Thesis

Social inequality of air pollution exposure in the Netherlands

Utrecht University

Faculty of Geosciences
Applied Data Science

Supervisors:
Simon Scheider
Tabea Sonnenschein

Student:
Sjors de Reu
(7027338)

2021/2022

Abstract

It is known that certain sub-populations with a lower socioeconomic status are exposed to higher than average levels of air pollution. This study explores whether these associations exist in the urban areas of the Netherlands, by investigating the existence of inequality in NO_2, $PM_{2.5}$ and PM_{10} across the socioeconomic factors income, ethnicity and gender. Having a understanding of this inequality can help policy makers to not let certain sub-populations be disproportionately exposed to air pollution. Using ElasticNet regression and an ANOVA test, it is found that the cities Rotterdam and The Hague show low income and high non-western ethnicity postcodes are considerably more exposed to air pollution. The reverse is happening for the capital Amsterdam, where high income and low non-western ethnicity postcodes are more exposed to air pollution. As for gender, the % female variable has a negative coefficient for all the pollutants, indicating that female dense postcodes could be exposed to lower levels of air pollution.

Contents

List of Tables

List of Figures

1 Introduction

Inequality of exposure to air pollution is a concept that derives from a much more known concept: 'health inequality' (*or health inequity*). The World Health Organization (WHO) [1] concludes that there is substantial evidence that socioeconomic factors such as education, employment status, income level, gender, and ethnicity have a significant impact on one's health. There are large variations in the health status of different social groups in all countries, whether low-, middle-, or high-income. The lower one's socioeconomic status (SES), the greater the likelihood of bad health.

One factor that contributes to this health inequality concept is air pollution. Air pollution can cause a variety of diseases including chronic obstructive pulmonary disease, stroke, trachea, bronchus, lung malignancies, lower respiratory infections and worsened asthma in both short and long-term exposure [2]. Exposure to air pollution has also been linked to obesity, dementia , type 2 diabetes, Alzheimer's disease and systemic inflammation, according to the WHO [3]. Moreover, air pollution, namely $PM_{2.5}$, has been identified as a main cause of cancer by the International Agency for Research on Cancer (IARC) [4] and chronic exposure can impact every organ in the body, worsening and intensifying existing health disorders, according to a global study [2], [5]. The Lancent Planetary Health [6] found that 92% of the deaths related to air pollution occur in low-income or middle-income countries(LMICs). Not only the deaths but also the economic burden that comes with air pollution occurs in LMICs. High income countries did well eliminating the worst forms of air pollution, while LMICs struggle to make air pollution a priority. Nonetheless, inequality of air pollution exposure still exist also in high income countries.

Altogether, the health consequences that evolve from air pollution are quite serious. Several studies found strong evidence that there is an inequality of air pollution exposure between socioeconomic groups. Low SES communities and communities of color are disproportionately exposed to air pollution [7]. Meaning that the health consequences of air pollution are not equally distributed across the population and thus putting low SES people in an even more disadvantageous position. The exposure to air pollution naturally exists based on where a person lives. Living in urban areas with lots of traffic, industry and agriculture will expose a person in most cases to more air pollution than living in a rural area. The place/neighborhood of living is however not always a choice and is constrained by someones socioeconomic status [8]. Ultimately, someones income decides whether they would be able to afford a nice house in a green neighborhood or a cheap apartment next to a polluting road. In urban areas, this is a far larger problem due to the scarcity of housing and density of polluting sources.

Air pollution inequality research has a lengthy history in the North America, but not so much in Europe [9]. Research to inequality of air pollution exposure is relevant to tackle the environmental justice. By understandings the exposure variations spatially and between sub-populations, environmental health policy makers could seek to reduce the population average risk, but more importantly don't let certain sub-populations be disproportionately exposed to air pollution compared the the overall population [10]. This research analyzes the inequality of air pollution exposure in the Netherlands. The three biggest cities of the so-called 'randstad' region will be considered (i.e., Amsterdam, Rotterdam and The Hague). These cities are interesting case studies as they consist of a great diversity in individuals with different socioeconomic statuses. Previous research on this topic tend to use low resolution air pollution maps and demographic statistics on large spatial districts, which can introduce considerable under or over estimation of effects. This research tries get closer to individual level exposure by using high resolution air pollution maps and having reliable demographic statistics on postcode level. In the Netherlands, 6-character postcode level consist of small areas and the average inhabitants within each postcode is ± 50 for the chosen cities. With this level of detail, it is hoped to have a better understanding of the true social inequality of air pollution exposure. The following research question is examined: "What is the relationship between socioeconomic factors and air pollution exposure in the urban areas of the Netherlands?"

