# Quantitative Analysis of Disparities in Healthcare Access

Linh Dam, Kasen Groth, Dan Nguyen

Spring 2024

# 1 Abstract

Quantifying inequities in the distribution of healthcare access and understating the drivers behind these disparities are crucial for promoting equitable healthcare delivery. This study investigates effective mathematical measurements to assess inequities in healthcare access and explores the underlying factors influencing the inequitable distribution of mental health providers across counties in Pennsylvania. We employ the Gini coefficient, Atkinson index, and the weighted Utilitarian social welfare function to evaluate disparities in the availability of mental health providers, hospital resources, and geographical accessibility metrics. Furthermore, we conduct regression analyses to investigate the impact of socio-economic factors, including rural-urban classifications, household income levels, income inequality, life expectancy, and uninsured population rates on the distribution of mental health providers. By quantifying these relationships, we aim to identify significant predictors contributing to mental health inequities. This comprehensive approach provides a rigorous framework for measuring healthcare access inequities and elucidates the underlying socio-economic determinants driving disparities in mental health service provision. Our findings offer valuable insights for policymakers and healthcare planners to develop targeted interventions promoting equitable access to essential mental health services across diverse communities within Pennsylvania.

# 2 Introduction

Access to healthcare is a fundamental aspect of societal well-being and a crucial determinant of public health outcomes. However, disparities in healthcare access persist across various regions, posing significant challenges for policymakers and public health officials. Furthermore, quantifying and evaluating equity in access poses significant methodological challenges. In this research paper, we explore novel mathematical approaches to measure healthcare access inequity, utilizing a modified Gini index, Atkinson measure, and the weighted Utilitarian social welfare function.

Traditionally, measures of economic inequality such as the Gini index have been applied to assess disparities in income distribution. In our study, we adapt these measures to analyze disparities in healthcare access among different counties in Pennsylvania, using a framework developed in the paper "Measuring Equity in Access to Health Care" by Hugh R. Waters [6], where the author applied the method for a study conducted using data gathered from an Ecuadorian survey. We refer to this paper as the "Ecuadorian Study." Rather than using individual income levels, we substitute healthcare access indicators for each county, including the number of hospitals, hospital beds, doctors, nursing home beds, and average distance to the closest hospital. By incorporating these indicators, we aim to provide a comprehensive assessment of healthcare access inequity.

Moreover, we go beyond mere measurement and delve into the underlying reasons for healthcare access inequity. We investigate potential correlations between healthcare access indicators and socioeconomic factors such as income levels and demographic distribution. By examining these relationships, we seek to identify factors contributing to disparities in healthcare access and inform targeted interventions to address them.

This research contributes to the broader discourse on healthcare equity by offering a rigorous mathematical framework for analyzing and understanding disparities in healthcare access. By elucidating the drivers of inequity, our findings have implications for healthcare policy and resource allocation, with the ultimate goal of promoting equitable access to healthcare for all individuals, regardless of geographic location or socio-economic status.

# 3 Literature Review: The Ecuadorian Study

The Ecuadorian study [6] presents a conceptual framework and practical approach for assessing equity in healthcare utilization. The author defines equity in this context as ensuring individuals can obtain appropriate healthcare services regardless of their demographic, socioeconomic, or geographic characteristics. Going beyond just measuring insurance coverage, this framework focuses on identifying and minimizing barriers that prevent certain groups from accessing the services they need.

One metric discussed is the modified Gini coefficient ($G$), which compares the cumulative distribution of healthcare use to the cumulative distribution of relevant equity stratified, such as income or education level. The value of $G$ ranges from –1 to 1, with 0 indicating perfect equity. Waters explains how this index can be calculated using household survey data on healthcare utilization and socioeconomic status. This methodology also allows for disaggregation by type of service, potentially pinpointing the most inequitable aspects of the healthcare system.

Quantifying equity in access, as described in this paper, could provide valuable insights to guide policies and interventions aimed at promoting fairness in health care delivery. For example, identifying service areas or population groups with low concentration index values would indicate regions or demographics that require more targeted efforts to reduce access barriers. Furthermore, tracking changes in the concentration index over time could help evaluate the impact of equity-focused reforms.

However, the author notes that further research is needed to refine measurement approaches and link equity metrics more directly to improvements in population health outcomes. While quantifying equity in access is an important first step, the ultimate goal is to ensure that equitable access translates into better health for historically underserved or disadvantaged populations. Continued work is required to establish clear connections between access, utilization, and meaningful health impacts.

Overall, the conceptual framework and practical methods outlined in the Ecuadorian study offer a promising approach for policymakers and researchers seeking to evaluate and address disparities in healthcare access. By providing a rigorous way to quantify equity, this work can inform the development of targeted interventions and monitor progress toward the crucial goal of equitable health care for all.

# 4 Methodology

## 4.1 Research Method

The research method for this study involves a multifaceted approach aimed at comprehensively analyzing the distribution of mental healthcare across counties in Pennsylvania. Initially, we employ established measures of inequality such as the Gini index, the Atkinson index, and the weighted Utilitarian social welfare function. These metrics are adapted by substituting income variables with indicators of healthcare access.

Potential accessibility is a valuable measure for evaluating the availability of healthcare services, particularly in regions facing shortages of physicians. This approach combines spatial analysis techniques with healthcare data to quantify the ease with which individuals can access medical care within a given geographic area.

Firstly, potential accessibility accounts for both the spatial distribution of health care providers and the population's spatial distribution. Traditional metrics, such as physician-to-population ratios often fail

to capture the nuances of accessibility because they do not consider geographic barriers that can hinder individuals from reaching healthcare facilities. Potential accessibility takes into consideration the distance individuals must travel to reach a healthcare provider, as well as the availability of providers within a reasonable distance.

Secondly, potential accessibility provides a more realistic representation of healthcare access compared to static measures. Physician-to-population ratios may suggest an adequate number of physicians within a region, but they do not reflect the geographical distribution of these providers. In areas where physicians are concentrated in urban centers while rural areas remain underserved, potential accessibility can reveal significant gaps in healthcare access that might otherwise go unnoticed.

Average distance serves as an additional metric for evaluating potential accessibility to health care services. While metrics like travel time provide valuable insights, average distance complements these measures by offering a straightforward assessment of the spatial distribution of healthcare facilities. By calculating the average distance of municipalities within a county to the nearest hospital, planners, and policymakers gain a comprehensive understanding of geographical disparities in healthcare accessibility. This information aids in identifying underserved areas and formulating targeted interventions to address disparities.

