Assessing environmental risk factors of noncommunicable diseases

Ambient and household air pollution

Monica Buczynski*

May 2020

Abstract

Over the past decade, the mortality rates of noncommunicable diseases have continued to rise, especially in lower developed countries. Furthermore, the purpose of this paper is to estimate the average marginal effects of ambient and household air pollution, measured by $PM_{2.5}$ concentration levels and the proportion of efficient cooking technologies, respectively, on the ratio of premature deaths from noncommunicable diseases (NCDs) to all deaths from NCDs. Using linear regression, I find a statistically significant relationship for both indicators of air pollution. Additionally, I validate the presence of a Health "Kuznets" curve, an inverse "U" shape relationship between the Human Development Index (HDI) and health inequalities.

Keywords: air pollution; development; disease; environment; NCD; noncommunicable; premature

1 Introduction

Bennett et al. [2018] stated in their report, *NCD Countdown 2030: worldwide trends in noncommunicable disease mortality and progress towards Sustainable Development Goal target 3.4*, that in 2016, an estimated 16.9 million people had died prematurely[1] worldwide from NCDs. Similarly, commenting on the human losses due to NCDs, but focusing on the global environmental impacts, the World Health Organization (WHO), a partner with the United Nations (UN) on the 2030 Sustainable Development Agenda, concluded in their 2016 report that an estimated 65% of the 12.6 million global deaths caused by the environment occurred from NCDs ("Preventing," 2017). The global 2016 NCD deaths attributed to the environment account for stroke (2.5 million), ischemic heart disease (2.3 million), cancers (1.7 million), chronic respiratory diseases (1.4 million) and others. More specifically, in 2016, 25% of stroke related deaths were caused by ambient air pollution and 26% of stoke-related deaths were attributable to household air pollution. Unfortunately, over the past decade, the mortality rates of NCDs have continued to rise.

It is critical to obtain accurate marginal estimates of environmental causes of death, like the impacts of ambient and indoor pollution on NCDs, so that leaders can make better informed policy decisions such as implementing more restrictive environmental policies. In total, the aforementioned 12.6 million global deaths attributed to the environment accounted for 23% of total global deaths in 2016. Put another way, an estimated 23% of all deaths worldwide could have been prevented by healthier environments. Moreover, these policies may open the medical community to more opportunities for funded research from governments and private agencies geared towards better quality health diagnostic tools, prevention mechanisms, and treatment.

Likewise, due to health inequalities[2], it is important to implement effective strategies to both prevent and combat NCDs as it is likely that diagnosis and treatment costs will become less affordable in the future, leaving developing countries most vulnerable to NCDs; more than 85% of premature deaths are concentrated in developing countries (Soheylizard et al., 2018; "U.S." 2019). In turn, global communities, country leaders

*Author contact information: Email: mbuczynski26@gmail.com | GitHub: https://github.com/monica-buczynski

[1] "Premature" deaths are denoted as deaths that have occurred under the age of 70 (Bennett et al., 2018).

[2] Health inequalities is a broad term that encompasses a variety of inequalities including those in the quality and access to care and treatments or the health statuses between socio-economic factors, genders, or geographies (Williams et al., 2020).

and private citizens worldwide should attempt to reduce risk factors of NCDs globally, regardless of their country's population health and economic standing, as it has been shown that the presence of NCDs in one country can negatively impact the world economy through disrupting international supply chains, depleting labor and capital supply, and decreasing the productivity of businesses (Abegunde and Stanciole, 2006). However, despite recent efforts to study the effects of the environment on NCDs, research is limited, and current estimates may be underestimated due to challenges such as determining lag times of exposure to pollutants, accuracy of primary records (especially in low-income countries) and assessing exposure to multiple toxins (Bennett et al., 2018).

This research paper aims to estimate the marginal effects of air pollution on premature NCDs because health determinants may impact conclusions for policies and health care services. Specifically, I measure the effects of ambient and household air pollution on the ratio of premature NCD deaths to all NCD deaths. Additionally, I will examine the validity of a Health "Kuznets" curve, the association between health inequalities and economic development.

The remainder of this paper is organized as follows: Section 2 will explore relevant studies concerning the impacts of air pollution on population health. Section 3 will layout the model used in my study, and Section 4 will include descriptions of data and its limitations. Section 5 will discuss the results. Section 6 will explore suggestions for future research. Section 7 will acknowledge all references mentioned in the paper. Finally, Section 8 will contain figures, maps and tables that are not provided in the main body of the paper.

2 Literature Review

$PM_{2.5}$ are fine particulate matter with a diameter less than 2.5 micrometers like combustion particles that can penetrate deep into the lungs and cause severe health problems (Meng, 2019). $PM_{2.5}$ particles are "short-lived" and have an average lifetime of 1 to 2 weeks; therefore, reducing $PM_{2.5}$ emissions could result in an immediate decline in air pollution problems. Zheutlin et al. [2014] find a statistically significant relationship (p-value $< .0001$) between $PM_{2.5}$ and diabetes prevalence while controlling for socioeconomic indicators like poverty rate and education level. For every additional 10 microgram/m^3 of $PM_{2.5}$, it is estimated that diabetes prevalence increases by 0.90%, though this may be a conservative result as the sample consisted of adults who are less susceptible to the exposure of pollutants than children as mentioned by Cosbey et al. [2005]. However, interestingly, Zheutlin et al. [2014] fail to find a statistically significant relationship (p-value $= .26$) between $PM_{2.5}$ and obesity prevalence, a risk factor for type-II diabetes.

While the WHO estimates that 3.7 million people die yearly from ambient air pollution-related causes, 4.3 million people die yearly due to household air pollution-related causes (Apte and Salvi, 2017). According to Apte and Salvi [2017], household air pollution, in general, is more profound in developing countries due to ill-ventilated housing structures and cooking with inefficient fuels and stoves: more efficient fuels produce less pollution, but are more expensive, leaving lower income and less developed countries more vulnerable to adverse health outcomes like chronic obstructive pulmonary disease. See Figure 1 and Figure 2 in Section 8 for maps of the distribution of the proportion of population with primary reliance on clean fuels and technologies across all countries and the distribution of the Human Development Index across all countries, respectively. When analyzing Figure 1 and Figure 2, it is visually apparent that countries with lower proportions of the population with primary reliance on clean fuels and technologies are generally less developed areas like African and Middle-Eastern countries, consistent with the study of Apte and Salvi [2007].

Coincidentally, developing countries, as mentioned in Soheylizard et al. [2018], typically have less access to medical facilities and have poorer quality diagnostic tools, leading to lower reported prevalence in diseases, but higher mortality rates. Specifically regarding premature NCD mortality, Figure 3 in Section 8 shows a scatter plot between the Human Development Index and the ratio of premature NCD mortality to all NCD mortality. There is a strong negative linear correlation with a Pearson correlation coefficient of -0.811. Figure 3 is consistent with the conclusion of Soheylizard et al. [2018] that countries with higher Human Development Index values are typically associated with experiencing lower mortality rates.

Exploring deeper the relationship between country development and health outcomes, Molini et al. [2010] explains that the HDI can explain an inverted "U" shape of health inequalities based on the traditional

Kuznets curve[3]. Molini et al. [2010] use Body Mass Index (BMI) rates as a proxy for health inequalities and adopt the following explanation for the inverse "U" relationship between health inequalities and economic development: at lower HDI levels, accessibility of medical products and services and high quality foods may be scarcer, increasing prices which, in turn, increases health inequalities. However, health inequalities may decrease when medical supplies become more affordable and accessible as country development increases.