To answer this question an ElasticNet regression analysis is performed on the socioeconomic variables income, gender and ethnicity in relation to the air pollutants NO_2, $PM_{2.5}$ and PM_{10} to find a possible inequality of exposure to air pollution. Furthermore, an ANOVA test is used for the income variable with respect to air pollution.

2 Literature review

Multiple studies on air pollution exposure inequality have been done. Many of them suggest there is relationship between different socioeconomic characteristics and air pollution exposure. To further understand this concept of air pollution exposure inequality, a literature review is needed. Air pollution exposure inequality can be measured with various methods and on various spatial scales.

As one of the biggest polluting country in the world, one would expect China to have a inequality problem as well. Different studies [11], [12] found that in China, air pollution levels are strongly correlated to the individual income. The inequality is seen across the whole country but becomes greater in the rural areas. Urban areas suffer mostly from indirect exposure of emissions, while rural areas are exposed to direct emissions from solid fuel combustion. The first study [11] utilized a hierarchical regression model to cover both micro level (individual) as well as macro level. A hierarchical regression model is possibly a more suitable option with the nested structure that exist in demographic data rather than the conventional linear regression which assumes all cases are independent of each other. The second study [12] uses the GINI coefficient together with descriptive statistics.

In North America multiple studies assessed the social inequality of air pollution exposure. A study [13] in Massachusetts (U.S.) analyzed the relationship between socioeconomic groups and ambient air pollution over a 8 year period. The greatest inequality has been found in the urban areas with the ethnic groups (e.g, White, Hispanic, Black) rather than the income and educational characteristics. To investigate inequality they used the Atkinson index and descriptive statistics. A study [10] in California's South Coast found that lower-income households, non-white and people that live in high population density areas are more exposed than the average. They only made use of descriptive statistics to analyze the inequality. Another study [14] examines inequality both on neighborhood level as well on individual level. By comparing group means and using a Intrinsic Conditional Auto-Regressive (ICAR) model, it found that the in the overall population, the higher the socioeconomic status, the lower the concentrations of NO_2, $PM_{2.5}$. In contrast to New York where the reverse is true. Moreover, these associations were stronger on neighborhood level than it was for individual level.

An Europe wide study [15] on this topic compared population weighted average PM_{10} concentrations with the mean household income. Large differences in air quality between Eastern- and Western Europe were found, although associations between low-income and poor air quality were mainly located in Eastern Europe. A study [16] in Czech Republic analyzed air pollution exposure with socioeconomic variables using principal component analysis (PCA). The first component with the variables SO_2, PM_{10}, low education level, and high unemployment, explained 44.7 percent of the data variability. This indicates that residents with poor socioeconomic position, that tend to live in smaller cities, are exposed to greater concentrations of combustion-related air pollutants. The second component with the variables NO_2, high pay, high education level, and large population, explained 28.1 percent of the data variability, implying that large cities with citizens with higher socioeconomic status are exposed to higher levels of traffic-related air pollution. In spite of these findings, they used a very small dataset and it's not common to have the dependent variable in the same PCA as the independent variables.

As for the current location of study - the Netherlands - there has been one study on the inequality topic with respect to air pollution. The study [17] explores the inequality on national, regional and city level using characteristics like deprivation, ethnicity, proportion of children within the Netherlands and England. It used descriptive statistics and correlations to describe the air pollution and demographic variables. To explore associations between the variables, a multiple linear regression has been done. It is found that air pollution inequalities are largely an urban issue. Moreover, ethnically diverse neighbourhoods found to have the highest air pollution levels. Unfortunately, this study doesn't analyze the important socioeconomic factor income and leaves out the $PM_{2.5}$ pollutant.

