by 1, the Poverty Headcount Rate will raise by 0.21 percentage points. This is a fairly strong positive correlation. We predicted a positive relationship between these two variables in our hypothesis. $\square_2$, the coefficient on LiteracyRate, is -0.47. This means that an increase in literacy rate by 1 percentage point will result in a decrease in Poverty Headcount Rate by .47 percentage points. This is a fairly strong negative correlation. We predicted an inverse relationship between these variables in our hypothesis. $\square_3$, the coefficient on log(GDPperCapita), is -5.32. This is an extremely strong negative correlation. It means that an increase in GDPperCapita by 1% will result in a decrease in Poverty Headcount Rate by 5.32 percentage points. We predicted a negative relationship between these two variables in our hypothesis $\square_4$, the coefficient on FreedomScore, is 0.14. This means that an increase in Freedom House score by 1 point will result in an increase in Poverty Headcount Rate by .14 of a percentage point. This shows a positive correlation between Freedom House score and Poverty Headcount Rate, which is the opposite of what we hypothesized. The coefficient on InfantMortality, $\square_4$, is -0.086.

This is a weakly negative relationship which says that a one percent increase in the infant mortality rate will decrease by .086 percentage points.

In this model, GINI and InfantMortality are insignificant at any level below 10% as determined by their P-values (significance levels for all variables for all models can be seen in the table labeled "Cross-Regression Significance Table". We decided to remove InfantMortality in order to refine the model because it was insignificant. We kept GINI in the model because it is our main explanatory variable and our SLR showed a significant relationship between Poverty and GINI. This alteration yields MLRII for which the data is below.

Multiple Linear Regression II (MLR II)				
Variable	**Coefficient**	**Std. Err.**	**t-score**	**P > |t|**
GINI	0.1895238	0.1532377	1.24	0.220
LiteracyRate	-0.4133549	.113681	-3.64	0.001
log(GDPperCapita)	-5.054608	1.497127	-3.38	0.001
FreedomScore	0.1368909	.0527581	2.59	0.011
Intercept	73.40045	11.53368	6.36	0.000
R^2	0.5849			

This regression yields the equation:

$$Poverty = 73.40 + 0.19GINI - 0.41LiteracyRate - 5.05log(GDPperCapita) + 0.14FreedomScore + u$$

This equation shows us that for a 1 point increase in GINI coefficient, there will be a 0.19 percentage point increase in Poverty Headcount Rate. This is a decently strong positive correlation and corresponds to what we found in our last two models. $\square_2$, the correlation coefficient on LiteracyRate, is -0.41. This shows that a 1 percentage point increase in a country's literacy rate will result in a decrease in Poverty Headcount Rate by 0.41 percentage points. $\square_3$, the coefficient on log(GDPperCapita), is -5.05. This is a very strong negative correlation which means that for a one percent increase in GDP per capita in a country, the Poverty Headcount Rate will decrease by 5.05 percentage points. Finally, $\square_4$, the coefficient on FreedomScore, is 0.14. This is a positive correlation that means that a one point increase in FreedomScore resulted in a 0.14 percentage point increase in Poverty Headcount Rate. It is important to note that for both MLR I and MLR II the coefficient on FreedomScore was positive, which means that higher Freedom House Score means higher level of poverty. This refutes our hypothesis and will be discussed more in detail later in the paper.

We determined the significance of all of our variables for each regression using the P-values in the above table. If the P-value is less than 0.01, then the variable's coefficient is significant at the 1% level. If the P-value is less than 0.05, then the variable's coefficient is significant at the 5% level. Finally, if the P-value is less than 0.10, then the variable's coefficient is significant at the 10% level. These significance levels are shown for all regressions in the tables below, along with the t-values and standard errors of each variable.

