

Research on Relationship Between Carbon Dioxide Emission and Mortality

Jiayi Zhang^{1,a,*}

¹*Chinese University of Hong Kong, Shenzhen, China*
a. 121010061@link.cuhk.edu.cn

**corresponding author*

Abstract: Carbon dioxide emissions have become an increasingly concerning issue, leading to a variety of complex and profound consequences that touch upon environmental, economic, and health aspects. This research delves into the correlation between carbon dioxide emissions and mortality rates within urban areas of China, emphasizing the significance of air quality on public health. By employing a multiple regression analysis and utilizing cross-sectional data from 2019 across 285 cities, the study scrutinizes the influence of CO₂ emissions, the extent of urban medical insurance enrollment, the accessibility of hospitals, and the combined effects of CO₂ emissions and hospital accessibility on mortality figures. The study's results underscore the imperative for additional investigative efforts to dissect the complex interplay among air pollution, accessibility to healthcare services, and death rates, particularly in the context of China's swift urban development and industrial growth. This is crucial for understanding the broader implications of environmental factors on health outcomes and for formulating effective public health policies.

Keywords: carbon dioxide emissions, mortality, multiple regression

1. Introduction

Climate change has become a pressing environmental concern, especially in China. The impact of air quality on human health is a critical public health topic with broad socioeconomic implications in such an increasingly polluted environment. China's rapid economic growth, industrialization, urbanization, and unique socioeconomic characteristics lead to rise in carbon dioxide (CO₂) emissions. These emissions contribute to climate change, exacerbating air pollution and leading to extreme weather events, worsened air quality, and increased health risks, including respiratory and cardiovascular diseases [1,2]. The most serious consequence of these issues is that they can lead to death. This highlights the need to investigate the relationship between CO₂ emissions and mortality in China.

Recent studies have shown that air pollution and CO₂ emission has plenty of adverse effects on human health. Klea (2003) pointed out the adverse effects of air pollution in Europe and North America, by conducting an epidemiological review [3]. Klea focused on studies involving time-series analysis, air quality monitoring, and cohort data to find out the short-term and long-term impacts of different pollutants. The World Health Organization emphasized that CO₂-induced climate change intensifies these risks by increasing the prevalence of heat-related illnesses and worsening air quality [1]. West et al. also demonstrated that mitigating global greenhouse gas emissions, including CO₂,

could significantly improve future air quality and reduce health risks [4]. They estimated that by 2030, several hundred thousand premature deaths related to air pollution could be avoided annually through effective emissions reduction.

Some existing literature also directly investigates the relationship between air quality and death rate. For instance, by studying the relationship between severe air pollution episodes in Europe and North America, David has found that high level of air pollution might be harmful to the human body and may increase mortality [5]. Wei et al. assessed the causal effects of air pollution on mortality rates in Massachusetts by employing a generalized propensity score (GPS) approach [6]. Wiesenthal et al. also determined the correlation between mortality and air pollution by conducting a cohort study using the Can CHEC database, involving over 7 million adults and 128 million person-years of follow-up data [7].

The researches directly linking CO₂ emissions to number of deaths, particularly through large-scale regression analysis tailored to China's unique socioeconomic and environmental context, remain scarce. Large-scale regression analysis enables the isolation of the impact of air quality on fatal diseases by accounting for specific Chinese socioeconomic confounders, including urban medical insurance and hospital availability. This method provides a more precise and actionable understanding of how air quality could influence number of deaths in China.

Therefore, this research employs regression methods to China's unique environmental context. According to the existing epidemiological and economic theories, it could be predicted that CO₂ emission poses a health risk through their contribution to climate change. Therefore, we hypothesize that high CO₂ emission will correlate with large number of deaths.

2. Method

2.1. Data source

The data for this study is obtained from multiple authoritative sources to ensure its accuracy and relevance. The primary sources include the State Statistics Bureau and Emissions Database for Global Atmospheric Research. These sources provide comprehensive data on CO₂ emissions across Chinese cities, which is crucial for understanding the relationship between CO₂ emissions and mortality.