After assessing the available air pollution inequality research, a regression analysis is by far the most used method to analyze associations between socioeconomic factors and air pollution exposure. Though, many variants of regression analysis are being used due to the nature of their dataset. Descriptive statistics seem to be present in most studies, which is used to get a better understanding of variance in air pollution exposure between groups. Some studies add inequality indices like GINI or Atkinson, to further understand the distribution of air pollution among the population. A combination of these methods will employed for the present study.

3 Data

3.1 Ambient air pollution for the Netherlands

To determine air pollution in the Netherlands, three pollutants are considered i.e., NO_2, $PM_{2.5}$ and PM_{10}. These pollutants can be seen as the most detrimental contributors to human health and are salient in urban environments. Nitrogen oxides (NO), which are mostly caused by the combustion of fossil fuels, have a significant role in changing the composition of the atmosphere[18]. Subsequently, there is Particulate matter (PM), which is a term used to describe a type of air pollution that consists of complex and varying mixtures of particles suspended in the breathing air that vary in size and composition(e.g., $PM_{2.5}$ and PM_{10}), and is produced by a variety of sources, including factories, power plants, refuse incinerators, motor vehicles, construction activity, fires, and natural windblown dust [19].

This research will use high resolution (5x5 M) annual average air pollution concentration maps for the Netherlands from the European Study of Cohorts for Air Pollution Effects (ESCAPE) project [20]. The measurements are calculated with land use regression (LUR) models using mean air pollution from 2009 and can be considered as long term average values due to spatial contrasts remaining stable for multiple years. The LUR models were estimated from traffic intensity, population density, traffic infrastructure and land use.

3.2 Demographic data

The demographic data is originated from the main source of statistics in the Netherlands i.e., Central bureau of Statistics (CBS). The CBS has reliable statistics on country, city, neighborhood, postcode level. This research will use the smallest level of detail available; 6-character postcode level(PC6) statistics from the year 2018. The year of 2018 is chosen since more recent years didn't contain complete information of the socioeconomic variables.

It has to be noted that the generalizability of the results is limited by an important factor, namely the modifiable areal unit problem (MAUP). By using information that is aggregated into spatial areas would always introduce the MAUP. Because of the MAUP, it could significantly impact the results and either under- or over estimate the relationships found. The MAUP is especially evident in health spatial literature, where aggregating on health issues is commonly done [21]. However, in this research the smallest area level of detail (i.e, postcode level 6) has been used, to reduce the effect of MAUP. With that being said, it is not possible to interpret the results on individual level, but should only be interpreted on postcode level. For each postcode the following information is used for further analysis:

- Inhabitants: The number of inhabitants
- Male/Female: The number of male/female
- Ethnicity: The percentage of inhabitants with a non-western background
- Median Income: Labeled as low, low-middle, middle, middle-high and high
- Social benefit receivers: The number of inhabitants that receive social benefit

3.3 Data preparation

To calculate the exposure of NO_2, $PM_{2.5}$ and PM_{10} in the different postcode areas, the raster with the air pollutants and a vector of the postcode boundaries is loaded in QGIS. Using zonal statistics, the sum, mean and median is calculated for each postcode. Postcodes with no residents are removed from the set.

The layer containing postcode- and air pollution statistics is exported from QGIS to a .csv file and loaded into R. Due to the fact that postcodes contain various number of inhabitants, relative numbers are needed for the socioeconomic factors. The social benefit receivers is divided by the inhabitants and the the same is done for the number of male and female for each postcode. Non-western ethnicity was already encoded as a percentage. Unfortunately, the median income variable is not a continuous variable but each postcode is assigned to one of the 11 income classes. Since it isn't a continuous value, it can't initially be used for regression analysis. The choice has been made to create dummy variables for the four most extreme categories: Low-, Low-middle-, Middle-high- and High income. The dataset is then split up into three subsets of the cities Amsterdam, Rotterdam and The Hague.