Cross-Regression Significance Table			
Variable	**SLR**	**MLR I**	**MLR II**
GINI	0.620799*** (.2009792)	0.2099268 (0.1599463)	0.1895238 (0.1532377)
LiteracyRate		-0.4718798*** (0.168305)	-0.4133549*** (0.113681)
log(GDPpeprCapita)		-5.318058*** (1.604515)	-5.054608*** (1.497127)
FreedomScore		0.1381265** (0.0530972)	0.1368909** (0.527581)
InfantMortality		-0.0864877 (0.1826027)	
Intercept	51.16705** (10.3217)	81.45898*** (20.58873)	73.40045*** (11.53368)
R^2	0.1090	0.5861	0.5849
Adjusted R^2	.0976	0.5582	0.5627
Significant at: * 10%, ** 5%, *** 1%			

From this table, we see that GINI is significant in the original SLR model, but loses its significant in the MLR models. This tells us that GINI is likely an important factor in poverty but is also very related to one of the other variables in our MLR models, likely log(GDPperCapita). Additionally, we see that LiteracyRate, log(GDPperCapita), and FreedomScore are highly significant in the models they are used in. However, InfantMortality was not significant in the model we included it in.

GINI coefficient was not significant in our MLR I and MLR II models. We will discuss this variable more extensively later in paper as it pertains to its relevance to Poverty. Across the board, log(GDPperCapita) was a significant variable with a negative coefficient. In terms of policy recommendations from this, we can say that governments should do their best to encourage policies such as large public works projects that employ citizens, increased trade, or encouraging widespread investment into the economy which will raise GDP per capita because our data demonstrated that growth in GDP per capita lowers Poverty Headcount Rate. Additionally, literacy rate was significant in all the models it was included in, which could suggest that countries or philanthropic organizations that want to support communities living in poverty should focus more heavily on improving early childhood education and the quality of public school systems.

V. Extensions:

F-Test:

We decided to complete an F-test in order to find the joint-significance of literacy rate and infant mortality rate. These are two variables which we believed would be have a strong negative correlation both due to our multicollinearity table as well as the fact that higher literacy rate typically means a better education system which leads to better healthcare systems and thus lower infant mortality rate. Proving this statistically would further the pragmatic policy applications of our research. In order to conduct the F-test, we considered MLR1, our unrestricted model, and a restricted model which did not include literacy rate or infant mortality. The Sum of Squared Residuals (SSR) for our unrestricted model is 6032.93, and 7117.94 for our restricted model. We considered the following hypotheses:

$$H_0 : \beta_2 = \beta_5 = 0$$

$$H_1 : \text{null hypothesis is false}$$

We then calculated our F-statistics using the following formula:

$$F = \frac{(SSRr - SSRur)/q}{SSRur/n-k-1}$$

$$F = \frac{(7117.94-6032.93)/2}{6032.93/80-5-1} = 6.65$$

An F-value of 6.65 is larger than the critical value of 3.12, which means the literacy rate and infant mortality are jointly related. This means that infant mortality rate, as a representative for a country's overall health system, is important in considering which factors affect poverty level even if it is not statistically significant in our MLR models. Therefore, countries should be concerned with the quality and adequacy of their health systems when working to reduce poverty within their borders.

Dummy Variable Extended:

The following chart shows the data from the regression using Poverty as the dependent variable and HighIncome, our dummy variable, as the independent variable.

Simple Linear Regression II

Variable	**Coefficient**	**Std. Err.**	**t-score**	**P > |t|**
HighIncome	-9.819792	2.915004	-3.37	0.001
Intercept	10.30417	1.843611	5.59	0.00
R^2	0.1270			

As seen in the above table, HighIncome has a strongly negative relationship with Poverty. The coefficient, -9.82, means that high income countries have a baseline poverty level that is 9.82 percentage points lower than countries that are not high income. Additionally, this coefficient is significant at the 1% confidence level as seen in the P value in the table. This has important implications for our data, specifically, it raises the question if our data would better be viewed as two distinct data sets: high income and not high income. Due to this, we decided to perform a Chow Test to analyze this question.

Chow Test:

Using HighIncome, our dummy variable, which takes on a value of 1 when a country is a high income country and 0 when it is not, we decided to conduct a Chow Test using our MLR II model as our pooled model. We selected MLR II over MLR III because MLR II still has GINI, our main explanatory variable, in it. We conduct this test by obtaining two data sets for our MLR II model, one for high income countries and one for all other countries.