The calculation of average distance involves determining the central point of all municipalities within a county and measuring the distance from each central point to the nearest hospital. These distances are then summed and divided by the total number of municipalities in the county to obtain the average distances. This methodology provides a quantitative measure of the average spatial separation between communities and healthcare facilities, allowing for comparative analysis across regions.

Despite its utility, average distance as a metric has inherent limitations that must be acknowledged. One significant limitation is the inability to account for road distance, leading to skewed results. Without considering road networks, the calculated distances may not accurately reflect the actual travel distance required to reach healthcare facilities. Moreover, the absence of road distance data precludes the conversion of distance into travel time, further limiting the utility of the metrics in assessing real-world accessibility. Additionally, variations in terrain, transportation infrastructure, and traffic patterns are not accounted for, potentially leading to inaccuracies in the assessment of accessibility.

From the data we found available, we decided to implement 6 healthcare access indicators:

- Number of primary care physicians per 1000 residents;

- Number of mental health providers per 1000 residents;

- Number of hospital beds per 1000 residents;

- Number of hospitals per 1000 residents;

- Number of nursing home licensed beds per 1000 residents;

- Average distance to the nearest hospital for each county.

Our analysis reveals that the distribution of mental health providers exhibits pronounced disparities across various metrics such as the number of primary care physicians, hospital beds, nursing home licensed beds, hospitals per 1000, and the average distance between towns in a county with the nearest hospital.

To further explore the factors influencing the distribution of mental healthcare providers, we conduct a comparative analysis between these indicators. We examine whether any particular indicator stands out as having a significant impact on the distribution patterns observed. This comparative analysis helps identify potential areas of focus for subsequent regression analysis, enabling us to delve deeper into understanding the drivers of healthcare inequity.

Following the identification of key indicators, we proceed to conduct regression analysis to investigate the factors that affect the distribution of mental healthcare providers. We include a range of explanatory variables in our regression models, including the Rural-Urban Continuum Code (RUCC) for each county, household income levels, income disparity within counties, life expectancy, and the percentage of uninsured individuals. By examining the relationship between these variables and the distribution of mental healthcare providers, we aim to uncover the underlying determinants of healthcare inequity.

Regression analysis allows us to quantify the impact of each explanatory variable on the distribution of mental healthcare providers, controlling for other factors. This enables us to identify significant predictors of healthcare inequity and assess their relative importance in influencing access to mental healthcare services. Through this comprehensive methodological approach, we seek to provide valuable insights into the factors driving disparities in mental healthcare distribution and inform targeted interventions aimed at promoting more equitable access to mental health services across Pennsylvania's counties.

## 4.2    Definitions

*Equity* is the fair and impartial distribution of health care resources and services aiming to address disparities and ensure that all individuals receive adequate and appropriate care regardless of their socio-economic status, race, ethnicity, or other factors. In the context of this study, equity in healthcare is specifically concerned with achieving parity in the distribution of healthcare resources and services among different population groups.

*Health care access* encompasses the ability of individuals to obtain timely and appropriate health care services when needed. It involves various features, including but not limited to geographical accessibility, financial affordability, and availability of healthcare facilities. For this research, healthcare access will primarily be evaluated and measured on geographical factors and the availability of healthcare infrastructure.

*Geographical accessibility* refers to the ease with which individuals can physically reach healthcare facilities from their place of residence. It is often measured in terms of distance or travel time to the nearest healthcare facility, such as hospitals or primary care clinics. In this study, geographical accessibility will be assessed based on the average distance from the municipality to the nearest hospital.

*The physician-to-population ratio* is a metric used to assess the availability of healthcare providers relative to the size of the population they serve. It is calculated by dividing the total number of primary health care physicians by the population in a given area. A higher physician-to-population ratio generally indicates better access to primary care services.

The *hospital beds-to-population ratio* is a measure of the availability of inpatient healthcare services relative to the size of the population. It is calculated by dividing the total number of hospital beds by the population in a given area. A higher hospital beds-to-population ratio suggests greater capacity for patients requiring hospitalization.

## 4.3    Gini coefficient

The modified Gini coefficient ($G$) is a measure used to assess equity in the distribution of access to health care across different socioeconomic groups. It adjusts the traditional Gini coefficient formula (which only compares values of a distribution against the other values in the same distribution) to explicitly compare two values from two different distributions.

The unmodified Gini coefficient is defined as the following formula:

$$G = \frac{1}{2I^2\mu} \sum_{i=1}^{I} \sum_{j=1}^{I} |h_i - h_j| \tag{1}$$

where:

- $G$ is the Gini coefficient,

- $h$ is the average access to health services,

- $\mu$ is the mean outcome for all groups,

- $I$ is the total number of groups.

In general, Gini coefficient tells us the difference we expect to find among the values, relative to the mean. To understand the formual above, we consider the following thought experiment. Imagine two people meeting randomly on the street. They compare their incomes and note the difference between them. How large would we expect this difference to be on average?.

The Gini coefficient calculates this expected difference between two randomly chosen people by taking the average difference between all possible pairs of people. In a society where incomes are distributed equally, we would expect the difference between two randomly selected people's incomes to be small. Conversely, in a highly unequal society, we would expect this difference to be large. However, when measured in absolute terms, the expected difference also depends on the overall distribution of incomes. In a society where even the wealthiest individuals have relatively low incomes, the absolute difference between people's incomes cannot be high.

To address this issue, the unmodified Gini coefficient expresses the expected absolute difference between people's incomes relative to the mean income of the population. Specifically, it is calculated as the expected difference as a proportion of twice the mean income. Twice the mean income represents the highest possible value for the average difference—a scenario of perfect inequality, where one person has all the income and everyone else has none. Therefore, in the case of maximum inequality, the Gini coefficient is 1.

