Impacts of air pollution on investors' moods*

Replicating 'Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide'

Zhongyu Huang

April 27, 2022

Abstract

The silent killer, air pollution, can cause not only health issues but psychological and economic effects. This paper replicates the 2021 article 'Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide,' and uses the results as evidence to examine the non-health effects of air pollution. Models in the article have shown that the daily stock returns decrease by 1.2% on average globally as PM2.5 concentration increases by 10 microgram per cubic metre. This analysis provides substantial evidence that air pollution might affect human behavior.

Keywords: Environmental Economics, Ambient Particular Air Pollution, Stock Price, PM2.5

1 Introduction

In recent years, extreme pollution events bring people's attention to studying the impacts due to air pollution. Exposure to pollutants such as nitrogen oxides (NOX) or sulfur oxides (SOX) can cause serious health issues including damaging lung function(Cardenas 2021). As the main component of smog, fine particulate matter and ground-level ozone (O3) can cause dysfunction of the eye, nose, throat, or lung(Cardenas 2021). Fine particulate matter has been identified as one of the risk factors for cardiovascular disease and premature death(Cardenas 2021). Other pollutants such as Carbon monoxide (CO) and Ammonia (NH3) are also harmful to human health(Cardenas 2021). Most of them are colorless gas to be not noticed. The vulnerable group, children and the elderly, can be harmed more easily at greater risk. In addition, indirect evidence was found to explain how air pollution influences mood and human cognitive functions.

The article 'Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide' by Simo Kiihamäki, Marko Korhonen, and Jouni Jaakkola investigates how daily concentrations of fine particulate matter influence daily stock market returns and volatility(Simo-Pekka Kiihamäki 2021a). They aim to test two hypotheses. The first is that the increases in the daily PM2.5 concentrations decrease daily stock returns(Simo-Pekka Kiihamäki 2021a). The second is that the increases in daily PM2.5 concentrations increase daily stock volatility(Simo-Pekka Kiihamäki 2021a). Linear models and meta-regression analysis are performed to generate results at the global level. Evidence from regression models for each city supports that short-term exposure to PM2.5 could lead to a decrease in stock returns(Simo-Pekka Kiihamäki 2021a).

This paper use R (R Core Team 2020) replicates the original analysis in the article by using the dataset provided by the authors. Data related to air pollution data, stock index, and meteorology are provided in the dataset. The dataset will be introduced and explained in more detail first in the Data section. Their methods will be explored in the Results section. Regression models created in the original analysis will be replicated while additional models are produced and explored. Results and discussion will be presented in the end.

*Code and data are available at: https://github.com/zhongyuhuang/Impacts-of-air-pollution-on-investors-moods.git

2 Data

2.1 Data source and collection

This paper uses the datasets conducted by Kiihamäki etc. The datasets are available as CSV files from https://data.mendeley.com/datasets/z8t3s8btxv/1. They store air pollution data, stock index data, and meteorological data for 47759 observations from 47 cities worldwide(Simo-Pekka Kiihamäki 2021b). Daily returns of stocks are collected from investing.com(Simo-Pekka Kiihamäki 2021b). This website is a financial markets platform offering access to real-time data and free financial tools across 250 exchanges around the world(Krishnan 2007). However, it is worthy to notice that this platform claims that data provided by them is not necessarily by any market or exchange, hence not necessarily real-time nor accurate(Krishnan 2007). Data related to air pollution are collected from various sources including the European Environment Agency, the U.S. Department of State, and so on(Simo-Pekka Kiihamäki 2021b). Only 47 out of 88 countries have sufficient PM2.5 concentration to be involved in the original analysis. Meteorological data for these countries are obtained from the Global Surface Summary of the Day(Simo-Pekka Kiihamäki 2021b). It is a product of the National Oceanic and Atmosphere Administration and it possesses averages of daily weather elements computed from global hourly station data(NOAA 2022).

2.2 Missing Values

The authors used two technics to deal with missing values in datasets. As for the missing values that exist in either stock data or air pollution data, the authors choose to omit observations from all day if any of the stock returns or PM2.5 concentration is missing(Simo-Pekka Kiihamäki 2021a). This decision is made based on that it is impractical to find relations in the absence of these data. As for meteorological variables, there exist a large amount of randomly missing data(Simo-Pekka Kiihamäki 2021a). These variables are served as confounding factors that affect the research question in the original paper. The authors decide to impute these missing values to perform adjustments due to confounders by using classification and regression training(Simo-Pekka Kiihamäki 2021a).