Moreover, assessing the relationship between health inequalities and economic development measured by the HDI is appropriate when studying the effects of air pollution and health outcomes. Chronic stress, due to higher probability of crime rates and violence from health inequalities, may increase the risk for poorer health outcomes, like ischemic heart disease, an NCD that accounted for more than 9 million global deaths in 2016 (Woodward and Kawachi, 2000; Nurias et al., 2016; Nowbar et al., 2019). However, concerning the link, specifically, between environment related NCD deaths and chronic stress, stress has been shown to lower the immune system's ability to fight against air pollution, as described by Apte and Salvi [2017].

One way that countries are globally contributing to the fight against NCDs is through the United Nation's Sustainable Development Goals (SDGs). Specifically, SDG target 3.4 calls for a "one-third reduction, relative to 2015 levels, of the probability of dying between 30 years and 70 years of age" from NCDs by 2030 (Bennett et al., 2018, p. 1072). However, as of 2018, three years after the adoption of the SDGs, less than 50% of countries were on track to reach the target.

When analyzing the study of Bennett et al. [2018], the authors only briefly cover ambient air pollution, but do not discuss the implications on NCDs beyond behavioral factors[4] such as greenhouse gases or $PM_{2.5}$; however, SDG target goal 3.9 does incorporate the reduction of deaths and illness caused by air pollution, but does not focus on NCDs. Even so, Bennett et al.[2018] excludes indoor air pollution via the fuels used in cooking and suggests policy implementations such as higher taxes on tobacco as a mechanism to decrease NCDs, neglecting the studies of Zheutlin et al. [2014], Soheylizard [2018] and others who have contributed to the studies that human health is negatively impacted by air pollution. Furthermore, I contribute to the literature by estimating the marginal effects of environmental determinants on premature NCDs and validate the Health "Kuznets" curve.

3 Model

I use a linear regression to estimate the marginal effects of ambient and household air pollution on health outcomes across countries worldwide. Because SDG target 3.4 is focused on tracking the NCD-related deaths (Bennett et al., 2018), the response variable in my model is the ratio of premature NCD deaths to all NCD deaths (NCD_P). The explanatory variables of interest are $PM_{2.5}$ concentration levels ($PM_{2.5}$) and the proportion of population with primary reliance on clean fuels and technologies as the primary source of domestic energy for cooking ($eTech$). Higher concentrations of $PM_{2.5}$ levels have been associated with poorer health while the usage of more efficient cooking technologies has been associated with less adverse health outcomes (Meng, 2019; Apte and Salvi, 2017). To hold constant countries' domestic financial power of combating NCDs, as Bokhari et al. [2007] suggests that increasing domestic medical research may increase population health, I control for the domestic general government health expenditures per capita ($DomHE$). Likewise, access to high-quality treatments and diagnostics may be dependent on the level of a country's development (Soheylizard et al., 2018). To account for this, I control for the Human Development Index (HDI).

My model is represented by the equation

$$log(NCD_P) = \beta_0 + \beta_1 log(PM_{2.5}) + \beta_2 log(eTech) + \beta_3 log(DomHE) + HDI + HDI^2 + u$$

[3]The traditional Kuznets curve is a hypothesis that is graphically expressed as an inverted "U" and explains that initial increases in a country's economic development (measured by GDP per capita) are associated with increases in a country's economic inequalities (Molini et al., 2010). Then, after some turning point, GDP per capita is associated with decreases in economic inequalities.

[4]Behavioral factors include activities such as tobacco smoke (Bennett et al., 2018).

where u is the error term with the usual assumptions that the error conditionally follows a normal distribution with 0 mean and constant variance. Concerning the functional form of the variables, NCD_P, $PM_{2.5}$, $eTech$, and $DomHE$ are given a logarithmic transformation so that their coefficients can be interpreted as elasticities. Conversely, HDI is not given a logarithmic transformation to replicate the methodology of Molini et al. [2010].

While the primary goal of this model is to obtain the average marginal effects of air pollution, measured by $PM_{2.5}$ and $eTech$, a secondary goal is to evaluate how health inequalities are dependent on a country's level of human development. To that end, I incorporate a quadratic term for the variable HDI to assess the validity of a Health "Kuznets" curve between health inequalities (measured by NCD_P) and economic development (measured by HDI). Then, if the Health "Kuznets" curve theory holds as theorized by Molini et al. [2010], I can further investigate at what value of a country's economic development health inequalities begin to decrease.

Moreover, NCD_P is a better proxy for health inequalities than the BMI indicator used by Molini et al. [2010] because "beyond a certain threshold" of the BMI measure, being overweight or underweight is considered unhealthy, whereas a "poverty and inequality analysis requires that a welfare indicator provides a monotonic ranking of individuals or households" (p. 1013). Conversely, every increase in NCD_P can be reasonably associated with poorer health outcomes for a population because of the economic losses related with premature death (Abegunde and Stanciole, 2006).

Data

Table 1 shows the five-number summaries and standard deviations of all variables used in my study.

	NCD_P	$PM_{2.5}$	eTech	DomHE	HDI
Minimum	16.00	5.89	5.00	4.23	0.37
1st Quartile	32.00	13.87	28.25	69.05	0.59
Median	47.00	21.43	85.00	358.99	0.73
Mean	45.28	26.93	64.04	974.90	0.71
3rd Quartile	57.75	33.34	95.00	1228.23	0.83
Maximum	76.00	98.06	95.00	8077.93	0.95
Standard Deviation	15.27	18.75	35.82	1372.52	0.15

Table 1: Five-Number Summary and Standard Deviation Values

Notably, $eTech$ has a negative skew with a mean and median of 64.04 and 85.00, respectively, which may be due to rounding techniques used by the WHO. The variables will each be discussed in greater detail below.

Premature deaths due to NCDs as a proportion of all NCD deaths (%)

The NCD Countdown 2030 study uses 2016 data of the mean probability of dying between 30 and 70 years of age for the SDG target 3.4. Because no data was found that would indicate the numeric value of premature deaths for easier interpretation, I use 2016 mean estimates of the premature NCD ratio to account for all premature deaths (Premature, 2016) from the WHO data bank. The indicator is measured as a percentage.

My study includes data for 170 countries and territories; 13 countries were excluded from the original dataset due to lack of data from the explanatory variables. The indicator is limited to the age of 80 as Bennett et al. [2018] notes that after age 80, the error in NCD mortality rates increase because it becomes more difficult to determine the cause of death; as a person ages, it is more likely that a person has multiple existing conditions leading to misidentification of the cause of death in death records. Standardized definitions and methods were used to ensure comparability among countries. Figure 4 in Section 8 is a histogram of NCD_P, which shows that there is no obvious skew and that NCD_P looks reasonably symmetric.

Measurement and monitoring gaps exist among high, middle, and low-income countries, especially in Africa where "there is strong reliance on verbal autopsy studies, most of which are not nationally representative samples" (Mathers et al., 2017). As a result, the WHO adjusts the estimates by assessing the "completeness of the death registration, the quality of the cause-of-death information, and the timeliness of publication" (Bennett et al., 2018, p. 1082). However, incomplete coverage could still impact results. For my analysis, only 49 of 170 (28.8%) countries and territories contain "high-quality" data while 68 (40.0%) countries and territories were characterized as "very low."