To understand the unmodified Gini coefficient in a different way, we introduce the Lorenz curve. The Lorenz curve plots the cumulative percentage of total income against the cumulative percentage of the population. The x-axis shows the cumulative percentage of households, and the y-axis shows the cumulative percentage of income. The line of perfect equality is a straight diagonal line from the origin (0,0) to (1, 1). This line represents the situation where every percentage of the population has the same percentage of total income. The area under the Lorenz curve is the actual distribution of income (the green area with label B). The Gini coefficient is essentially the ratio of the area between the line of perfect equality and the Lorenz curve (the grey area with label A) over the total area under the line of perfect equality. When Gini coefficient is 0, the distribution is represented by the line of perfect equality. Thus, value 0 represents perfect equality and value 1 represents perfect equality.
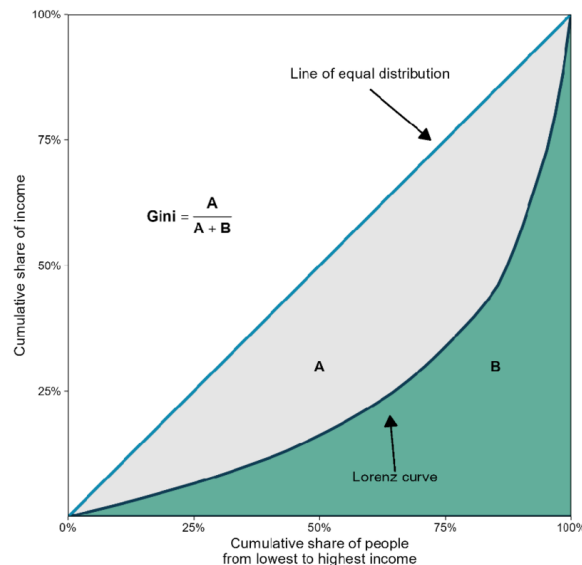
We have a rewritten formula for Gini coefficient as

$$G = 1 + \frac{1}{I} - \frac{2}{I^2 \mu}(h_1 + 2 * h_2 + \cdots + i * h_i + \cdots + I * h_I) \tag{2}$$

for $h_1 \geq h_2 \geq \cdots \geq h_i \geq \cdots \geq h_I$.

In this formula, the Gini coefficient weighs the outcome for each group by the rank of that group in the overall distribution of the outcome. In the paper, the groups are sorted by income from highest to lowest. Value of $h_i$ is the access to health care for the $i^{th}$ rank group by income. A negative value for the Gini coefficient means that there is health inequity but the inequity favors those with lower incomes. Conversely, a positive value for the Gini coefficient means that there is health inequity but the inequity favors those with higher incomes. The modified Gini coefficient can have negative values. The negative values occur when the sum of $h_1 + 2 * h_2 + \cdots + I * h_I$ is larger than the product of $I^2$ and mean value of all groups. Because we are organizing the order based on income, the value of $h_i$ is not guaranteed to be in decreasing order. If a lower income county $i$ has higher value of $h_i$, the coefficient of $h_i$ will be larger than the coefficient of a county $j$ with higher income and lower value of $h_j$. Thus, the result will be negative.

In table 1 below, we are using the number of hospitals per 1,000 residents as a health care access measurement. In this example, there is health inequity because one county has 10 hospitals while the other two have 80 and 90 hospitals. County $A$ with the fewest hospitals is also the richest. Thus, this data has health inequity and the inequity favors the lower income county. The Gini coefficient should be negative for this example.

| County | Number of hospitals | Income |
|--------|---------------------|--------|
| A | 10 | 50000 |
| B | 80 | 20000 |
| C | 90 | 10000 |

Table 1: Example for Gini coefficient

There are three counties thus $I = 3$. The average access to health care is $\mu = \frac{10+80+90}{3} = 60$. Next, we calculate the modified Gini coefficient and get a negative value as we expected.

$$G = 1 + \frac{1}{3} - \frac{2}{3^2 * 60}(10 + 2 * 80 + 3 * 90) = -0.296$$

## 4.4   Atkinson Index

The Atkinson index is a measure of economic inequality that takes into account both the extent of inequality and the society's aversion to inequality. In this study, we utilized the measure to assess the extent to which total healthcare access could be reduced without diminishing social welfare, under the assumption that the reduced healthcare is equally distributed. Essentially, it quantifies the overall welfare loss resulting from an unequal distribution of health care.

The Atkinson measure merges an inequality indicator with society's degree of aversion to inequality. It is defined by the following formula:

$$A = 1 - \left(\sum_{i=1}^{I} \frac{1}{I} \left(\frac{h_i}{\bar{\bar{h}}}\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \tag{3}$$

where:

- $A$ is the Atkinson index,
- $I$ is the number of counties,

- $h_i$ is the healthcare access of individual county $i$,

- $\bar{h}$ is the mean healthcare access of the Pennsylvania,

- $\varepsilon$ is the Atkinson inequality aversion parameter, with $\varepsilon > 0$ and $\varepsilon \neq 1$.

As $\varepsilon \to \infty$, only the outcome for the individual with the lowest score contributes to the calculation of $A$. When $\varepsilon = 0$, the Atkinson index becomes the Gini coefficient. $A$ equals zero when outcomes are perfectly evenly distributed, and approaches 1.0 as the distribution becomes increasingly unequal or $\varepsilon$ approaches infinity. Here, $\varepsilon$ is set to 1.5 to indicate a relatively strong aversion to inequality.

The Atkinson index is derived from the social welfare function and measures the equivalent healthcare access level that would provide the same level of social welfare in a society with perfect equality. The Atkinson index provides a comprehensive measure of inequality that reflects both the distribution of healthcare access and societal preferences regarding inequality.

## 4.5 Weighted Utilitarian social welfare function

Because the indices $G$ and $A$ do not take into account the maximization of overall welfare, we also employ a weighted Utilitarian social welfare function ($U$). Both indices $G$ and $A$ would indicate perfectly equitable scores if no county had access to healthcare services — a scenario far from ideal for social planners or public health officials. Therefore, to address this limitation, the first two indicators are complemented here by the inclusion of a weighted Utilitarian social welfare function. It is a measure that calculates the sum of individual utilities across a population, aiming to maximize the total utility or happiness of society. It is a foundational concept in economic and social analysis [11] and serves as a benchmark for evaluating policy decisions and their impact on collective well-being. The weighted Utilitarian social welfare function ($U$) is:

$$U = \sum_{i=1}^{I} \alpha_i h_i \tag{4}$$

where:

- $U$ is the weighted Utilitarian social welfare function,

- $I$ is the number of counties,

- $h_i$ is the healthcare access of individual county $i$,

- $\alpha_i$ is weights assigned to each county.

The weights assigned ($\alpha_i$) are determined based on the household expenditure quintile, where households in the fifth quintile (the wealthiest) are assigned a weight of 0.20, those in the third quintile a weight of 0.33, and those in the first quintile a weight of 1.00. In this study, we assume that the marginal benefit of receiving healthcare is consistent across counties, translating to an equal level of welfare.

## 4.6 Multiple Regression

Multiple regression is a statistical technique used to understand the relationship between a dependent variable and two or more independent variables. It's an extension of simple linear regression, which involves only one independent variable.