4 Methods

Scripts used for the analysis available on GitHub: https://github.com/Sjors189/thesis_airpollutionexposure

4.1 Analysis

Various methods are employed to investigate inequality of air pollution exposure between different socioeconomic statuses. Descriptive statistics are used to understand the distribution of the pollutants across the cities. To understand the multicollinearity between the variables a Pearson's correlation is computed for the continuous social demographic variables and pollutants. As found in the literature review, another way to examine the distribution is the use of the Atkinson index. This index is initially made for the inequality of wealth, but has been used in other environmental health studies [13], [22]. The index has an inequality aversion parameter, which is introduced to set the societal concern of the inequality. In line with other environmental health related literature a 0.75 parameter is used [22], [23]. The index won't give insights in the inequality of exposure between different socioeconomic status, but it will give a understanding of the inequality of exposure on the total population. Using a Lorenz curve, it is visually displayed whether there is an inequality or equality in the population. The Lorenz curve is weighted by the inhabitants of each postcode area.

To find possible relationship between the socioeconomic variables and pollutants, a regression analysis is done. It has to be noted that the regression analysis is used to understand positive or negative correlations and not casual influence. With the assumption of multicollinearity within the predictor variables a normal Ordinary Least Squares (OLS) regression would introduce overfitting. To overcome this problem regularization is necessary. The methods Lasso and Ridge both use a regularization technique by adding a penalty and thereby, shrinking the beta coefficients to zero. Elastic Net combines both methods where it introduces α which is the mixing parameter between ridge ($\alpha = 0$) and lasso ($\alpha = 1$). In total nine regression models are made, which derives from every city (i.e., The Hague, Rotterdam, Amsterdam) times every pollutant (i.e., NO_2, $PM_{2.5}$, PM_{10}). Every model is tuned using cross-validation for the most optimum α and λ values. The coefficients of the regression models are the main results used to argue any positive or negative effects for different sub-populations.

Since income is an important socioeconomic factor, but not a continuous variable in this study, a one-way Analysis of Variance (ANOVA) test is performed for comparing the means in air pollution exposure of the categorical variable median income. Rather than just comparing the means, this test allows to check whether the variations between group means is higher than the variations within the group, resulting in the so-called F-value. After confirming a significant F-value, the Tukey HSD test is used. This test is commonly used afterwards to see the mean differences between the groups in an absolute value. Using these two tests, the found relationship can be tested on statistical significance.

5 Results

First, let's take a look at the mean air pollution concentrations and demographic variables of Amsterdam, The Hague and Rotterdam in Table 1. The air pollution levels in the cities differ quite much from the average of the Netherlands. This can be expected since the Netherlands also has rural areas with limited air pollution. Between the cities there is not much variations, only in The Hague where NO_2 levels are around 20% higher compared to the other cities. The social and demographic variables ethnicity, education, income and social benefit receivers of the cities show also much similarity.

	Population	Western background	Non-western background	Low/middle education	People in lowest 40% income	Social benefit receivers	Social assistance (unemployed)	NO_2 $\mu g/m^3$	$PM_{2.5}$ $\mu g/m^3$	PM_{10} $\mu g/m^3$
Netherlands	17282163	10,3%	13,4%	69,8%	7,7%	2,5%	1,3%	11,33	9,14	16,29
Amsterdam	862965	18,8%	35,7%	52,0%	13,3%	4,7%	1,5%	28.8	16.6	27.5
The Hague	537833	18,6%	36,0%	65,5%	13%	4,7%	1,4%	34.6	16.9	27
Rotterdam	644618	13,0%	38,6%	69,7%	14,1%	5,7%	1,5%	28.2	16.9	25.6

Table 1: Descriptive statistics of socioeconomic characteristics and air pollution on city level.

To have a better understanding of the distribution of air pollution among the postcodes, a histogram for each city and pollutant is made and can be seen in Figure 1. As expected, it shows a normal distribution for all cases, although some difference can be noted. The NO_2 pollutant seems to have the most 'perfect' normal distribution across all cities, ranging between $\pm$ 15 - 66 $\mu g/m^3$. The $PM_{2.5}$ pollutant definitely shows a right (positive) skew across all cities. As for the PM_{10}, it only shows a slight right (positive) skew in The Hague and Rotterdam. Another way to further understand the distribution are Lorenz curves combined with the Atkinson index. Figure 2 shows a Lorenz curve with the cumulative population compared to the cumulative NO_2 emission. This graph gives insight in how the environmental hazard from air pollution is distributed across the population (i.e., The Hague, Rotterdam, Amsterdam). The NO_2 pollutant showed a slight inequality. The Atkinson index (= 0.01) shows there is almost full equality, with zero being total equality and one total inequality. The Lorenz curve therefor shows minimal distance from the diagonal equality line. Concerning the slight inequality, it shows that 45% of the population is exposed to 50% of the total NO_2 emissions. Almost no inequality was found for the pollutants $PM_{2.5}$ and PM_{10} using Atkinson index. In contrast to the histogram, the Lorenz curve is weighted by the inhabitants for each postcode and therefor gives additional insight in the real distribution among the population.