Table X provides the breakdown of the "Vital registration data quality" for the indicator of premature deaths due to NCDs as a proportion of all NCDs deaths to measure health outcomes. Row 3 corresponds with the percentage of the total number of countries in the respective data quality group to the total number of countries considered for assessment for "vital registration data quality" standards by the WHO.

	Very Low	Low	Medium	High
Country Count	68	25	28	49
Percentage %	40.0	14.7	16.5	28.8

Table 2: "Vital registration data quality"

However, while my primary goal in this study is to investigate the marginal effects of air pollution on the ratio of premature NCD deaths to all NCD to account for premature deaths, no data was available or no calculations (based on the available WHO health data) could be made without questioning the parameters of the dataset that would produce a value that would indicate the numeric value of premature deaths in the units of people in my study. Moreover, easier interpretation of the practical significance of my results could be made. For example, I could have used my health indicator, "Premature deaths due to noncommunicable diseases (NCD) as a proportion of all NCD deaths (%)", and another health indicator from the WHO, "Total NCD deaths (in thousands)" to calculate the number of premature deaths in units of people. However "Total NCD deaths (in thousands)" does not explicitly describe whether or not the indicator measures deaths from all NCDs or from only NCD4; only after intensive research on other WHO documents, I discovered that my health indicator, "Premature deaths due to noncommunicable diseases (NCD) as a proportion of all NCD deaths (%)", accounted for the premature *NCD4* deaths while the types of NCDs that is accounted for by the WHO indicator, "Total NCD deaths (in thousands)" still remains unknown based on my investigation (Beaglehole et al., 2014, *Total NCD*, 2016). Furthermore, I use the ratio of "Premature deaths due to noncommunicable diseases (NCD) as a proportion of all NCD deaths (%)" to account specifically for premature deaths since the *NCD Countdown 2030* campaign accounts for premature deaths, ages 30 to 70.

Additionally, while the marginal changes in the premature NCD death *ratio* may indicate changes to the "premature" deaths or the "non-premature" deaths (or both), it may be difficult to accurately identify if $PM_{2.5}$ and $eTech$ has a greater association with the number "premature" deaths or the "non-premature" deaths. However, using the Ratio of premature deaths due to NCDs as a proportion of all NCD deaths as the response variable is still useful because the information obtained from the ratio roughly indicates the proportion of "the most economically productive age span," people ages 30 to 70, and the future members of "the most economically productive age span" group as defined by the *NCD Countdown 2030* campaign (Bennett et al., 2018).

$PM_{2.5}$ concentration levels

I use the 2016 estimates of the $PM_{2.5}$ mean annual concentration exposure levels, measured by micrograms/m^3, from the World Bank Group (WBG) to represent ambient air pollution ($PM_{2.5}$ air, 2016). The Global Burden of Disease study calculated the estimates by using combined data from atmospheric models, satellite observations, and ground-level monitoring. Exposure levels are derived from both urban and rural areas and then aggregated. The WHO also provides data for $PM_{2.5}$ concentration levels, and, on average, these values were lower than the data obtained from the WBG. Data from the WBG was used in my study due to more availability of data across years that could be used for further studies; the World Bank Group provides data of $PM_{2.5}$ concentration levels from the years 1900, 1996, 2000, 2010-2017 while the World Health Organization provides data for only the year 2016. Despite the WBG's range of data across 11 years, I only use 2016 data for $PM_{2.5}$ since the WHO only uses 2016 data for their SDG target 3.4 health indicator for the *NCD Countdown 2030* campaign.

Proportion of population with primary reliance on clean fuels and technologies

I used 2016 annual estimates of the proportion of population with primary reliance on clean fuels and technologies as the primary source of domestic energy for cooking from the WHO to measure the ability of each respective country to combat household air pollution (Percentage, 2016). The indicator is measured as a percentage; the number of people who use efficient cooking technologies is divided by the total country population. Several census and surveys on national and regional levels are used and thus, inconsistent definitions among countries of "clean fuels and technologies" may be used, distorting full comparability.

Domestic general government health expenditures per capita

I also control for the 2016 domestic general government health expenditures per capita (Domestic, 2016) from the World Bank Group. The data is measured in international dollars at purchasing power parity (PPP) to increase compatibility for cross-country analysis. Regarding national funding toward health care efforts, external (international aid) are not included in this study, so I could hold constant the financial sustainability of a country to combat NCDs on a national level; however, data for external healthcare financial can be acquired at The Word Bank database.

Human Development Index

I control for the 2016 HDI levels calculated by the United Nations Development Programme (UNDP) (United, 2016). The HDI is measured on a scale with values ranging from 0 to 1 where higher values represent a higher levels of country development and consists of the following indicators: Life expectancy at birth, expected years of schooling, mean years of schooling and GNI per capita (PPP) ("Human", n.d.). Furthermore, the Human Development Index is a stronger measure of the "development of a country" than per-capita income because it measures non-income attributes and not only economic growth (Human Development Index (HDI), n.d.; Davies and Quinlivan, 2006). Research suggests that all three dimensions impact one's risk of dying from a NCDs (Soheylizard et al. 2018). However, the HDI does not account for measures such as the quality of products or income inequality.

Results

Summary Statistics

Table 2 shows the summary statistics of the estimated linear regression model.

	Estimate	Std. Error	Pr(>|t|)	95% CI [LB,UB]
(Intercept)	2.5952	0.3397	$1.727e^{-12}$*	1.9246, 3.2658
log($PM_{2.5}$)	0.1362	0.0459	0.003479*	0.0455, 0.2269
log(eTech)	-0.0781	0.0279	0.005751*	-0.1333, -0.0230
log(DomHE)	0.0428	0.0305	0.1634	-0.0175, 0.1030
HDI	4.4156	0.7897	$9.232e^{-08}$*	2.8567, 5.9745
HDI^2	-4.4588	0.5763	$9.795e^{-13}$*	-5.5963, -3.3212

Table 3: Summary Statistics

All statistical significance is based at the .05 level. Robust standard errors were used for inference due to evidence of heteroskedasticity[5],however, Table X shows the summary statistics when using the uncorrected standard errors. Based on my slope coefficient estimates, shown in column 1 of Table X, all the covariates show a positive relationship with NCD_P except for $eTech$. The positive and negative coefficients on HDI and HDI^2, respectively, indicate an inverted "U" relationship with health inequalities. To validate the Health "Kuznets" theory, I perform an F test for exclusion restrictions and conclude that the variables HDI and HDI^2 are jointly statistically significant (p-value = $1.394 * 10^{-13}$).

[5]Figure 5 in Section 8 shows a residuals vs. fitted values plot for the estimated model. From this plot, there is some visual evidence indicating heteroskedasticity, which is further verified when conducting a Breusch-Pagan test (p-value = 0.0005818).

While the normal Q-Q plot shown in Figure 6 in Section 8 provides some evidence that the residuals diverge from normality, particularly in the tails, the sample size is large enough to justify valid statistical inference.

The variables $PM_{2.5}$ (p-value = .003479) and $eTech$ (p-value = .005751) were found to be statistically significant, while $DomHE$ (p-value = .1634) was not.