Doing so will give us two models: one for high income countries and one for low income countries. Respectively, these models are:

$Poverty = \mathrm{A}GINI + \mathrm{B}LiteracyRate + \mathrm{C}log(GDPperCapita) + \mathrm{D}FreedomScore + u$

$Poverty = \mathrm{A}_1GINI + \mathrm{B}_1LiteracyRate + \mathrm{C}_1log(GDPperCapita) + \mathrm{D}_1FreedomScore + u$

Where: A, A_1, B, B_1, C, C_1, D, D_1 are the correlation coefficients for their respective models.

Chow Test								
	MLR II when HighIncome = 1				MLR II when HighIncome = 0			
Variable	Coef.	Std. err.	t-value	P>|t|	Coef.	Std. err.	t-value	P>|t|
GINI	0.063144	0.017419	3.62	0.001	0.69985	0.21656	3.23	0.002
LiteracyRate	0.042259	0.051952	0.81	0.968	0.00615	0.15336	0.04	0.968
log(GDPperCapita)	-0.00123	0.172398	-0.01	0.994	-14.5187	2.62823	-5.52	0.000
FreedomScore	-0.02980	0.01708	-1.75	0.092	0.149317	0.067318	0.22	0.826

To conduct the Chow Test, we use the following null hypothesis, alternative hypothesis, and formula:

$$H_0 = A = A_1, B = B_1, C = C_1, D = D_1$$

$$H_a = H_o \text{ is not true}$$

$$F = \frac{[SSRp-(SSR1+SSR2)]/(k+1)}{[SSR1+SSR2]/n-2(k+1)}$$

Substituting our data into the equation gives us:

$$F = \frac{[6051.22-(5.75+3922.68)]/5}{[5.75+3922.68]/80-2(5)} = 7.57$$

Our F-value is significantly larger than the critical value, 2.50. This significance means that we can reject our null hypothesis that data from our pooled model, MLR II, does not have significantly different true coefficients when split into two separate models which differentiate between high income countries and all other countries. The differences in true coefficients contribute to the story that data paints for development economists. As seen in our table, the GINI coefficient is statistically significant for countries of all income levels. Income inequality is an important structural issue to tackle no matter a country's income level. However, a deeper look into the data reveals where policy paths should diverge depending on a given country's income level. The coefficient for log(GDPperCapita) for high income countries is shown by the table as not significant at any confidence level under ten percent. However, this coefficient is significant at the one percent level for non-high income countries and is strongly negative. This means that for countries high income, percent changes in GDP per capita strongly affect Poverty Headcount rate. Furthermore, FreedomScore is significant at a ten percent level when looking at data for high income countries; however, it would be significant at an 82.6% confidence level when it comes to non-high income countries' poverty regression, making it virtually irrelevant. Additionally, the coefficient on FreedomScore for high income countries, -0.092, is negative. This means that for a one point increase in freedom score in high income countries, Poverty Headcount rate decreases by 0.092 percentage points. This negative coefficient is in line with our hypothesis that FreedomScore would be negatively related to Poverty.

It is clear from these data dichotomies that policies aiming for poverty alleviation will have radically different degrees of effectiveness depending on the income level of the recipient country. Specifically, poverty elimination for high income countries should be focused on factors like political and social rights that affect Freedom House score.. On the other hand, this data tells us that poverty alleviation in non-high income countries should primarily be concerned with growth in GDP per capita. This means that governments of countries which are not high income should be first be focused on increasing GDP per capita if they want to alleviate poverty. Both sets of countries have GINI as a variable that is significant which tells us that all governments should focus on alleviating income inequality when they are trying to combat poverty.