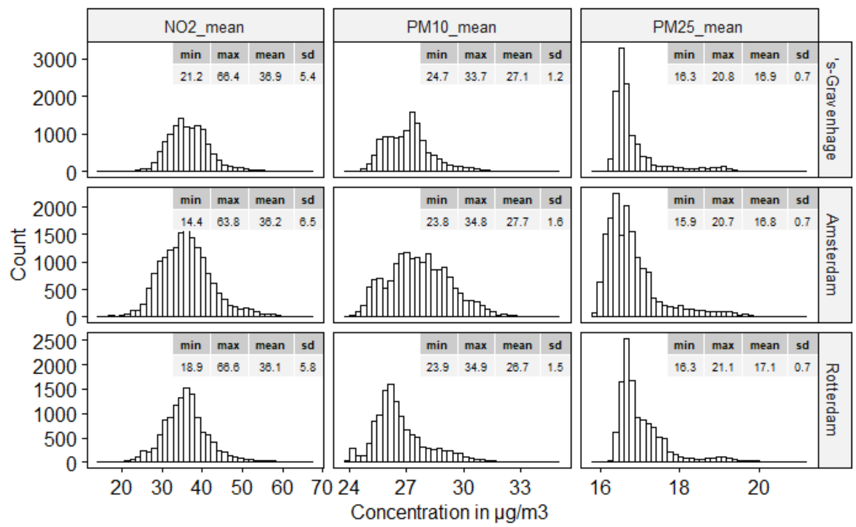


Figure 1: Histograms showing the distribution of air pollution concentration for each pollutant and each city.

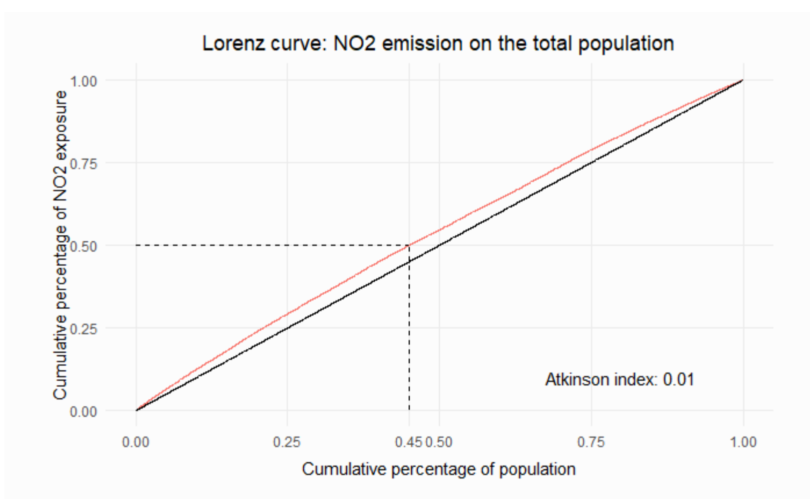


Figure 2: Lorenz curve showing the cumulative percentage population compared to the cumulative NO_2 emission. The population are all inhabitants of The Hague, Rotterdam and Amsterdam.

5.1 Pearson's correlation

To understand the associations and investigate multicolinearity between the variables, a correlation matrix is made, and can be seen in Table 2.