2.3 Visualization of variables

Most of the observations are from London, Toronto, New York City, Copenhagen, and Helsinki.(Figure 1). They are almost equally distributed each day of the week. The dependent variables here is the logarithmic daily returns. The daily return of stock indicates the dollar change in a stock's price as a percentage of the previous day's closing price. We use the logarithmic daily returns of stock instead the daily return to ensure the normality of the dependent variable. From the plots of explanatory variables, a significant amount of outliers are presented(Figure 2). It is possible to have problematic observations in the dataset which could affect the further models. Categorical Variables would be transformed to factors in further analysis.

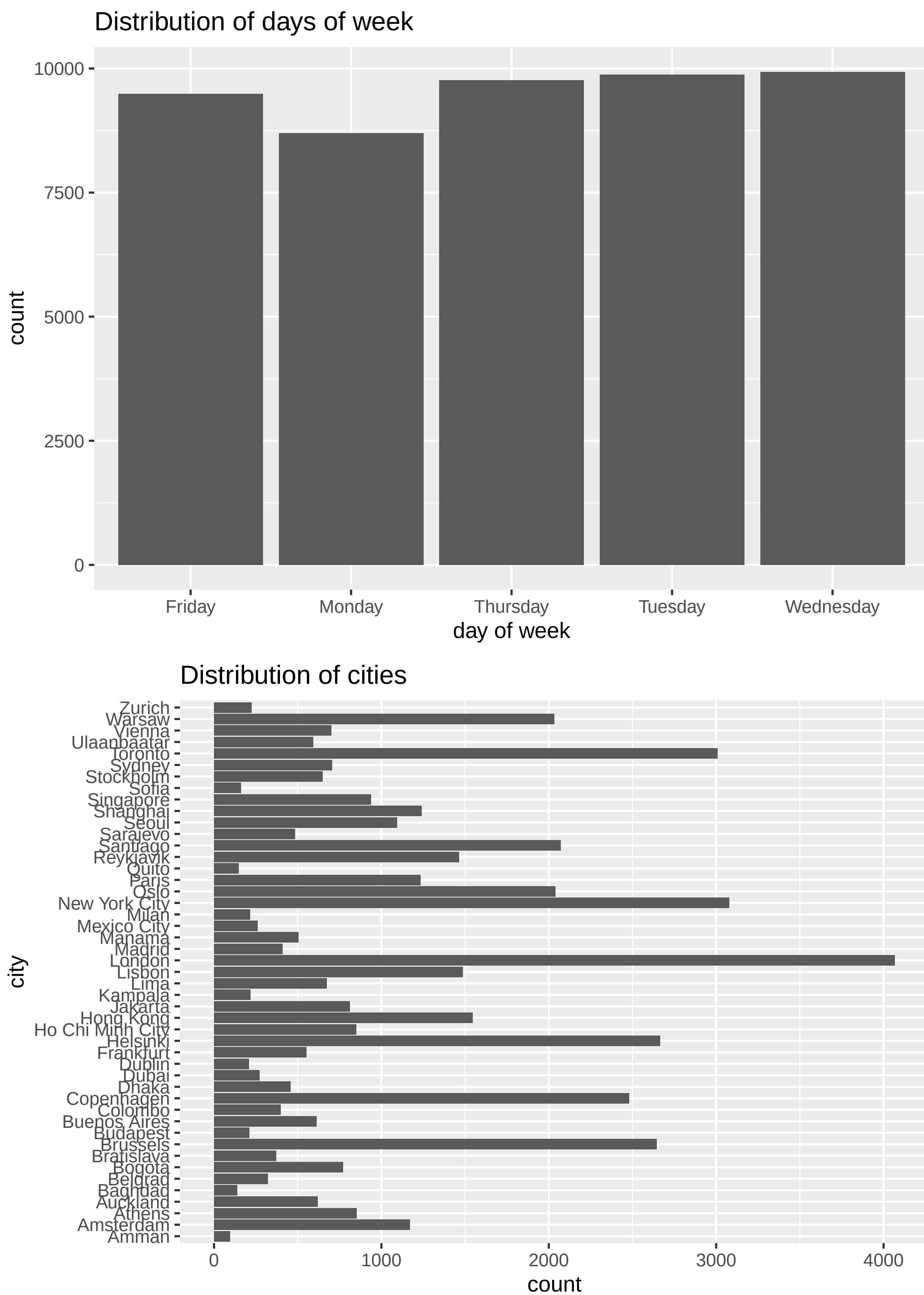


Figure 1: Cities and days of week of observations

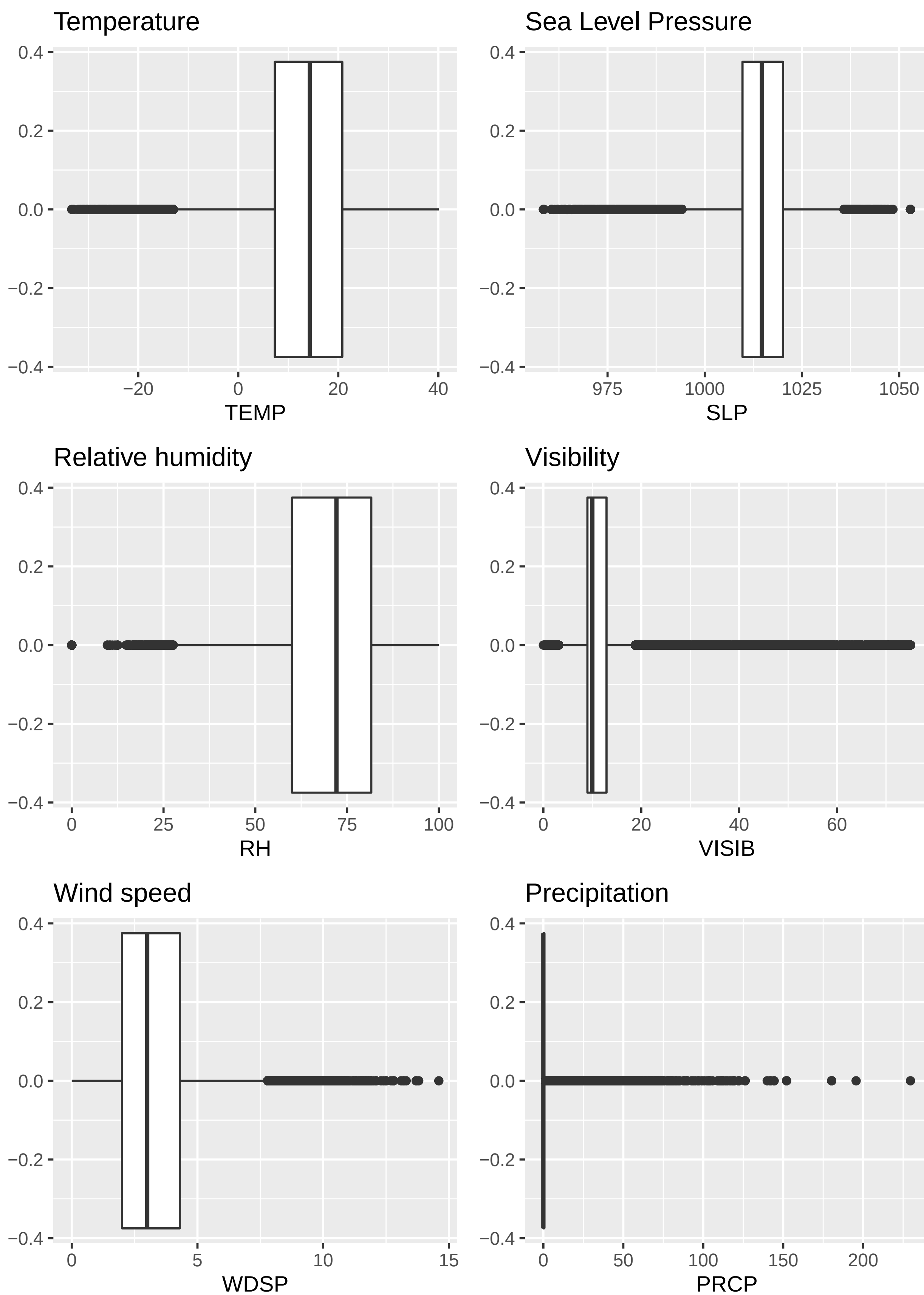


Figure 2: Visualization of explanatory variables

3 Model

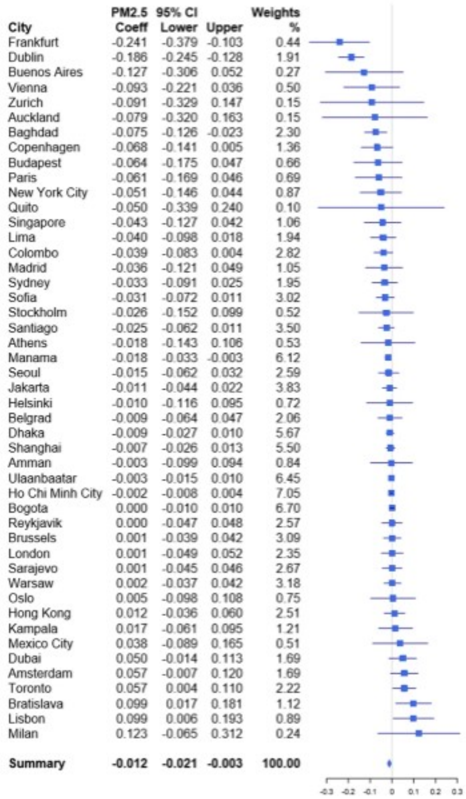