VI: Conclusions

The overall findings from our models were slightly different than our original hypotheses. While the variables representing literacy rate, change in GDP per capita, and Freedom House score were significant in our multiple regression models, the variable representing infant mortality rate was not. Although the variable GINI was not significant at our desired confidence level in the multiple linear regression, GINI was significant in our simple regression as well as in the two Chow tests we performed. Our findings indicate that, as hypothesized, income inequality is still an important factor affecting poverty levels. As policymakers grapple with complex proposals for new income taxes, these findings on the relationship between inequality and poverty potentially support a larger tax on the ultra-wealthy. In summary of our findings, focusing on improving literacy rates, growing GDP per Capita, and reducing inequality are important focuses for any countries who want to combat poverty. For high income countries, an additional focus on improving civil liberties and political rights is also critical.

There are a vast amount of complex factors that can affect poverty levels within a country; however, our model focuses on many variables related to poverty that can provide potential policy outlets that governments could undertake. Determining the significance of income inequality, literacy rate, infant mortality rate, and Freedom House score can help governments and policymakers determine the most effective and strategic plans to combat poverty and improve their country's global standing.

In the future, we would like to test the relationship between poverty and other indicators of healthcare system quality. While infant mortality rate was not individually significant, other health-related variables may be. Moving forward, these additional findings could be utilized as policymakers grapple with which healthcare systems are best in terms of how they relate to poverty and how to allocate government spending.

Appendix

Countries Used in Research:

Argentina, Armenia, Austria, Bangladesh, Belarus, Belgium, Benin, Bhutan, Bolivia, Brazil, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, France, Gabon, Gambia, Germany, Georgia, Ghana, Greece, Honduras, Hungary, Indonesia, Iran, Ireland, Israel, Italy, Kazakhstan, Kenya, Kosovo, Latvia, Liberia, Lithuania, Luxembourg, Malawi, Malaysia, Malta, Mexico, Moldova, Mongolia, Myanmar, Netherlands, Norway, Namibia, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russia, Rwanda, Serbia, Slovenia, Spain, Sweden, Switzerland, Thailand, Tunisia, Turkey, Uganda, United Kingdom, United States, Uruguay, Vietnam

Stata Output:

SLR I

```
regress Poverty GINI

      Source |       SS           df       MS      Number of obs   =        80
-------------+----------------------------------   F(1, 78)        =      9.54
       Model |  1588.74069         1  1588.74069   Prob > F        =    0.0028
    Residual |  12988.1642        78  166.514925   R-squared       =    0.1090
-------------+----------------------------------   Adj R-squared   =    0.0976
       Total |  14576.9049        79  184.517783   Root MSE        =    12.904

------------------------------------------------------------------------------
     Poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        GINI |    .620799   .2009792     3.09   0.003     .2206802    1.020918
       _cons |   -16.7198   7.615093    -2.20   0.031    -31.88028   -1.559314
------------------------------------------------------------------------------
```

MLR I

```
. regress Poverty GINI LiteracyRate logGDPperCapita FreedomScore InfantMortality
```

Source	SS	df	MS
Model	8543.97506	5	1708.79501
Residual	6032.92981	74	81.5260786
Total	14576.9049	79	184.517783

Number of obs	=	80
F(5, 74)	=	20.96
Prob > F	=	0.0000
R-squared	=	0.5861
Adj R-squared	=	0.5582
Root MSE	=	9.0292

Poverty	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
GINI	.2099268	.1599463	1.31	0.193	-.1087731	.5286268
LiteracyRate	-.4718798	.168305	-2.80	0.006	-.8072349	-.1365247
logGDPperCapita	-5.318058	1.604515	-3.31	0.001	-8.515124	-2.120991
FreedomScore	.1381265	.0530972	2.60	0.011	.0323281	.2439249
InfantMortality	-.0864877	.1826027	-0.47	0.637	-.4503316	.2773563
_cons	81.45898	20.58873	3.96	0.000	40.43504	122.4829

MLR II

```
. regress Poverty GINI LiteracyRate logGDPperCapita FreedomScore
```

Source	SS	df	MS
Model	8525.68605	4	2131.42151
Residual	6051.21883	75	80.6829177
Total	14576.9049	79	184.517783

Number of obs	=	80
F(4, 75)	=	26.42
Prob > F	=	0.0000
R-squared	=	0.5849
Adj R-squared	=	0.5627
Root MSE	=	8.9824