2.2. Descriptive Statistics

This study employs a multiple regression model to analyze the relationship between CO₂ emissions and mortality across Chinese cities using 2019 cross-sectional data. Data for 41 cities were excluded due to incomplete datasets, resulting in a final sample of 285 cities. The dependent variable in this study is the number of deaths, while the key independent variables include CO₂ emissions (measured in metric tons per year), number of participants in urban basic medical insurance (in number of people) and number of hospitals.

The following table presents the descriptive statistics for the variables used in the analysis, including the mean and variance for each key variable, such as the number of deaths, annual CO₂ emissions, the number of participants in urban basic medical insurance, and the number of hospitals.

Table 1: Variables used in the analysis

	Mean	Std. dev
Number of Deaths	5361.54	22672.74
Annual CO ₂ emissions	3.43e7	3.66e7
Number of participants in insurance	1066231.	1939205
Number of hospitals	113.7003	106.182

2.3. Multiple regression model

The baseline regression model is specified as follows.

$$death = \beta_0 + \beta_1 * Annual\ CO_2\ emissions + \beta_2 * urban_medical_insurance + \beta_3 * hospital + \beta_4 * interaction_term + u_i \quad (1)$$

The model components are summarized in table 2.

Table 2: Model components

Dependent variable	
Signal	Implication
<i>Death</i>	The total number of deaths in city ii
Independent variable	
<i>Annual CO₂ emissions</i>	Annual CO ₂ emissions (metric tons per year) for each city
<i>urban_medical_insurance</i>	Number of participants in urban basic medical insurance in city ii
<i>hospital</i>	Total number of hospitals in city ii
<i>interaction_term</i>	An interaction term capturing how CO ₂ emissions and hospital availability jointly affect mortality
Variable Interpretations	
β_1	Measures the effect of a 1-metric-ton increase in CO ₂ emissions on the number of deaths
β_2	Reflects the change in number of death due to an additional participant in urban medical insurance
β_3	Captures the impact of an additional hospital on number of death
β_4	Explains the interaction effect between CO ₂ emissions and hospital availability

The key assumptions of the Molde can be seen in table 3.

Table 3: Key assumptions

No.	Items	Implication
1	Linearity	The relationship between the dependent and independent variables is linear.
2	Random Sampling	The data represents a random sample of cities.
3	No Omitted Variables	All key variables influencing mortality are included.
4	No Multicollinearity	Independent variables are not strongly correlated.
5	Homoscedasticity	The variance of residuals is constant across all levels of independent variables.

2.4. Addressing Statistical Challenges

To ensure unbiased and consistent parameter estimates, diagnostic tests were conducted. The variance inflation factor (VIF) test was used to assess multicollinearity, and the White test was applied to check for heteroskedasticity. In case of heteroskedasticity, robust standard errors were employed.

3. Results and discussion

3.1. Model fit statistics

The regression model's fit statistics are presented in table 4. The overall model explains 61.24% of the variation in the number of deaths $R^2 = 0.6124$), with an adjusted R^2 of 0.6070. The F-statistic is 111.41, with a corresponding p-value of 0.0000, indicating that the model as a whole is statistically significant.

Table 4: Fit statistics

Number of observations	287
F-statistic (4, 282)	111.41
Prob > F	0.0000 (statistically significant)
R-squared	0.6124
Adjusted R-squared	0.6070

3.2. Empirical Results

The results of the regression model are presented in Table 5. The dependent variable is the number of deaths, and the independent variables include CO₂ emissions, urban medical insurance, hospital availability, and an interaction term between CO₂ emissions and hospitals. The regression output reveals the following key results.

The coefficient for CO₂ emissions is -0.0001146, which suggests that holding all other factors constant, a one-unit increase in CO₂ emissions is associated with a decrease of approximately 0.0001146 deaths. The result is statistically significant with a p-value of 0.002. This suggests a negative relationship between CO₂ emissions and the number of deaths, indicating that higher CO₂ emissions are associated with fewer deaths. The 95% confidence interval for this coefficient ranges from -0.0001876 to -0.0000415, confirming that this relationship is robust.