	NO2	PM2.5	PM10	% Female	% Male	% Non-Western	% Social benefit receivers
NO2	*						
PM2.5	0.570	*					
PM10	0.790	0.710	*				
% Female	-0.066	-0.031	-0.048	*			
% Male	0.050	0.016	0.034	-0.711	*		
% Non-Western	-0.042	-0.069	-0.102	0.017	0.061	*	
% Social benefit receivers	-0.043	-0.060	-0.079	-0.014	0.060	0.513	*

Table 2: Correlation matrix with air pollutants and demographic variables of Amsterdam, The Hague and Rotterdam combined

The three air pollutants NO_2, $PM_{2.5}$ and PM_{10}, show moderate to high positive Pearson's correlations with each other. As one would expect, % Female is highly correlated with % Male. A moderate correlation is found between the demographic variables non-western ethnicity and social benefit receivers. This relationship is found across all the cities ($r > 0.5$). In terms of correlation between air pollution concentrations and socioeconomic factors, no convincing correlations have been found. Only weak correlations between non-western ethnicity and NO_2 have been found in The Hague ($r = 0.26$) and Amsterdam ($r = -0.3$).

5.2 ElasticNet regression

The coefficients (in Table 3) of the ElasticNet models show whether there is a positive or negative relationship with the pollutant and to what extend this contributes. At first glance, the coefficients suggest that the strongest effects appear around the NO_2 pollutant. This is however not that simple, since the range of concentrations is far larger for NO_2 pollutant than it is for $PM_{2.5}$ and PM_{10}, as can be seen in Figure 1. Looking at the non-western ethnicity variable, it shows quite the distinction between The Hague and Rotterdam compared to Amsterdam. In the first two cities, non-western ethnicity positively correlates with almost all the pollutants, while in Amsterdam it negatively correlates with the pollutants. A reason that could explain this difference is that in Amsterdam, the city centre is to a greater extend occupied by Dutch or western background, but due to the increased traffic of the city centre, there are relatively high levels of air pollution. This can visually be confirmed by looking at the non-western ethnicity map of

Amsterdam (Figure 3a) and the NO_2 map (Figure 4a). By way of contrast, this does not apply for The Hague, where the high NO_2 levels are mainly located in the non-western communities. Second, the social benefit variable negatively correlates with the pollutants across all cities. The third socioeconomic variable is income, where four categories of median income are used within the model. The general trend that can be seen is that low / low-middle class postcodes have a positive relationship with all the air pollutants. In contrast to the high / high-middle class residents which generally has a negative relationship. Lastly, for all cities the % men positively correlates with the pollutants and the % female shows a negative correlation. The coefficients for % female suggest a stronger effect compared to % men.

	The Hague			Rotterdam			Amsterdam		
	$PM_{2.5}$	PM_{10}	NO_2	$PM_{2.5}$	PM_{10}	NO_2	$PM_{2.5}$	PM_{10}	NO_2
Intercept	*16.85*	*27.16*	*37.25*	*17.25*	*26.92*	*36.83*	*17.02*	*28.25*	*38.52*
% Non-Western ethnicity	0.11	0.77	4.66	-0.06	0.03	1.57	-0.47	-1.96	-7.40
% Government social assistance	-0.17	-0.61	-3.82	-0.52	-1.28	-5.01	-0.19	-0.2	-1.09
% Men	0.03	0.13	0.77	0.01	0.05	0.02	0	0.25	1.27
% Female	-0.06	-0.56	-3.41	-0.18	-0.16	-1.64	-0.11	-0.3	-1.79
Low income	0.09	0.4	1.84	0.08	0.27	1.41	0.15	0.42	1.7
Low-middle income	0.05	0.21	0.85	-0.02	0	0.35	0	0.08	0.43
High income	0.01	-0.07	-0.97	-0.25	-0.32	-1.89	-0.03	0.06	-0.09
High-middle income	-0.07	-0.3	-1.46	-0.11	-0.16	-1.21	0	-0.08	-0.89

Table 3: ElasticNet coefficients for the pollutants NO_2, $PM_{2.5}$ and PM_{10} using postcode 6 statistics of the cities The Hague, Rotterdam and Amsterdam.

5.3 ANOVA test

The relationship between income and air pollution is further examined using the ANOVA test and Tukey HSD. It again gives evidence for the inequality of air pollution based on income. In Table 4 can be seen that all the F-values are significant, signaling a difference in means between the different income groups. Upon further analysis with the Tukey HSD, the absolute differences between the groups can be found. The Hague shows the most clear inequality based on income, where the mean of high income postcodes is for all pollutants lower than for the low income postcodes. Rotterdam shows the same trend, although the mean differences are lower and not always significant. Conversely, in Amsterdam postcodes of high income have higher mean values than low income postcodes across all pollutants.