The dependent variable ($Y$) is the variable that you are trying to predict or explain. The independent variables ($X_1, X_2, ...$) are the variables you believe to have an influence on the dependent variable. They are also called the predictor variables, explanatory variables, or features. The relationship between the dependent variable and independent variables is expressed through a linear equation.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \varepsilon$$

Here $\beta_0$ is the intercept and $\beta_1, \beta_2, ..., \beta_n$ are the coefficients or weights for the independent variables $X_1, X_2, ..., X_n$, and $\varepsilon$ is the error term representing the variability in $Y$ that is not explained by the model.

The goal is to find the best-fitting line that minimizes the difference between the predicted values ($Y$) and the actual values, using the least squares method. Each coefficient ($\beta$) in the regression equation represents the change in the dependent variable for a one-unit change in the corresponding independent variable, holding all other variables constant. For example, if $\beta_1$ is 2.5, it means that for every one-unit increase in $X_1$, $Y$ is expected to increase by 2.5 units, assuming all other variables remain constant.

The standard error ($SE$) is a measure of the uncertainty or variability in the estimate of the coefficient. It tells you how much the coefficient may vary from the true population value if you were to estimate it from a different sample of data. A low standard error indicates that the estimate of the coefficient is relatively precise, while a high standard error indicates that the estimate is less precise and has more uncertainty.

Multiple regression assumes that there is a linear relationship between the independent variables and the dependent variable, the residuals (differences between predicted and actual values) are normally distributed, and there is not a high correlation among independent variables.

# 5    Findings

## 5.1    Gini coefficient

Figure 1 presents the Gini coefficients for different healthcare variables distribution. A higher Gini value suggests greater inequity in the distribution of that particular healthcare resource among the population.

The Gini coefficient for hospital beds per 1,000 residents and hospital bed provider per 1,000 residents shows that there is a relatively high degree of inequity in the distribution of hospital beds. This means some areas may have a higher concentration of beds than others. Nursing home licensed bed per 1,000 residents and average distance to nearest hospitals have lower Gini coefficients, which means these variables are more evenly distributed compared to the other resources. This means there is a more even distribution of nursing beds and the distance to the nearest hospital among all the counties.
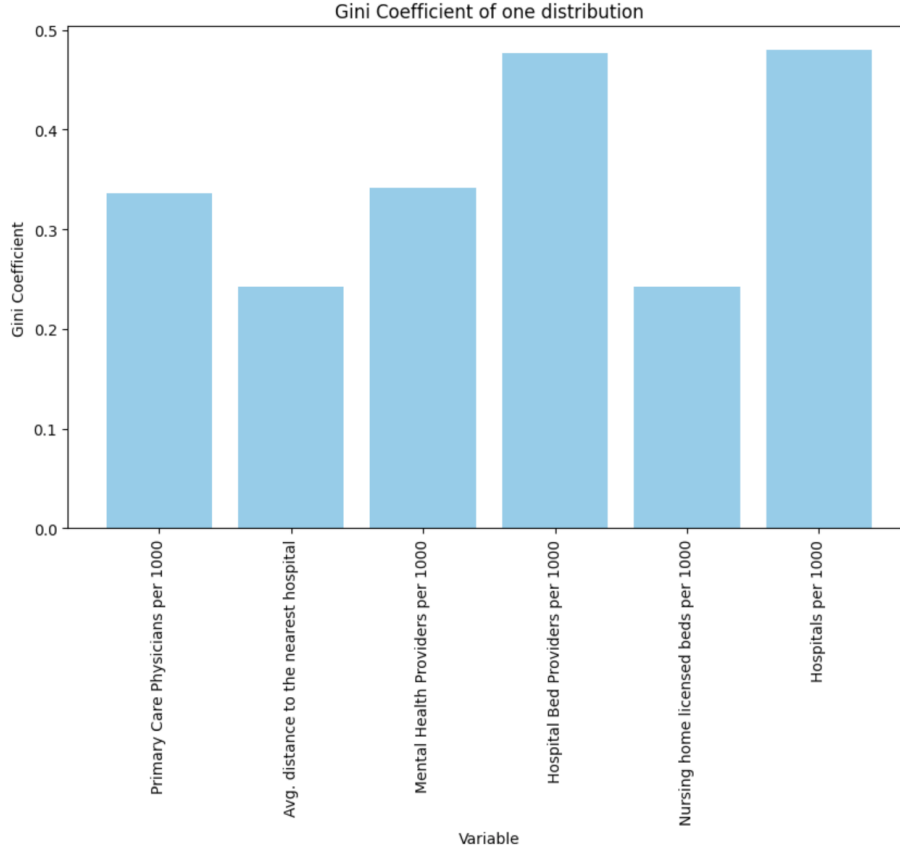
Figure 1: Gini coefficient of one distribution

Figure 2 shows the Gini coefficient when we compare values of income distribution with healthcare resources distribution. We first rank the counties by the richest to the poorest. Thus, the value $h_1$ will be the value of access to health care for the richest county. A negative Gini coefficient suggests a distribution of healthcare resources that favors counties with lower incomes. Conversely, a positive Gini coefficient suggests a distribution of healthcare resources that favors counties with higher incomes.

When we use the Gini index to compare a variable with income, the Gini index has both positive and negative values. The number of mental health providers per 1000 residents and the number of primary care physicians per 1000 residents are the only two variables with positive Gini coefficient, showing health inequity favoring counties with higher incomes. The other variable have negative values, showing health inequity favoring those with lower incomes.

The negative Gini coefficient for number of hospitals per 1000 residents, number of hospital beds per 1000 residents and number of nursing home licensed beds per 1000 residents confirm there are disparities in access to healthcare facilities, but the disparity favors those with lower incomes. Out of these variables, number of hospitals per 1000 residents and the average distance to the closest hospital have higher absolute values of the Gini coefficient. Thus, there is health disparity but the disparity favors those with lower incomes when it comes to the number of hospitals and distance to the closest hospital.