Poverty	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
GINI	.1895238	.1532377	1.24	0.220	-.1157413	.494789
LiteracyRate	-.4133549	.113681	-3.64	0.001	-.6398191	-.1868908
logGDPperCapita	-5.054608	1.497217	-3.38	0.001	-8.037217	-2.072
FreedomScore	.1368909	.0527581	2.59	0.011	.0317914	.2419904
_cons	73.40045	11.53368	6.36	0.000	50.42417	96.37672

Correlation Matrix

```
. correlate GINI LiteracyRate logGDPperCapita FreedomScore InfantMortality HighIncome
(obs=80)

             |     GINI Litera~e logGDP~a Freedo~e Infant~y HighIn~e
-------------+------------------------------------------------------
        GINI |   1.0000
LiteracyRate |  -0.2082   1.0000
logGDPperC~a |  -0.3671   0.7505   1.0000
FreedomScore |  -0.1561   0.4463   0.6845   1.0000
InfantMort~y |   0.3592  -0.8862  -0.8081  -0.4832   1.0000
  HighIncome |  -0.4437   0.4631   0.8079   0.7392  -0.5997   1.0000
```

SLR II

```
. regress Poverty HighIncome

      Source |       SS       df       MS              Number of obs =      80
-------------+------------------------------           F(  1,    78) =   11.35
       Model |  1851.42352     1  1851.42352           Prob > F      =  0.0012
    Residual |  12725.4814    78  163.147197           R-squared     =  0.1270
-------------+------------------------------           Adj R-squared =  0.1158
       Total |  14576.9049    79  184.517783           Root MSE      =  12.773

------------------------------------------------------------------------------
     Poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  HighIncome |  -9.819792   2.915004    -3.37   0.001    -15.62312   -4.016464
       _cons |   10.30417   1.843611     5.59   0.000      6.63382    13.97451
------------------------------------------------------------------------------
```

F-Test

Unrestricted

```
. regress Poverty GINI LiteracyRate logGDPperCapita FreedomScore InfantMortality

      Source |       SS       df       MS              Number of obs =      80
-------------+------------------------------           F(  5,    74) =   20.96
       Model |  8543.97506     5  1708.79501           Prob > F      =  0.0000
    Residual |  6032.92981    74  81.5260786           R-squared     =  0.5861
-------------+------------------------------           Adj R-squared =  0.5582
       Total |  14576.9049    79  184.517783           Root MSE      =  9.0292

------------------------------------------------------------------------------
        Poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
           GINI |   .2099268   .1599463     1.31   0.193    -.1087731    .5286268
   LiteracyRate |  -.4718798    .168305    -2.80   0.006    -.8072349   -.1365247
logGDPperCapita |  -5.318058   1.604515    -3.31   0.001    -8.515124   -2.120991
   FreedomScore |   .1381265   .0530972     2.60   0.011     .0323281    .2439249
InfantMortality |  -.0864877   .1826027    -0.47   0.637    -.4503316    .2773563
          _cons |   81.45898   20.58873     3.96   0.000     40.43504    122.4829
------------------------------------------------------------------------------
```

Restricted

```
. regress Poverty GINI logGDPperCapita InfantMortality

      Source |       SS           df       MS      Number of obs   =        80
-------------+----------------------------------   F(3, 76)        =     25.14
       Model |  7260.61495         3  2420.20498   Prob > F        =    0.0000
    Residual |  7316.28992        76  96.2669727   R-squared       =    0.4981
-------------+----------------------------------   Adj R-squared   =    0.4783
       Total |  14576.9049        79  184.517783   Root MSE        =    9.8116

------------------------------------------------------------------------------
        Poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
           GINI |    .136043    .165366     0.82   0.413    -.193312     .465398
logGDPperCapita |  -3.553341   1.398941    -2.54   0.013   -6.339573   -.7671098
InfantMortality |   .3413878   .1332119     2.56   0.012    .0760733    .6067023
          _cons |   28.87026   16.14872     1.79   0.078   -3.292691    61.03322
------------------------------------------------------------------------------
```

```
. test LiteracyRate InfantMortality

 ( 1)  LiteracyRate = 0
 ( 2)  InfantMortality = 0

       F(  2,    74) =    6.65
            Prob > F =    0.0022
```