The coefficient for urban medical insurance is -0.0008657, but with a p-value of 0.147, it is not statistically significant at the 5% level. This suggests that urban medical insurance may not have a direct impact on the number of deaths in this model, though this result should be interpreted cautiously due to the high p-value.

The coefficient for the number of hospitals is 94.51029, which indicates that holding other variables constant, an increase of one hospital is associated with an increase of approximately 94.51 deaths. with a very small p-value (0.000), indicating that the number of hospitals has a strong positive effect on the number of deaths. This suggests that more hospitals are associated with more deaths, which could be due to increased accessibility to medical care or other factors.

The interaction term has a coefficient of 1.01e-06, and it is statistically significant (p-value = 0.000). This suggests that the impact of CO₂ emissions on deaths decreases as the number of hospitals increases. In other words, in cities with more hospitals, the effect of air pollution on mortality is somewhat moderated.

The constant term is 13,963.56, with a p-value of 0.000, indicating that the baseline number of deaths, when all other variables are held constant, is approximately 13,963.56.

Table 5: Results of the regression model

Variable	Coefficient	Robust Standard error	t-value	P value	95% confidence interval
CO ₂ emissions	-0.0001146	0.0000371	-3.09	0.002	[-0.0001876, -0.0000415]
Urban Medical Insurance	-0.0008657	0.0005954	-1.45	0.147	[-0.0020377, 0.0003062]
Hospitals	94.51029	14.01268	6.74	0.000	[66.92756, 122.093]
Interaction	1.01e-06	1.56e-07	6.46	0.000	[7.00e-07, 1.31e-06]
Constant	13,963.56	1,744.063	8.01	0.000	[10,530.53, 17,396.6]

3.3. Complex relationship between CO₂ emission and urban health quality

The negative relationship between CO₂ emissions and deaths is somewhat unexpected, as one would typically assume that higher pollution levels would lead to more deaths. However, the effect size is very small, and given the magnitude of the coefficient, it suggests that the direct impact of CO₂ emissions on mortality might be relatively weak. Alternatively, the negative sign may reflect a counteracting effect from other factors not captured in the model. For instance, cities with higher CO₂ emissions are often more economically developed and have better healthcare infrastructure. In such cities, higher economic development and better medical conditions might contribute to better living standard and longer average life expectancies. These cities could have more resources to mitigate the health risks associated with pollution, resulting in a smaller overall impact on deaths. However, those factors such as living standard, ability to access to quality health care and social status which may reduce mortality are difficult to quantify or measure in this study. This may consequently lead to the fact that those factors which play a crucial role in explaining the lower number of deaths may are not explicitly capture in this model.

Furthermore, it is important to consider that Chinese governments have implemented stringent CO₂ emission control policies, which have reduced the overall emission levels in recent years. For instance, The Energy Production and Consumption Revolution Strategy, adopted in 2019, focuses on promoting cleaner energy sources and improving energy efficiency across industries in China [7]. Additionally, the Air Pollution Prevention and Control Action Plan, launched in 2013 and reinforced in subsequent years, has targeted a reduction in air pollution, which indirectly curtails CO₂ emissions in China [8]. Therefore, the current levels of CO₂ emissions in the dataset might not be high enough to pose a significant threat to public health.

To assess multicollinearity, Variance Inflation Factors (VIF) were calculated for the explanatory variables in the model. The VIF values for the variables were found to be below 10, with the highest being 5.51 and mean 3.26 for the interaction term, suggesting that there are moderate multicollinearity which is not a concern in the model. Therefore, the model can be considered reliable for further analysis.