Despite high multicollinearity which increases the difficulty of achieving statistical significance, the variables $PM_{2.5}$ and eTech obtained low p-values of .003479 and .005751, respectively, indicating statistically significant effects.

Table 3 shows the variance inflation factors (VIF) for the model.

log($\mathrm{PM}_{2.5}$)	log(eTech)	log(DomHE)	HDI	HDI^2
1.665224	5.850765	11.859549	144.527084	128.516761

Table 4: Variance Inflation Factor

The variance would be approximately 1.66, 5.85, 11.86, 144.53 and 128.52 times as large as it would be if there was no multicollinearity, respectively. It is expected that HDI and HDI^2 have large VIF values of 144.53 and 128.52, respectively, since HDI^2 is a function of HDI.

Log (DomHE) also has a relatively large VIF value of 11.86, which may explain its lack of statistical significance and is likely due to its close relationship with HDI and HDI^2 (Soheylizard et al., 2018). I also conduct an overall F test to show that the covariates are collectively statistically significant (p-value $< 2.2 * 10^{-16}$).

Altogether, my results largely corroborate the existing literature and my initial hypotheses. Consistent with Meng [2019] and Zheutlin et al. [2014] who find in their studies that an increase in the concentration of $PM_{2.5}$ is correlated with adverse health outcomes, I find that a 1% increase in $PM_{2.5}$ is estimated to be associated with a .1362% (95% CI [0.0455, 0.2269]) increase in NCD_P, ceteris paribus.

In addition, I find that a 1% increase in $eTech$ is estimated to be associated with a .0781% (95% CI [-0.1333, -0.0230]) decrease in NCD_P, ceteris paribus. This is consistent with Apte and Salvi [2017] who suggest that households with efficient cooking stoves and use clean fuels for cooking are less susceptible to deaths attributed to household air pollution.

Inconsistent with Bokhari et al. [2007] who finds that an increase in domestic health expenditures is correlated with positive health outcomes, I find that a 1% increase in $DomHE$ is estimated to be associated with a .0428% increase in NCD_P, ceteris paribus, which may be attributed to corrupt healthcare systems (Kiross et al., 2020). The sign is not as expected but since the effect is fairly small and not significant (p-value = .1634), it should not be over-interpreted as $DomHE$ was not the focus of my research.

Health "Kuznets" Curve

The inverse "U" relationship is shown by Figure 7 in Section 8. I find that HDI has an increasing marginal effect on health inequalities (as measured by NCD_P) until HDI reaches .495[6]. When a country reaches an HDI level of .495, health inequalities between lower and higher income groups are expected to stop increasing and will begin to decrease. Even though my model shows that a Health "Kuznets" curve exists when using NCD_P and HDI as indicators, only 17 out of 170 countries in my dataset are characterized by a 2016 HDI score lower than .495. However, based on past research, these impacts may still be globally significant (Abegunde and Stanciole, 2006). Chronic stress, for example, a by-product of health inequalities, may result in higher susceptibility to the adverse effects of air pollution, like NCD deaths, because of lower immune system function when aggregating the research of Woodward and Kawachi [2000], Nurius et al. [2016] and Apte and Salvi [2017].

[6] The vertex turning point of my Health "Kuznets" curve is obtained as follows:

$$|\hat{\beta}_4/2\hat{\beta}_5| = |4.4156/2 \times -4.4588| \approx 0.495$$

Discussion

Concerning the effectiveness of the model, the standard error of the regression (SER) and the R^2 are 0.2033 and 0.7214, respectively. The SER of NCD_P is 0.2033. In other words, on average, the observed values are 0.2033 units from the regression line. When interpreting the R^2 value, 72.14% of the variation in NCD_P is explained by the regressors, $PM_{2.5}$, *eTech*, *DomHE*, *HDI* and HDI^2. Since my primary goal is to estimate the marginal effects of $PM_{2.5}$ and *eTech* on NCD_P, as opposed to prediction, R^2 is of lesser importance in my analysis.

The results of this analysis have some important implications. Even though the marginal effects of $PM_{2.5}$ and *eTech* on NCD_P may seem small and practically insignificant, my results indicate that $PM_{2.5}$ and *eTech* are both individually statistically significant at the .05 level. This provides evidence that the current rising trend of NCDs deaths may be slowed if effective measures are taken to combat air pollution and increase the distribution of efficient cooking fuels and technologies ("Preventing", 2017). As was highlighted earlier, an estimated 23% of global deaths in 2016 could have been prevented by living in healthier environments.

However, as mentioned in Section 4, because the World Health Organization's ambiguous data descriptions for some indicators, such as the indicators relating to NCD mortality, discussing the practical significance of my results is difficult due to lack of appropriate comparisons especially when estimating the marginal effects of a health indicator that is a ratio. No data was available or no calculations could be made without questioning the parameters of the data set that would produce a value that would indicate the numeric value of premature deaths in the units of people in any given country instead of using a ratio.

Furthermore, even though the marginal effect of a 0.1362% increase on NCD_P and a -0.0781% decrease on NCD_P ceteris paribus, from $PM_{2.5}$ and *eTech*, respectively, may seem small and practically insignificant, my results indicate that the lives of humans, regardless of the numerical amount, could be saved before they turn 71 years old and surpass the "premature" threshold if measures are taken to combat air pollution and increase the distribution of efficient cooking fuels and technologies. As mentioned in Section 2, even though Abegunde and Stanciole [2006] find that NCD related deaths can hurt the global economy, death can also have adverse effects for the health of friends and family members of the deceased person who are still alive which may in turn increase their risk for NCDs. Coping with death may cause increased levels of stress, emotional trauma, poorer diets and higher levels of alcohol intake (Hairston, 2019; Smith, 2019). Consequently, the mentioned reactions of the emotional responses of death increases the susceptibility of poorer health outcomes, like NCDs ("Risk factors", n.d.). Specifically, according to Johns Hopkins Medicine, chronic stress can lead to serious heart problems, including heart disease, an NCD that accounted for an estimated 9 million global deaths in 2016 (Nowbar et al., 2019). As a result, nations should still strive to fight against air pollution related NCD deaths because NCDs are "among the leading causes of preventable illness and related disability," and human death has both adverse economic and emotional effects on population health in addition to the loss of life ("U.S.", 2019, para. 3).

Conclusion

NCDs are "among the leading causes of preventable illness and related disability" and accounted for 16.9 million premature deaths in 2016 ("U.S.", 2019, para. 3). My study adds value to Bennett et al. [2018] and the NCD Countdown 2030 campaign as I have shown that ambient air pollution is associated with increases, and efficient cooking technologies are associated with decreases in the premature NCD ratio. Additionally, I find that HDI may be correlated with increasing marginal effects of health inequalities at low HDI levels but that the marginal effects in health inequalities peak and then begin to decrease at higher HDI levels; the Health "Kuznets" theory supports the statistic that over the past decade the number of NCD deaths have continued to increase as more than 85% of premature deaths are concentrated in developing countries ("Preventing", 2017; "U.S.", 2019).