## 5.2   Atkinson Measure

In our study, we calculated the Atkinson measure across a range of 80 values of $\varepsilon$, spanning from 1 to 5, to assess the sensitivity of each healthcare access indicator to changes in society's aversion to inequality. This

9

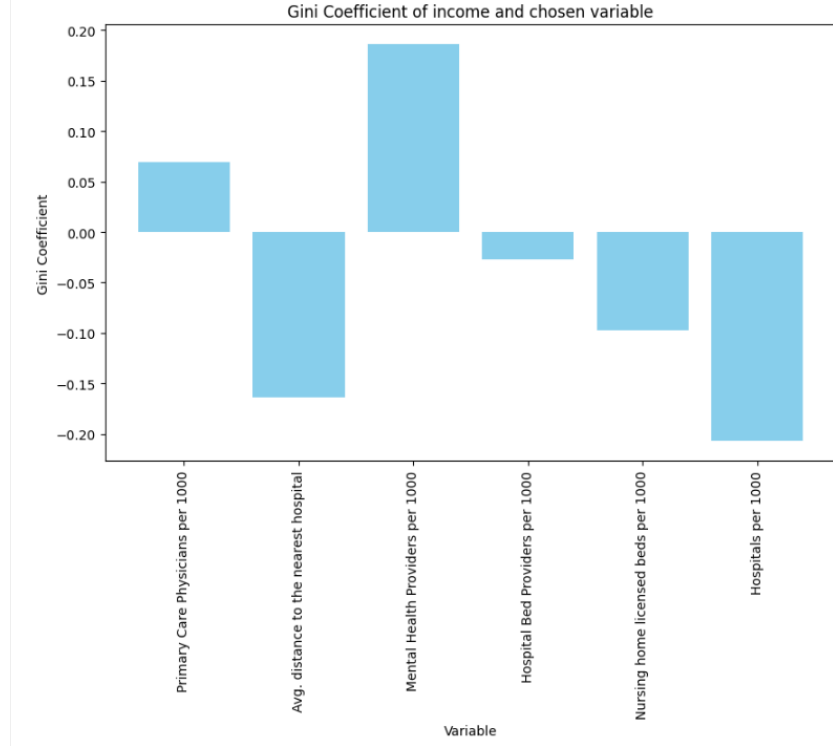Figure 2: Gini coefficient of two distributions

range of $\varepsilon$ allowed us to focus on the more meaningful portion of the curve where Atkinson measures were less likely to be cut off at 1, providing more accurate and reliable data for our analysis. By concentrating on values of $\varepsilon$ within this range, we were able to avoid the issues associated with extreme values close to 1, which could distort the results and complicate the process of finding clear graphs of best fit.

By examining such a broad spectrum of values of $\varepsilon$, we want to evaluate how each healthcare access indicator responds to shifts in the emphasis placed on different parts of the distribution. Through this comprehensive analysis, we can better understand how sensitive healthcare access is to changes in societal preferences for equality, providing valuable insights for policymakers aiming to create more equitable healthcare systems.
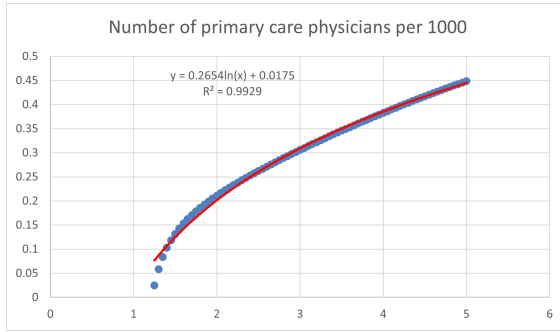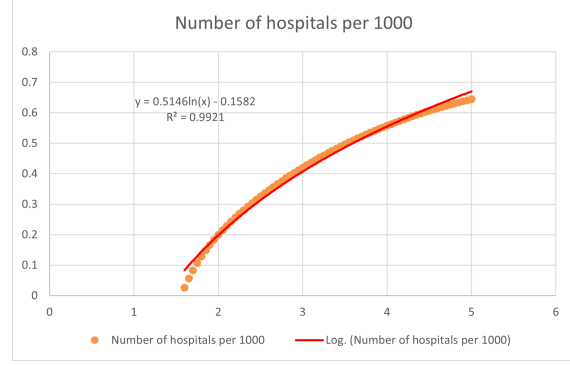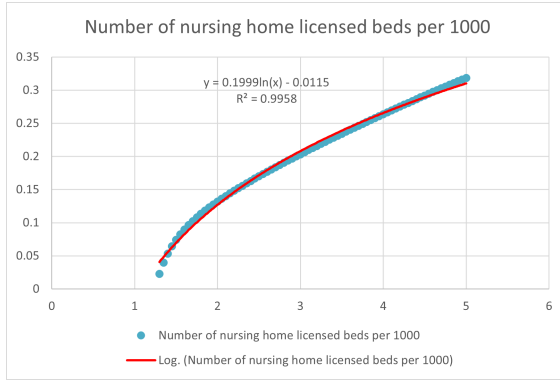


(a) Primary Care Physician per 1000
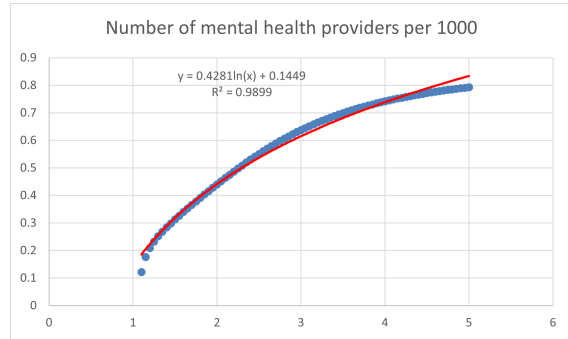
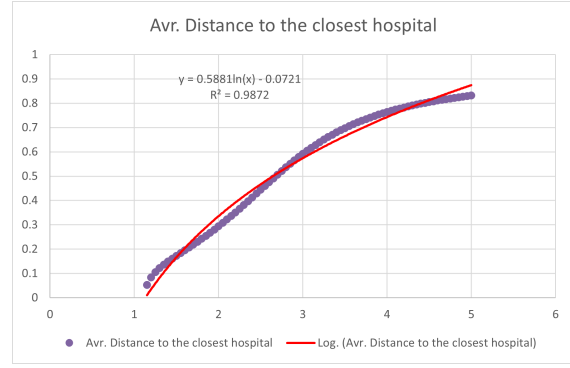(b) Mental Health Providers per 1000

10
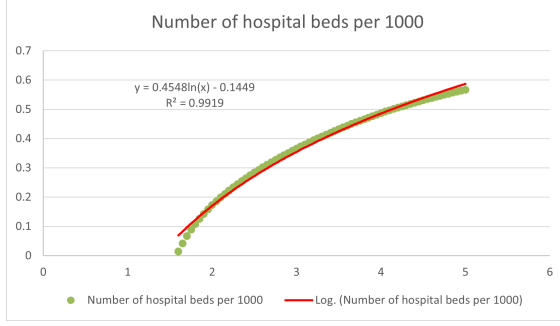
(c) Hospital Beds per 1000



(d) Hospitals per 1000



(e) Nursing Home Licensed Beds per 1000



(f) Avr. Distance to the Closest Hospital

Logarithmic regression is a statistical method used to model the relationship between a dependent variable and an independent variable when the relationship is not linear. Instead of fitting a straight line to the data, logarithmic regression uses the logarithm of the independent variable, which can be useful for transforming the data and capturing non-linear relationships more accurately. In this form of regression, the model equation is often represented as

$$y = a + b \cdot \log(x) \tag{5}$$

where:

- $y$ is the dependent variable;

- $x$ is the independent variable;

- $a$ is the intercept;

- $b$ is the slope of the line.