In this paper, only the stock return model has been replicated. Here, the stock return is defined as: - $RET_t = (ln(P_t) - ln(P_t - 1) * 100$, where Pt is the closing price of the city's corresponding stock index on day t. The complete model is:

- $RET_t = \beta_0 + \beta_1 * PM2.5_t + \beta_2 * RET_t - 1 + \beta_3 * RET_t - 2 + \beta_4 * woy_t + \beta_5 * dow_t + \beta_6 * PRCP_t + \beta_7 * WDSP_t + \beta_8 * VISIB_t + \beta_9 * SLP_t + \beta_1 0 * TEMPbin_t + \beta_1 1 * DEWPbin_t + \beta_1 2 * AD + \epsilon$

Woy and dow representing week of year and day of week are included as dummy variables in the model. 1-day and 2-day lagged returns are fitted in the model to account for autocorrelation in the stock return. PRCP denotes precipitation. WDSP denotes wind speed. SLP denotes sea level pressure. VISIB denotes visibility. SAD denotes the daylength in hours. They are all treated as continuous variables in the model. TEMP_bins denotes temperature and it is treated as dummy bin. Same for dew point (DEWP_bins). ϵ is the Newey-West standard errors. Traffic emissions are considered as confounding factors here. Daily NO2 concentrations are fitted into the model to perform adjustments.

4 Results

Regression models for each city are performed individually. The estimated coefficients with their 95% confidence intervals are summarized(Figure 4). We can see the effects varies among cities from -0.241 to 0.123. Frankfurt has the largest negative effect of PM2.5 on stock returns. Milan has the largest positive effect of PM2.5 on stock returns. The summary effect estimate for PM2.5 based on all 47 city is -0.012. This number indicates that a 10 microgram per cubic metre increase in concentration of PM2.5 leads to an average 1.2% reduction in daily stock returns. It provides statistically significant evidence to not reject hypothesis 1. That is, exposure to PM2.5 would lead to a decrease in stock returns. Traffic emissions are treated as confounding factors here. In absence of direct volumes of traffic emissions, concentrations of NO2 are used here. However, the effect estimates are not significantly influenced in models after fitting the data of NO2(Simo-Pekka Kiihamäki 2021a).