Future research should test which sectors of our industrialized society like transportation or energy production have the largest impacts on $PM_{2.5}$. Greenhouse gasses like nitrogen and sulfate oxides, by-products of using gas powered vehicles, are $PM_{2.5}$ precursors[7] and contribute to 4 to 37% and 7 to 47%, respectively, to $PM_{2.5}$

[7] A $PM_{2.5}$ precursor is "any chemical that contributes to the formation of $PM_{2.5}$ particles but is not emitted directly from

formation (Hodan and Barnard, 2005). As a result, manufacturers could deflect the costs of pollutants more accurately.

Additionally, given my results, specifically the positive and statistically significant marginal effect of $PM_{2.5}$ on NCD_P, the collection and publication of data should be done more carefully and expeditiously because tracking trends in marginal estimates of the air pollution effects on population health is of great concern. For example, NCD_P data is released by the WHO every two to three years and the Environmental Protection Agency synthesizes research to assess air quality regulations only every five years. (Premature, n.d.; "Air", n.d.).

The results of this study, specifically the effects of ambient and household air pollution, should be beneficial for organizations like the UN to better understand the global implications of environmental health determinants. Hopefully these results as well as further research in epidemiology analytics contribute to the NCD Countdown 2030 campaign and help the campaign to reach its goal to "reduce by one third premature mortality from non-communicable diseases" by 2030 (Bennett et al., 2018, p. 1072).

its source as $PM_{2.5}$" (Hodan and Barnard, 2005, p. ii).

7 References

Air Monitoring, Measuring, and Emissions Research. (n.d.). *Environmental Protection Agency.* Retrieved March 05, 2020, from **https://www.epa.gov**.

Abegunde, D & Stanciole, A. (2006) *An estimation of the economic impact of chronic noncommunicable diseases in selected countries.* World Health Organization. Retrieved March 15, 2020, from **http://www.who.int/chp**.

Apte, K. & Salvi, S. (2016). Household air pollution and its effects on health [version 1; referees: 2 approved]. *F1000Research*, 5, **10.12688/f1000research.7552.1.**.

Beaglehole, R., Bonita, R., Ezzati, M., Alleyne G., Dain K., Kishore, S.P., & Horton, R., (2014). NCD Countdown 2025: accountability for the 25 × 25 NCD mortality reduction target. *The Lancet*, 384(9938), 105-107, **10.1016/S0140-6736(14)61091-6**.

Bennett, J.E., Stevens, G.A., Mathers C.D., Bonita R., Rehm J., Kruk M.E., Riley L.M., Dain K., Kengne A.P., Kalipso C., Beagley, J., Kishore, S.P., Chen, W., Saxena, S., Bettcher, D.W, Grove, J.T., Beaglehole, R., & Ezzati, M. (2018). NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. *The Lancet*, 392(10152), 1072-1088, **10.1016/S0140-6736(18)31992-5**.

Bokhari, F.A., Gai, Y., & Gottret, P. (2007). Government health expenditures and health outcomes. *Health Economics*, 16(3):257-73, **10.1002/hec.1157**.

Cosbey, A., Peterson, L.E., Pintér, L. (2005). Environmental Health and International Trade: Linkages and Methodologies. *International Institute for Sustainable Development.* Retrieved February 11, 2020, from **https://www.iisd.org**.

Davies, A.& Quinlivan, G.,M. (2006) A Panel Data Analysis of the Impact of Trade on Human Development *Journal of Socio-Economics.* 35(5):868-876, **10.1016/j.socec.2005.11.048**.

Hairston, Stephanie. (2019). How Grief Shows Up In Your Body. *WebMD.* Retrieved April 27, 2020, from **https://www.webmd.com**.

Hodan, W.M., & Barnard, W.R. (2004). Evaluating the Contribution of PM 2.5 Precursor Gases and Re-entrained Road Emissions to Mobile Source PM 2.5 Particulate Matter Emissions. *Environmental Science.* Retrieved March 11, 2020, from **https://www3.epa.gov.**.

Human Development Index (HDI). (n.d.). *United Nations Human Development Programee.* Retrieved March 29, 2020, from **http://hdr.undp.org**.

Mathers, C., Stevens, G.A., Mahanani, W.R., Fat, D.M., & Hogan,D. (2017). WHO methods and data sources for country-level causes of death 2000-2016. *World Health Organization.* Retrieved February 20, 2020, from **https://www.who.int**.

Meng, J., Yang, H., Yi, K., Liu, J., Guan, D., Liu, Z., Mi, Z., Coffman, D., Xuejun, W., Zhong, Q., Huang, T., Meng, W., & Tao, S. (2019). The Slowdown in Global Air-Pollutant Emission Growth and Driving Factors. *One Earth*, 1(1), 138-148, **10.1016/j.oneear.2019.08.013.**.

Molini, V., Nubé, M., & Boom, B. (2010). Adult BMI as a Health and Nutritional Inequality Measure: Applications at Macro and Micro Levels. *World Development.* 38. 1012-1023. **10.1016/j.worlddev.2009.12.003**.

Nowbar, A., Gitto, M., Howard, J.P.,Francis, D.P., Al-Lamee, R. (2019). Analysis of Data From the World Health Organization and Coronary Artery Disease Risk Factors From NCD Risk Factor Collaboration. *Cardiovascular Quality and Outcomes.* 8(2), 241–256. **10.1161/CIRCOUTCOMES.118.005375**.

Nurius, P. S., Green, S., Logan-Greene, P., Longhi, D., & Song, C. (2016). Stress pathways to health inequalities: Embedding ACEs within social and behavioral contexts. *International public health journal.* 12(6). **PMID: 27274786; PMCID: PMC4891624**.

Preventing noncommunicable diseases (NCDs) by reducing environmental risk factors. (2017). *Geneva: World Health Organization* (WHO/FWC/EPE?17.1). **License: CC BY-NC-SA 3.0 IGO**.

United Nations Development Programme. (2016). *Human Development Data (1990-2018)* [Data file]. Retrieved April 4, 2020, from **hdr.undp.org/**.

The U.S. Government and Global Non-Communicable Disease Efforts (2019). *Kaiser Family Foundation* Retrieved April 19, 2020, from **https://www.kff.org/**.

Smith, C. (2019, July 24). *Alcohol Rehab Guide*. Retrieved April 22, 2020, from **https://www.alcoholrehabguide.org**.

Soheylizad, M., Khazaei, S., Jenabi, E., Delpisheh, A., & Veisani, Y. (2018). The Relationship Between Human Development Index and Its Components with Thyroid Cancer Incidence and Mortality: Using the Decomposition Approach. *International Journal of Endocrinology and Metabolism*, 16(4);e65078, **10.5812/ijem.65078.**.

Williams, E., Buck, E., Babalola, G. (2020, 18 February) *KingsFund*. Retrieved April 30, 2020, from **https://www.kingsfund.org.uk**.

Woodward, A., & Kawachi, I. (2000). Why reduce health inequalities?. *Journal of epidemiology and community health*, 54(12), 923–929. **https://doi.org/10.1136/jech.54.12.923**.

World Bank Group. (2016). \textit{Domestic general government health expenditure per capita, PPP (current international $)} [Data file]. Retrieved March 13, 2020, from **https://www.who.int**.

World Bank Group. (2016). $PM_{2.5}$ *air pollution, mean annual exposure (micrograms per cubic meter)* [Data file]. Retrieved March 13, 2020, from **https://www.who.int**.