Chow Test

```
. regress Poverty GINI LiteracyRate logGDPperCapita FreedomScore if HighIncome == 1

      Source |       SS           df       MS      Number of obs   =        32
-------------+----------------------------------   F(4, 27)        =      6.60
       Model |  5.62801423         4  1.40700356   Prob > F        =    0.0008
    Residual |  5.75417327        27  .213117529   R-squared       =    0.4945
-------------+----------------------------------   Adj R-squared   =    0.4196
       Total |  11.3821875        31  .367167339   Root MSE        =    .46165

------------------------------------------------------------------------------
        Poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
           GINI |    .063144   .0174195     3.62   0.001    .0274021    .0988859
   LiteracyRate |   .0422586   .0519526     0.81   0.423   -.0643393    .1488566
logGDPperCapita |  -.0012361   .1723986    -0.01   0.994   -.3549689    .3524967
   FreedomScore |  -.0298041   .0170767    -1.75   0.092   -.0648426    .0052343
          _cons |  -3.007329   5.644846    -0.53   0.599    -14.5896    8.574939
------------------------------------------------------------------------------
```

```
. regress Poverty GINI LiteracyRate logGDPperCapita FreedomScore if HighIncome == 0
```

Source	SS	df	MS			
Model	8791.42191	4	2197.85548	Number of obs	=	48
Residual	3922.67725	43	91.2250524	F(4, 43)	=	24.09
				Prob > F	=	0.0000
				R-squared	=	0.6915
				Adj R-squared	=	0.6628
Total	12714.0992	47	270.512748	Root MSE	=	9.5512

Poverty	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
GINI	.6998515	.2165593	3.23	0.002	.263118	1.136585
LiteracyRate	.0061532	.1533611	0.04	0.968	-.3031289	.3154354
logGDPperCapita	-14.51869	2.62823	-5.52	0.000	-19.81902	-9.218358
FreedomScore	.0149317	.0673177	0.22	0.826	-.1208273	.1506907
_cons	100.3767	14.80225	6.78	0.000	70.52511	130.2283

References

"Adult Literacy Rate." *UNESCO UIS*, UNESCO, 12 Sept. 2019, http://uis.unesco.org/en/glossary-term/adult-literacy-rate.

Ahmad, Nisar, et al. "The Long Run and Short Run Relationship between Poverty and Literacy Rate in Pakistan." *Pakistan Business Review*, vol. 20, no. 4, Jan. 2019, pp. 875–885. *EBSCOhost*, doi:https://pbr.iobm.edu.pk/archives/.

Arndt, Channing, Andy McKay, and Finn Tarp. *Growth and poverty in sub-Saharan Africa*. Oxford University Press, 2016.

Chambers, Dustin, and Shatakshee Dhongde. "A Non-Parametric Measure Of Poverty Elasticity." *Review of Income and Wealth*, vol. 57, no. 4, 2011, pp. 683–703., doi:10.1111/j.1475-4991.2011.00446.x.

Fosu, Augustin Kwasi. "Inequality, Income, and Poverty: Comparative Global Evidence." *Social Science Quarterly*, vol. 91, no. 5, 2010, pp. 1432–1446. *JSTOR*, www.jstor.org/stable/42956468.

"Global Billionaire Population: The Billionaire Census 2018." *Wealth-X*, Wealth-X, 15 May 2018, https://www.wealthx.com/report/the-wealth-x-billionaire-census-2018/?utm_campaign=billionaire-census-2018&utm_source=press&utm_medium=referral&utm_term=bc-2018-press.

"Health and Development." *World Health Organization*, World Health Organization, 9 Dec. 2010, https://www.who.int/hdp/en/.

OECD (2015), In It Together: Why Less Inequality Benefits All, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264235120-en