This approach is particularly helpful when the data has a multiplicative relationship or when there are diminishing returns, as it can better capture these trends.

The $R$-squared ($R^2$) value is a statistical measure that indicates the proportion of variance in the dependent variable that is explained by the independent variable(s) in the model.

$$R^2 = 1 - \frac{\text{var}(y - \hat{y})}{\text{var}(y)} \tag{6}$$

where:

11

- $R^2$ represents R-squared;

- var(y) represents the variance of the observed values (y);

- $\hat{y}$ represents the predicted values;

- var$(y - \hat{y})$ represents the variance of the residuals $(y - \hat{y})$.

In other words, it indicates how well the regression model fits the data. An $R^2$ value close to 1 suggests that the model explains a large proportion of the variance, while an $R^2$ value close to 0 indicates that the model explains very little of the variance.

We chose to adapt logarithmic regression to our data due to the behavior of the Atkinson value as $\varepsilon \to \infty$. As $\varepsilon$ increases, the Atkinson value approaches 1. Given this trend, using logarithmic regression allows us to better model the non-linear relationship between $\varepsilon$ and the Atkinson value. This approach captures the decreasing rate of change in the Atkinson value as $\varepsilon$ increases and helps us understand how sensitive the Atkinson measure is to variations in societal aversion to inequality. By employing logarithmic regression, we aim to provide a more nuanced analysis of how the Atkinson value responds to changes in $\varepsilon$, offering deeper insights into the distribution of healthcare access and potential policy implications.

In analyzing the Atkinson measure across the six healthcare access indicators, a comparative perspective reveals distinct patterns of inequality as values of $\varepsilon$ increase. Firstly, the $R^2$ values for all the indicators are above 0.987, indicating the strong reliability of the method. All indicators show positive relationships with increasing $\varepsilon$, indicating that as societal aversion to inequality rises, the Atkinson measure for these indicators increases. However, the rate of change differs: the average distance to the closest hospital exhibits the strongest increase in inequality (slope of 0.5881) compared to the weakest of the nursing home licensed beds (slope of 0.1999). This suggests that some indicators are more sensitive to societal preferences for reducing disparities than others are.

The comparative analysis of the six healthcare access indicators provides valuable insights for policymakers when considering how to address inequalities in healthcare distribution. The higher the slope coefficient in the relationship between the Atkinson measure and values of $\varepsilon$, the more sensitive an indicator is to society's aversion to inequality. In simpler terms, this implies that greater social welfare is lost due to the unequal distribution of an indicator with a higher slope value compared to one with a lower slope value.

For instance, "Mental health providers per 1000", "Hospitals per 1000", and "Average distance to the nearest hospital" have higher slope coefficients, suggesting that these aspects of healthcare access are more sensitive to changes in societal preferences for reducing in. This means that as society becomes more averse to inequality, the disparities in these healthcare access indicators lead to greater losses in social welfare. Unequal access to these services can result in poorer health outcomes for certain populations, increasing the overall burden on the healthcare system and exacerbating existing disparities.

Policymakers can take these findings into account when designing interventions and allocating resources. Addressing inequalities in the indicators with higher slope values, such as mental health providers, hospitals, and hospital access based on distance, should be a priority. By focusing on these areas, policymakers can mitigate the adverse effects of unequal distribution, thereby improving social welfare and promoting a more equitable healthcare system.

## 5.3 Weighted Utilitarian social welfare Function

When calculating $U$ across different healthcare access indicators, it is essential to standardize the data to account for variations in scale and range. The six indicators under study have different scales and ranges, which can obscure the analysis and comparisons if not properly addressed. To standardize the data, we first scaled all values using the mean and the standard deviation. This process involves subtracting the mean from each data value and then dividing by the standard deviation, according to the formula

$$z = \frac{(x - \mu)}{\sigma}, \tag{7}$$

where $z$ represents the standardized data value, x is the original data value, $\mu$ is the mean of the dataset, and $\sigma$ is the standard deviation of the dataset.

By standardizing the data in this way, we transformed all indicators into a common scale with a mean of zero and a standard deviation of one. This approach allows for a fair and accurate comparison of the different healthcare access indicators within $U$. It ensures that each indicator contributes equally to the analysis, regardless of its original scale and range. This standardized method provides a more robust and meaningful assessment of social welfare across different healthcare access measures.

Since a greater average distance to the closest hospital indicates poorer healthcare access, we will multiply every scaled value of the distance indicator by -1. This transformation effectively reverses the direction of the scale, making smaller values represent worse access and larger values represent better access. By doing this, we ensure consistency in interpreting the results across all healthcare access indicators, allowing us to maintain the same positive correlation between higher values and better healthcare access for all measures.

In the context of $U$, larger or smaller values have significant implications for the overall welfare of a society. A larger value of $U$ indicates that the sum of individual utilities is higher, suggesting that the society as a whole is experiencing greater overall welfare. This means that resources, services, and opportunities are more evenly distributed across the population, leading to higher levels of satisfaction and well-being for most individuals. Policies that result in larger weighted Utilitarian social welfare function values are generally seen as successful in promoting a high standard of living and improving quality of life. Conversely, a smaller value of $U$ implies that the sum of individual utilities is lower. Policies that lead to smaller weighted Utilitarian social welfare function values may need to be adjusted to address these disparities.

| Healthcare Access Indicator | U |
|---|---|
| Primary care physicians per 1000 | -2.2668 |
| Mental Health Providers per 1000 | -9.8175 |
| Hospital Beds per 1000 | 0.5698 |
| Avr. Distance to the closest hospital | -12.1366 |
| Nursing Home Licensed Beds per 1000 | 5.7315 |
| Hospitals per 1000 | 5.7799 |

Table 2: Values for Six Healthcare Access Indicators

These values provide insights into the overall welfare and well-being of society based on the availability and accessibility of healthcare services. The two highest values among the six indicators are the "Number of hospitals per 1000" and the "Number of nursing home licensed beds per 1000" with Utilitarian values of 5.7799 and 5.7315, respectively. They suggest relatively favorable distributions of the indicators. This positive value implies that the availability and distribution of hospitals and nursing home beds may be well-aligned with the needs of the population, contributing positively to social welfare.