World Health Organization. (2016). *Percentage of the population with primary reliance on clean fuels and technologies at the household level* [Data file]. Retrieved March 13, 2020, from **https://www.who.int**.

World Health Organization. (2016) *Premature deaths due to noncommunicable diseases (NCD) as a proportion of all NCD deaths* [Data file]. Retrieved March 13, 2020, from **https://www.who.int**.

World Health Organization. (2016). *Total NCD Deaths (in thousands)* [Data file]. Retrieved April 18, 2020, from **https://www.who.int**.

Zheutlin, A.R., Adar, S.D., & Park, S.K. (2014). Carbon Dioxide Emissions and Change in Prevalence of Obesity and Diabetes in the United States: An Ecological Study. *Environmental International*, 74(1), 111-116, **10.1016/j.envint.2014.07.012.**.

8 Appendix

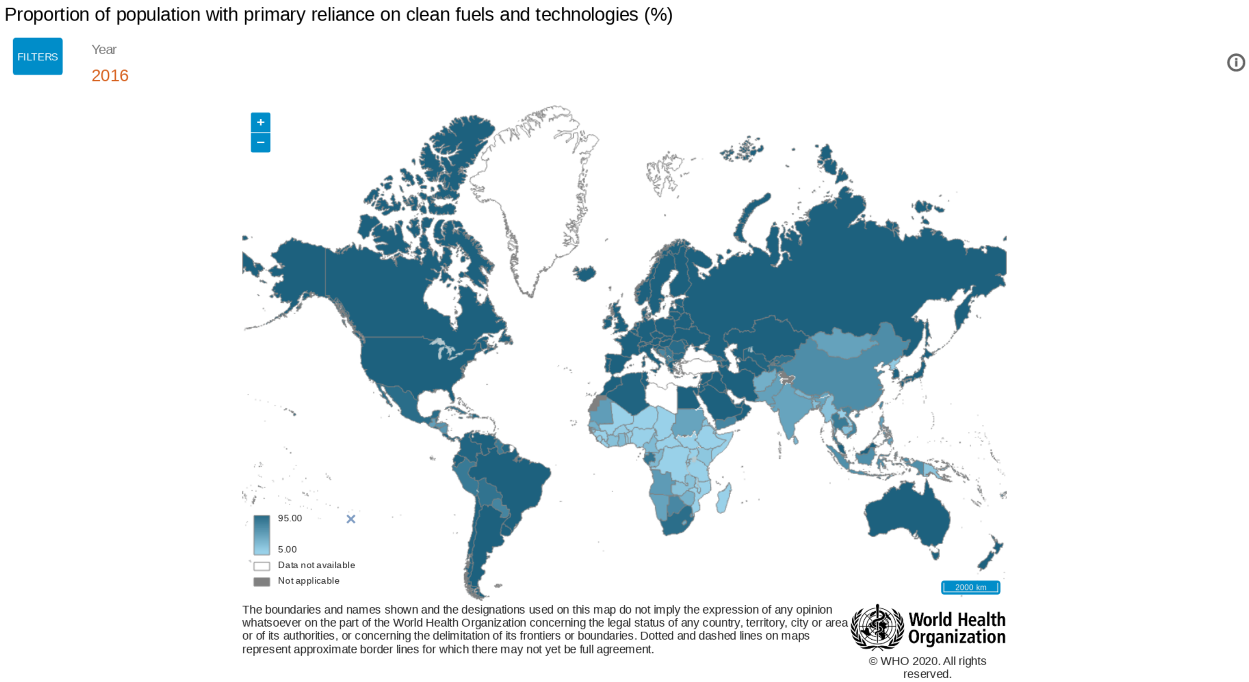

Figure 1: The 2016 Premature NCD mortality distribution across all countries. (Source: World Health Organization)

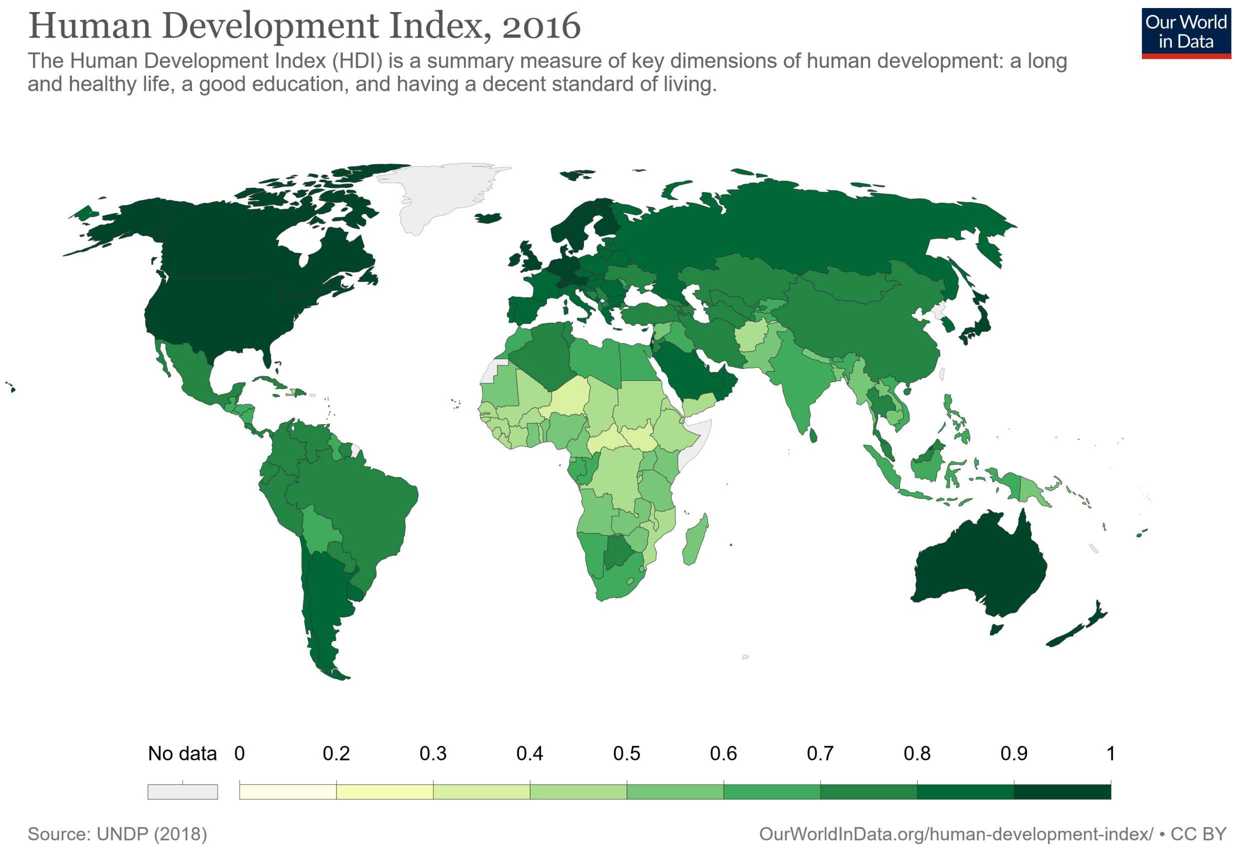

Figure 2: The 2016 $PM_{2.5}$ concentration levels distribution across all countries (Source: World Health Organization)