City	PM2.5 Coeff	95% CI Lower	Upper	Weights %
Frankfurt	-0.241	-0.379	-0.103	0.44
Dublin	-0.186	-0.245	-0.128	1.91
Buenos Aires	-0.127	-0.306	0.052	0.27
Vienna	-0.093	-0.221	0.036	0.50
Zurich	-0.091	-0.329	0.147	0.15
Auckland	-0.079	-0.320	0.163	0.15
Baghdad	-0.075	-0.126	-0.023	2.30
Copenhagen	-0.068	-0.141	0.005	1.36
Budapest	-0.064	-0.175	0.047	0.66
Paris	-0.061	-0.169	0.046	0.69
New York City	-0.051	-0.146	0.044	0.87
Quito	-0.050	-0.339	0.240	0.10
Singapore	-0.043	-0.127	0.042	1.06
Lima	-0.040	-0.098	0.018	1.94
Colombo	-0.039	-0.083	0.004	2.82
Madrid	-0.036	-0.121	0.049	1.05
Sydney	-0.033	-0.091	0.025	1.95
Sofia	-0.031	-0.072	0.011	3.02
Stockholm	-0.026	-0.152	0.099	0.52
Santiago	-0.025	-0.062	0.011	3.50
Athens	-0.018	-0.143	0.106	0.53
Manama	-0.018	-0.033	-0.003	6.12
Seoul	-0.015	-0.062	0.032	2.59
Jakarta	-0.011	-0.044	0.022	3.83
Helsinki	-0.010	-0.116	0.095	0.72
Belgrad	-0.009	-0.064	0.047	2.06
Dhaka	-0.009	-0.027	0.010	5.67
Shanghai	-0.007	-0.026	0.013	5.50
Amman	-0.003	-0.099	0.094	0.84
Ulaanbaatar	-0.003	-0.015	0.010	6.45
Ho Chi Minh City	-0.002	-0.008	0.004	7.05
Bogota	0.000	-0.010	0.010	6.70
Reykjavik	0.000	-0.047	0.048	2.57
Brussels	0.001	-0.039	0.042	3.09
London	0.001	-0.049	0.052	2.35
Sarajevo	0.001	-0.045	0.046	2.67
Warsaw	0.002	-0.037	0.042	3.18
Oslo	0.005	-0.098	0.108	0.75
Hong Kong	0.012	-0.036	0.060	2.51
Kampala	0.017	-0.061	0.095	1.21
Mexico City	0.038	-0.089	0.165	0.51
Dubai	0.050	-0.014	0.113	1.69
Amsterdam	0.057	-0.007	0.120	1.69
Toronto	0.057	0.004	0.110	2.22
Bratislava	0.099	0.017	0.181	1.12
Lisbon	0.099	0.006	0.193	0.89
Milan	0.123	-0.065	0.312	0.24
Summary	**-0.012**	**-0.021**	**-0.003**	**100.00**

\begin{figure} \caption{Regression coefficients and 95% confidence intervals for 47 cities} \end{figure}

5 Discussion

5.1 Cleaner air indicates higher stock returns?

It might be possible for stock returns to be higher when there is no pollution. However, we cannot make such a statement without further investigation. The results shown in the analysis are not necessarily saying "Let's rise the stock return by cleaning the air." Indeed, they indicate that we will see more accurate stock returns and volatility with cleaner air. In this case, the volatility will follow the market fundamentals. A strong relation between air pollution and stock volatility actually indicates the inefficiency of the market.

5.2 Other non-health outcomes

Often we could hear discussions on how human behaviors produce or reduce air pollution. It is rare to talk about this topic in the opposite perspective, especially on psychiatric aspect. Only a few researches had examined these effects linked to air pollution. However, evidence had been found to suggest that air pollutants interfere with functioning of the nervous system(Gary W. Evans 1981). Other than that, although the ability to withstand anxiety depends on individual, medical effects of air pollution can indirectly affect psychological health(Gary W. Evans 1981). Research from 2017 had found the relationship between ambient air pollution and increased aggressive behavior in teenagers(Health 2017).

5.3 Weaknesses and next steps

Although we found a link between the daily concentration of PM2.5 and the daily stock return, we did not prove the causality between them. Other than that, there exists confounding factors such as other pollutants can effect the models and results established in the analysis. Whether the effect of PM2.5 on stock volatility is temporary or permanent is not examined. As described in the previous section, a significant amount of metro data are estimated. All of the results are made under potential bias including data error. More accurate and informative data are required to make a stonger conclution.