The two indicators with the lowest $U$ values are "Mental health providers per 1000" and the "Average distance to the closest hospital", with values of -9.8175 and -12.1366, respectively, showing a significantly negative impact on social welfare. This points to substantial disparities in the availability of mental health services and significant issues with hospital accessibility geographically. Such an imbalance can have serious implications for the mental health, well-being, and timely accessibility to medical services of the population.

## 5.4  Analysis

| Healthcare Indicator | Measures | | |
|---|---|---|---|
| | Gini Index | Atkinson Index ($\varepsilon = 2$) | Utilitarian Index |
| Primary Care Physicians per 1000 | 0.069 | 0.2110 | -2.2668 |
| Mental Health Providers per 1000 | 0.185 | 0.4399 | -9.8175 |
| Hospital Beds per 1000 | -0.027 | 0.1725 | 0.5698 |
| Nursing Home Licensed Beds per 1000 | -0.097 | 0.1319 | 5.7315 |
| Hospitals per 1000 | -0.207 | 0.1994 | 5.7800 |
| Average Distance to the closest hospital | -0.163 | 0.2939 | -12.1366 |

Table 3: Healthcare Access Indicators and Measures of Inequality

For comparative analysis purposes, the $\varepsilon$ value of the Atkinson measure is set to 2 to represent a relatively strong aversion to inequality. When we analyzed the various healthcare access indicators in Table 3, one area that stood out was the distribution of mental health providers per 1000 residents. The Gini index for mental health providers was 0.185, suggesting a moderate level of inequality in their distribution favoring those with higher incomes. Moreover, the Atkinson index at a value of $\varepsilon$ of 2 was at 0.4399 (the highest of all indicators). This highlights an area that demands urgent attention.

Moreover, the weighted Utilitarian social welfare value for mental health providers was -9.8175, one of the lowest across all healthcare access indicators. This negative value points to a significant loss in social welfare due to the uneven distribution of mental health providers. Such disparities can lead to inadequate access to vital mental health services for many individuals, exacerbating mental health challenges and limiting opportunities for a healthy, productive life.

Given these findings, we have decided to delve deeper into the analysis of mental health care providers' distribution in the next parts. By focusing on mental health, we hope to contribute to a more balanced and equitable healthcare system that prioritizes both physical and mental well-being.
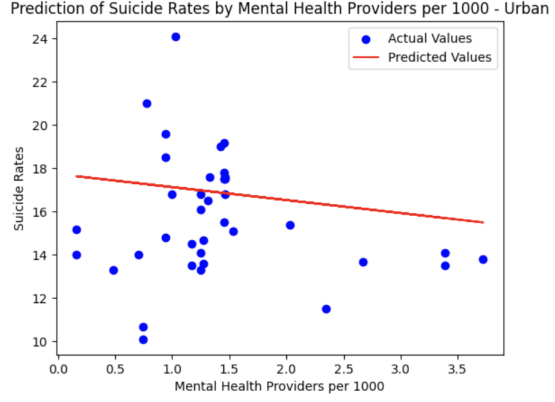
## 5.5  Mental Health Provider and Suicide Rate

Next, we look into the correlation between "Mental health provider per 1000" and the "Suicide Rate" in Pennsylvania. We split the data to two areas: the urban area, which has Rural-Urban Continuum Codes of 1 to 3, and the Rural-Urban Continuum Codes of 4 to 9. We run a linear regression model and calculate the $R^2$ value for both Rural and Urban areas.
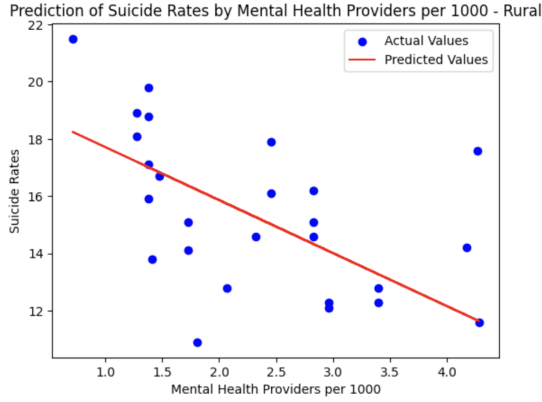
The downward slope of the regression line suggests a negative correlation between the number of "Mental health providers per 1000" people and the suicide rates in both rural and urban areas: As the number of mental health providers increases, there is a general trend of decreasing suicide rates. The $R^2$ for rural is at 0.245 meaning approximately 24.57% of the variation in suicide rates in rural areas can be explained by the number of mental health providers rate. For urban areas, the negative $R^2$ means the model fits the data worse than a horizontal line representing the mean of the dependent variable.

## 5.6   Mental Healthcare Access Factors

The multiple regression analysis investigating the association between the density of mental health providers across Pennsylvania counties and several key factors, including the "Hospitals per 1000", "Hospital beds per 1000", percentage of uninsured, rural-urban classification, and median household income, revealed compelling insights. The overall regression model exhibited robust statistical significance and accounted for 62.2% of the variance in mental health provider availability across the counties. This substantial explanatory power emphasizes the relevance of the selected predictor variables in capturing the underlying dynamics that shape the distribution of these critical healthcare resources.

|  | Coeff. | Std. Err. |
|---|---|---|
| Intercept | 2.2935 | 0.472 |
| Numbers of Hospitals per 1000 | 1.3177 | 6.547 |
| Numbers of Hospital Beds per 1000 | 0.1581 | 0.051 |
| Percentage of Uninsured | -0.0769 | 0.051 |
| RUCC | -0.1206 | 0.051 |
| Median Household Income ($10k) | 0.013 | 0.004 |

Table 4: coefficients of Linear Model

Upon closer examination of the individual coefficients, several noteworthy patterns emerge. The positive coefficient associated with the number of "Hospitals per 1000" (1.3177) suggests a positive correlation between a county's healthcare infrastructure and the availability of mental health services. Counties with more hospitals tends to have a higher density of mental health providers.

The positive coefficients associated with the number of "Hospital beds per 1000" (0.1581) residents suggest a positive correlation between a county's healthcare infrastructure and the availability of mental health services. Counties with a more robust presence of hospitals and hospital beds tend to have a higher density of mental health providers, indicating a potential spillover effect or synergistic relationship between the broader healthcare system and mental health service provision.

Conversely, the negative coefficient for the percentage of uninsured residents (-0.0769) highlights the potential barriers posed by a lack of health insurance coverage. Counties with a larger uninsured population appear to have fewer mental health providers per capita, emphasizing the importance of access to affordable health care in ensuring the availability of mental health services.