To test for the potential presence of omitted variables in the model, we conducted the Ramsey RESET test. The test results show a significant $F(3, 279) = 9.09$ with a p-value of 0.0000, which leads us to reject the null hypothesis of no omitted variables. This suggest that there might exist an omitted variable bias in this model which may be main cause of negative relation between CO₂ emission and

number of deaths. Therefore, further model adjustments may be necessary to account for potential omitted variable bias.

White test has also been conducted to check for Heteroskedastic. The test results indicate a significant presence of heteroskedasticity in the model. Robust standard errors are therefore be used to correct for the potential bias in the estimation of the standard errors. A Q-Q plot is used to verify the normality of the residuals. As shown in figure 1, the plot shows that most of the points lie close to the 45-degree reference line, suggesting that the residuals are approximately normally distributed.

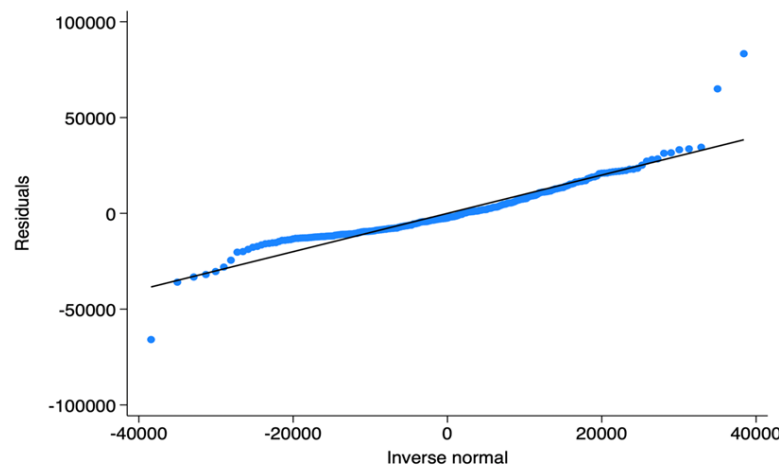


Figure 1: Q-Q plot of the modle.

4. Conclusion

In this study, the relationship between air quality and number of deaths in China is analyzed with regression analysis. Data used in this study is from 285 different cities in 2019. The analysis revealed a negative, albeit weak relationship between CO₂ emissions and number deaths. This suggests that while CO₂ emissions do not seem to have a direct, large-scale impact on number of deaths, other factors such as regional economic development, healthcare access, and air pollution may play more significant roles. Future studies could address these limitations by incorporating the data of a year while there are large amount of co2 emission instead of using the 2019's data where government already start to control the co2 emission. Besides, future study should explore the specific health impacts of air pollution, and accounting for omitted variables that may confound the results. With more comprehensive data, a more nuanced understanding of the dynamics between air quality and mortality rates could be achieved.

References

- [1] West, J. J., Smith, S. J., Silva, R. A., Naik, V., Zhang, Y., Adelman, Z., ... & Lamarque, J. F. (2013). Co-benefits of mitigating global greenhouse gas emissions for future air quality and human health. *Nature Climate Change*, 3, 885-889.
- [2] World Health Organization. (2003). *Climate Change and Human Health: Risks and Responses*. <https://www.who.int/publications/i/item/924156248x>
- [3] Weichenthal S., Pinault, Christidis, T, Burnett R. T., Brook J. R., Chu, Weagle c. L. (2022). How low can you go? Air pollution affects mortality at very low levels. *Science Advances*, 8, 1-9.
- [4] Davis D. L. (2002). *When Smoke Ran Like Water*. New York: Basic Books.
- [5] Wei S., Li X. Liu, Y, & Smith,S. (2023). Causal effects of air pollution on. mortality rate in Massachusetts. *American Journal of Epidemiology*, 189, 1316-1323.
- [6] Katsouyanni, K. (2003). Ambient air pollution and health. *British Medical Bulletin*, 68, 143-156.
- [7] National Development and Reform Commission (NDRC) of China. (2016). *Energy Production and Consumption Revolution Strategy (2016-2030)*

- [8] *State Council of the People's Republic of China (2013). Air Pollution Prevention and Control Action Plan.*