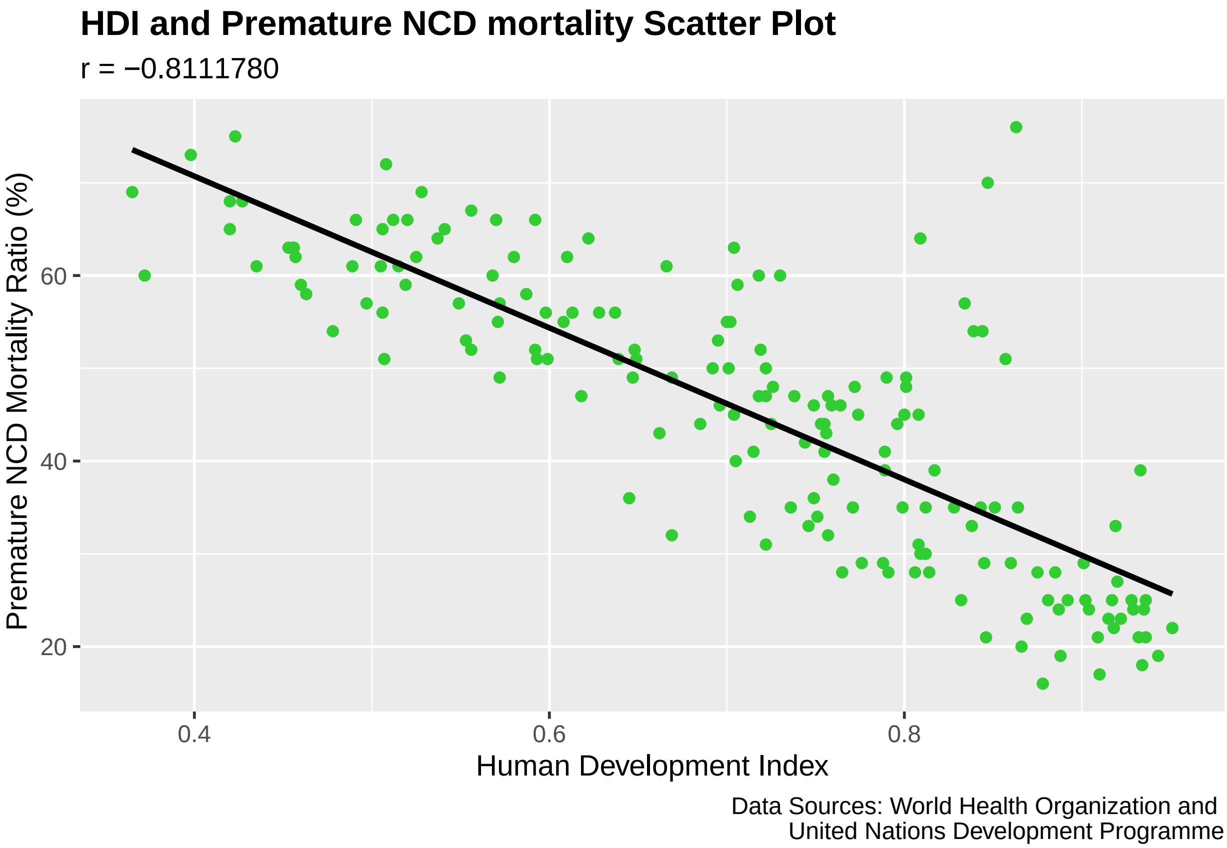

Figure 3: Scatter Plot between Human Development Index and Premature NCD mortality (2016)