Appendix

.1 Model Card

- Model Details. Basic information about the model.

  - Person or organization developing model

    Simo-Pekka Kiihamäki, Marko Korhonen & Jouni J. K. Jaakkola

  - Model date

    21 April 2021

  - Model version

    2

  - Model type

    Linear regression model

  - Paper or other resource for more information

    "Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide"

  - Citation details

    Kiihamäki, SP., Korhonen, M. & Jaakkola, J.J.K. Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide. Sci Rep 11, 8628 (2021). https://doi.org/10.1038/s41598-021-88041-w

  - License

    Creative Commons Attribution 4.0 International License

  - Where to send questions or comments about the model

    contact original authors of the paper "Ambient particulate air pollution and daily stock market returns and volatility in 47 cities worldwide."

- Intended Use. Use cases that were envisioned during development.

  - Primary intended uses

    used to examine relationships between concentration of PM2.5 and dailt stock returns for 47 cities.

  - Primary intended users

    Simo-Pekka Kiihamäki, Marko Korhonen & Jouni J. K. Jaakkola

  - Out-of-scope use cases

    Not applicable

- Factors.

  - Relevant factors

    The foreseeable salient factors for which model performance may vary includes the accuracy of the dataset.

  - Evaluation factors

    Traffic emissions are treated as confounding factors to adjust the model.

- Metrics. Metrics should be chosen to reflect potential realworld impacts of the model.

  - Model performance measures

    Not applicable

  - Decision thresholds

    Not applicable

  - Variation approaches

    Approaches including confidence intervals

- Evaluation Data. Details on the dataset(s) used for the quantitative analyses in the card.

  - Datasets

    Datasets used to evaluate the model were produced by the original authors,

  - Motivation

    These datasets obtain air pollution data, stock returns data, and meteorological data which are needed to run the model.

  - Preprocessing

    Omitted and imputed missing values.

- Training Data. May not be possible to provide in practice. When possible, this section should mirror Evaluation Data. If such detail is not possible, minimal allowable information should be provided here, such as details of the distribution over various factors in the training datasets.

Not applicable

- Quantitative Analyses

  - Unitary results

    The summary effect estimate for PM2.5 based on 47 city specific estimates is -0.012 (95% confidence interval -0.021, -0.003), which is statistically significant at the 5% level.

  - Intersectional results

    Not applicable

- Ethical Considerations

The model did not use any sensitive data. The model did not intended to inform decisions about matters central to human life or flourishing. No risk mitigation strategies were used during model development. Risks present in model usage remain unknown.

- Caveats and Recommendations

The ideal characteristics of an evaluation dataset for this model would be complete data for air pollution, stock returns, and meteorology from reliable sources.

References

Cardenas, Shirley. 2021. *Air Pollution: The Silent Killer Called Pm2.5.* https://www.mcgill.ca/newsroom/channels/news/air-pollution-silent-killer-called-pm25-329428.

Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean 'Lahman' Baseball Database.* https://CRAN.R-project.org/package=Lahman.

Gary W. Evans, Stephen V. Jacobs. 1981. "Air Pollution and Human Behavior." *Social Issues.* https://doi.org/10.1111/j.1540-4560.1981.tb01059.x.

Health, USC Environmental. 2017. "Increased Air Pollution Linked to Aggressive Behavior in Teens." https://envhealthcenters.usc.edu/2017/12/air-pollution-linked-to-bad-teenage-behavior.html.

Krishnan, Barani. 2007. *About Investing.com.* https://www.investing.com/about-us/.

NOAA. 2022. *Global Surface Summary of the Day - GSOD.* https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Simo-Pekka Kiihamäki, Jouni J. K. Jaakkola, Marko Korhonen. 2021a. "Ambient Particulate Air Pollution and Daily Stock Market Returns and Volatility in 47 Cities Worldwide." *Sci Rep* 11: 8628. https://doi.org/10.1038/s41598-021-88041-w.

———. 2021b. "Ambient Particulate Air Pollution and Daily Stock Market Returns and Volatility in 47 Cities Worldwide." https://doi.org/10.17632/z8t3s8btxv.1.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.