The rural-urban classification code (RUCC) coefficient (-0.1206) reveals a concerning disparity, with more rural counties tending to have fewer mental health providers compared to their urban counterparts. This finding aligns with broader concerns about the challenges of providing adequate healthcare services in rural and remote areas, necessitating targeted efforts to address these inequities.

Interestingly, the positive coefficient associated with median household income (0.0130) suggests a potential socio-economic component, with counties having higher median incomes exhibiting a greater density of mental health providers. This relationship may reflect the influence of socio-economic factors on the demand for and accessibility of mental health services, as well as the potential for resource allocation patterns to contribute to disparities in health care provision.

The intercept of the model (2.2935) represents the predicted number of mental health providers per capita when all the indicator variables are set to zero, providing a baseline against which the effect of the predictor variables can be interpreted.

This comprehensive analysis not only quantifies the relationships between various factors and mental health provider availability but also offers valuable insights for policymakers and healthcare planners. By identifying the significant predictors and their relative impacts, targeted interventions can be designed to address the disparities in mental health service provision, ultimately promoting more equitable access to these essential services across Pennsylvania's diverse communities.

# 6    Conclusions

Based on the Gini coefficient, there is health disparity in all variables that we choose to measure healthcare access. There high disparity in the distribution of hospital bed providers and the number of Hospitals. The number of hospitals strongly favors the lower income while Mental Health Provider strongly favors the higher income.

Our comparative analysis revealed a significant disparity in the distribution of mental health providers per 1000 population. The Gini index of 0.185 indicates a moderate level of inequity, favoring wealthier regions. Notably, the Atkinson index at an $\varepsilon$ value of 2 was the highest among all indicators, underscoring the need for immediate action in this area. Additionally, the weighted Utilitarian social welfare value of -9.8175 for mental health providers highlights a substantial loss in social welfare due to the uneven distribution. These findings suggest that focusing on the distribution of mental health providers in future research could help promote a more balanced and equitable healthcare system that supports both physical and mental health.

Based on the linear regression between mental health provider rate and suicide rate, there is a negative correlation between the number of mental health providers and suicide rates. This means the higher the mental health provider rate, the lower the suicide rate will be. The low $R^2$ value indicates that the number of mental health providers is not the sole factor influencing suicide rates, and other variables may be significant contributors to suicide rates.

The findings from this study have significant implications for policymakers and healthcare planners. By identifying the specific areas of healthcare access that exhibit the most pronounced inequalities and are most sensitive to societal preferences for reducing disparities, targeted interventions can be designed and implemented. Addressing the unequal distribution of mental health providers, hospital resources, and geographical accessibility should be prioritized to promote equitable access to improve overall population health outcomes.

Moreover, this research highlights the importance of considering socioeconomic factors and societal values in healthcare resource allocation. Policies and initiatives should aim to reduce barriers to access for disadvantaged communities, ensuring that essential healthcare services are available to all, regardless of income level or geographic location.

While this study focused on Pennsylvania, the methodological framework can be applied to other regions, allowing for a comprehensive assessment of healthcare access disparities and informing evidence-based policymaking. Continued research in this area is crucial to refine measurement approaches, evaluate the impact of interventions, and ultimately achieve the goal of equitable and accessible health care for all individuals.

In expanding our study, we aim to explore potential solutions for addressing the inequitable distribution of healthcare access indicators, particularly focusing on mental health providers and the average distance to the closest hospital. One key question is what effective strategies can be implemented to rectify these imbalances and ensure equitable access to healthcare services for all. Furthermore, we need to develop robust methods for assessing and predicting the effectiveness of these solutions before they are implemented, to maximize their impact and avoid unintended consequences. Another area of investigation will involve examining the underlying issues that contribute to unfavorable statistics in average distance to the closest hospital, such as transportation barriers and geographic disparities. By addressing these questions, we can contribute to the development of targeted policies and initiatives that promote a more equitable and accessible healthcare system for everyone.

# 7 References

[1] Luo, W. (2004). Using a GIS-based floating catchment method to assess areas with shortage of physicians. In *Health & Place, Vol. 10* (pp. 1-11).

[2] Caballo, B., Dey, S., Prabhu, P., Seal, B., Chu, P., & Kim, L. (2021). The Effects of Socioeconomic Status on the Quality and Accessibility of Healthcare Services.

[3] Cristia, J. (2009). Rising Mortality and Life Expectancy Differentials by Lifetime Earnings in the United States. In *Working Paper, No. 665*, Inter-American Development Bank, Washington, DC.

[4] Gravelle, H., Wildman, J., & Sutton, M. (2000). Income, Income Inequality and Health: What can we Learn from Aggregate Data? In *Discussion Papers in Economics, No. 2000/26*, Department of Economics and Related Studies, University of York, York.

[5] Lynch, J., Smith, G., Harper, S., Hillermeier, M., Ross, N., Kaplan, G., & Wolfson, M. (2004). Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. In *The Milbank Quarterly, Vol. 82, No. 1*, University of Michigan, University of Bristol, Pennsylvania State University (pp. 5-99).

[6] Waters, H. (2000). Measuring equity in access to health care. In *Social Science & Medicine, Vol. 51*, Sanger, Texas (pp. 599-612).

[7] Behavioral Health Services in Pennsylvania (2023) Centers for Disease Control and Prevention. Available at: `https://www.cdc.gov/childrensmentalhealth/stateprofiles-providers/pennsylvania/index.html`.

[8] Bureau, U.C. (2023) Current population survey (CPS), Census.gov. Available at: `https://www.census.gov/programs-surveys/cps.html`.

[9] Health Facilities (no date) Department of Health. Available at: `https://www.health.pa.gov/topics/HealthStatistics/HealthFacilities/Pages/health-facilities.aspx`.

[10] Pennsylvania Department of Human Services (DHS) (2020) Uninsured population census data 5-Year estimates for release years 2017-current County Human Services and Insurance: PA Open Data Portal, Commonwealth of Pennsylvania Open Data Portal. Available at: `https://data.pa.gov/Health/Uninsured-Population-Census-Data-5-year-estimates-/neqb-cw4e/about_data`.

[11] Hammond, P. (n.d.). Harsanyi's Utilitarian Theorem: A Simpler Proof and Some Ethical Connotations. Available at `https://web.stanford.edu/~hammond/HarsanyiFest.pdf#:~:text=Harsanyi%E2%80%99s%20utilitarian%20theorem%20states%20that%20the%20social%20welfare`

[12] Hasell, J., & Roser, M. (2023). Measuring inequality: What is the Gini coefficient? Retrieved from `https://ourworldindata.org/what-is-the-gini-coefficient#:~:text=The%20Gini%20coefficient%20captures%20how,the%20'line%20of%20equality'`.