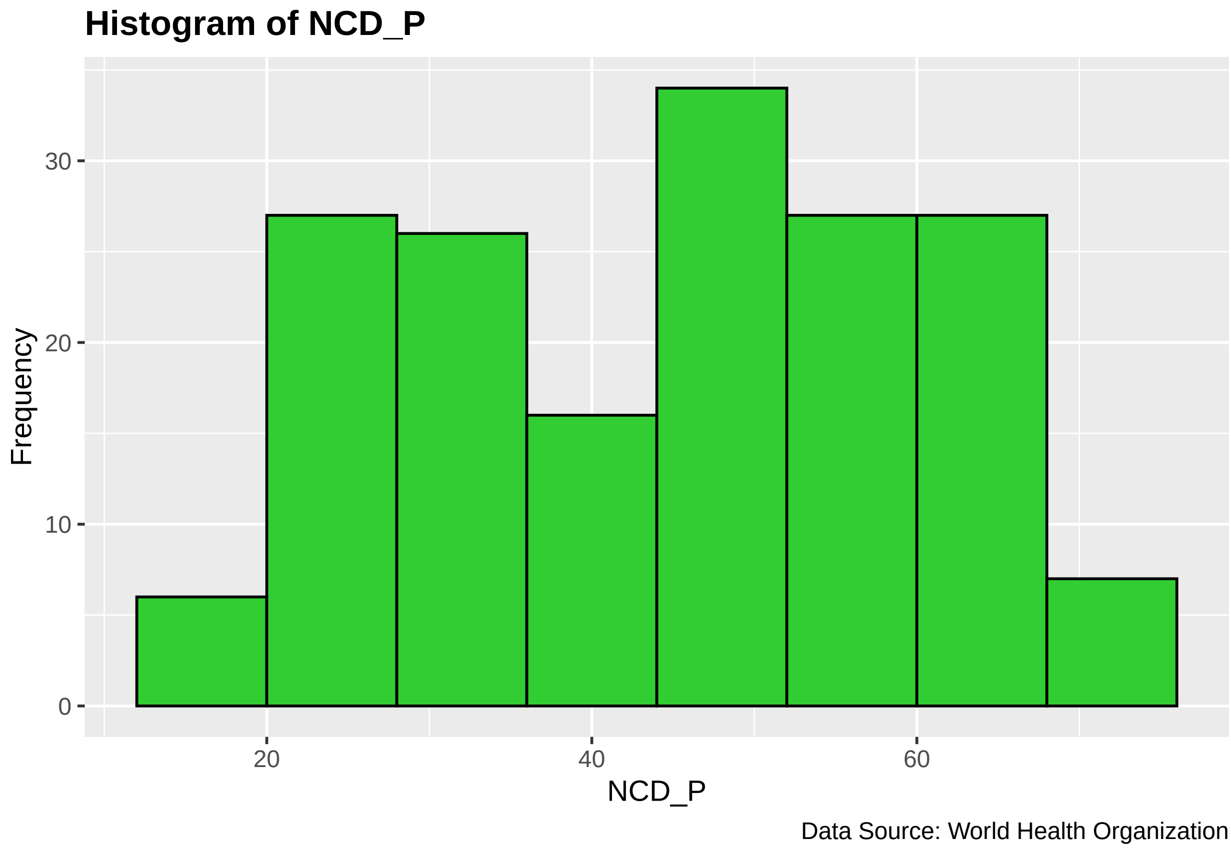

Figure 4: Histogram of the 2016 Premature NCD mortality distribution

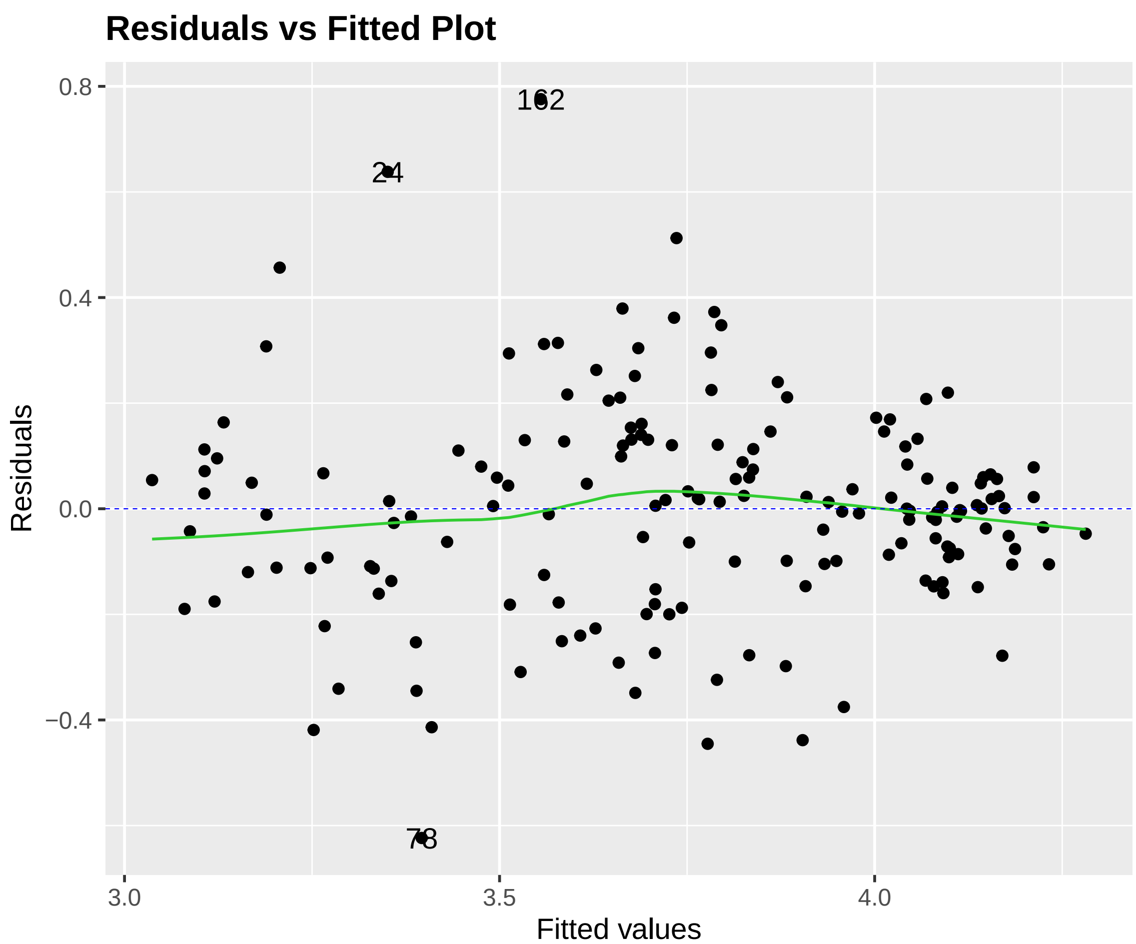

Figure 5: Validation of Homoskedasticity and Linearity in Parameters

Figure 6: Normality of the error term validation

	Estimate	Std. Error	Pr(>|t|)	95% CI [LB,UB]
(Intercept)	2.5952	0.4112	$2.48e^{-09}$*	1.7835, 3.4069
log($PM_{2.5}$)	0.1362	0.0319	$3.24e^{-05}$*	0.0733, 0.1991
log(eTech)	-0.0781	0.0360	0.0313*	-0.1492, -0.0071
log(DomHE)	0.0428	0.0307	0.1658	-0.0179, 0.1034
HDI	4.4155	1.2384	0.0005*	1.9705, 6.8601
HDI^2	-4.4588	0.8487	$4.58e^{-07}$*	-6.1341, -2.7834

Table 5: Summary Statistics with Biased Standard Errors.

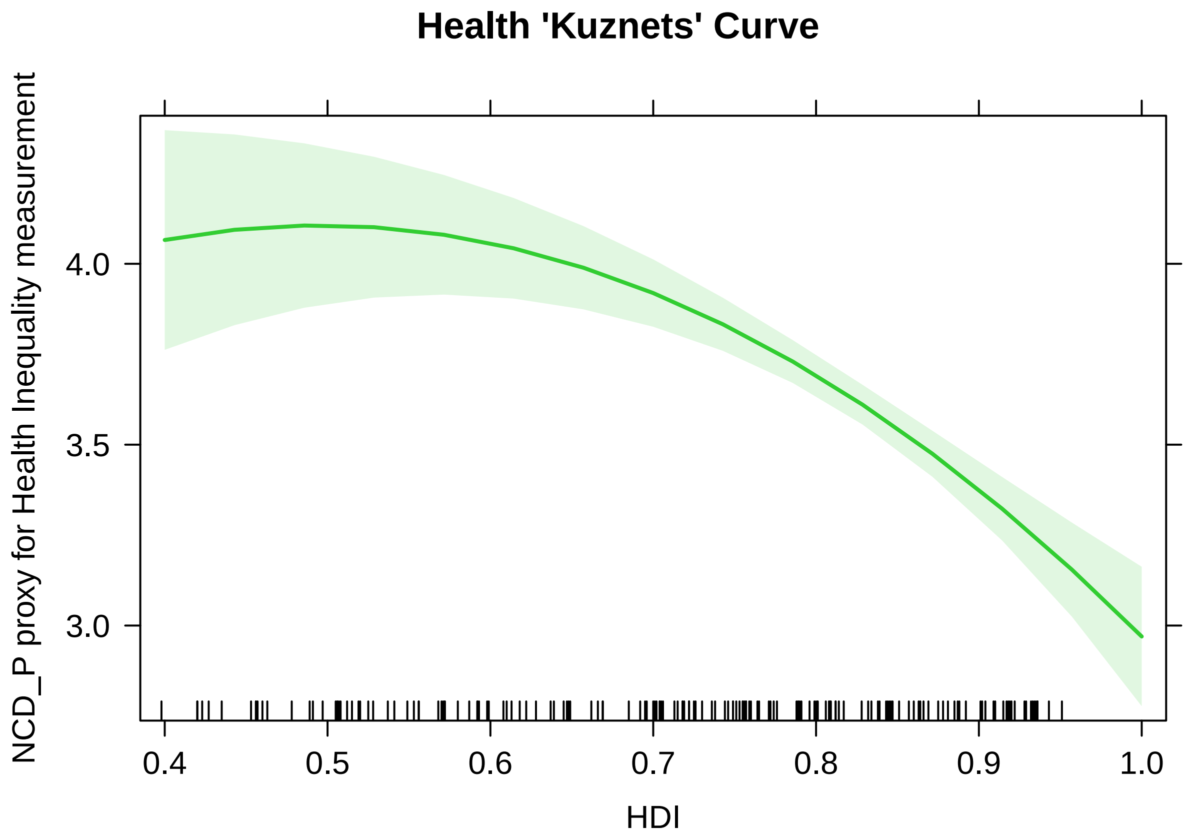

Figure 7: Health 'Kuznets' Curve: Association between health inequality (measured by the ratio of premature deaths) and economic development (measured by the Human Development Index)