Like the regression analysis, it again shows a clear difference between The Hague and Rotterdam compared to Amsterdam. This can also be traced back to the city centre of Amsterdam, where due to the expensive house prices, mainly wealthy people are located.

	The Hague			Rotterdam			Amsterdam		
	PM25	PM10	NO2	PM25	PM10	NO2	PM25	PM10	NO2
ANOVA									
F-value	10.66***	97.68***	140.5***	8.53***	8.15***	40.24***	41.9***	104.6***	71.39***
Tukey HSD - Difference									
High - Low	-0.11*	-0.7***	-4.12***	-0.2***	-0.36***	-2.97***	0.19	0.46***	1.37**
High - Middle Low	-0.06	-0.46***	-2.94***	-0.11*	-0.11	-1.99***	0.20***	0.87***	2.93***
Middle High - Low	-0.18***	-0.92***	-4.56***	-0.06	-0.2	-2.27***	0.04	0.26*	0.36
Middle High - Middle Low	-0.13***	-0.69***	-3.38***	0.03	0.05	-1.29***	0.22***	0.68***	1.93***

Table 4: Results of the ANOVA test and Tukey HSD with the (mean) pollutants as dependent variable and the categorical variable median income as independent variable. Multiple tests have been performed for the pollutants NO_2, $PM_{2.5}$ and PM_{10} and the cities The Hague, Rotterdam and Amsterdam. Significance codes: 0: '***' , 0.001: '**' , 0.01: '*'

6 Discussion

The results of the study indicate that there exists a social inequality on air pollution exposure in the Netherlands, although it depends on the city, social variable and pollutant. The most apparent city of inequality is The Hague. The regression analysis (Table 3) reveals that low income postcodes together with high percentages of non-western ethnicity are disproportionately exposed to air pollution in The Hague. This phenomenon can also visually be confirmed when inspecting the map of The Hague with respect to non-western ethnicity in Figure 3b. It shows a graduated change from left to right, really indicating a division between high and low socioeconomic statuses. When inspecting the NO_2 map of The Hague (Figure 4b) with respect to non-western ethnicity, it becomes clear there is indeed a social inequality of air pollution exposure. As for Rotterdam, the results also point towards a social inequality of air pollution exposure based on ethnicity and income. Conversely, in the city of Amsterdam the opposite is happening; non-western postcodes are less exposed to air pollution. For all the pollutants the regression analysis shows a negative coefficient on the non-western ethnicity variable, with NO_2 having a great coefficient of -7.40. The income variable gives mixed signals between the regression analysis and ANOVA test in Amsterdam. The regression analysis suggests that lower income postcodes are more exposed to air pollution, but the ANOVA test says the reverse. Since the ANOVA test is statistically significant together with the maps of Amsterdam showing similar postcodes of high non-western ethnicity and low income, it is believed that low income postcodes are potentially exposed to air pollution in Amsterdam. This is also in line what is found in other literature, for example, in Rome (Italy) where people with higher socioeconomic status were exposed to higher volumes of NO_2, PM_{10}, because they lived in the centre where there was more pollution [24]. Similarly, this phenomenon is found to happen in more metropolitan areas like New York [14]. Additionally, the social variables gender and social benefit receivers have been evaluated. Across all cities and pollutants, the %Female coefficients of the regression model are negative, indicating that postcodes with high percentage of females are less exposed to air pollution. This could be explained due to the fact that women tend to have stronger neighborhood preferences than men and therefor end up in higher SES neighborhoods [25]. As for the variable social benefit receivers, it is surprising that the regression analysis shows negative coefficients across all cities and pollutants, since non-western ethnicity and social benefit receivers are moderately correlated. It could be explained by two reasons. First, this variable contains all the social benefit packages that exist in the Netherlands, where some don't necessarily mean a person has a low SES. Second, as seen in Table 1, only around 5% of the people receive one of these benefits, resulting in many postcodes that don't have social benefit receivers, which makes it prone for outliers.

A limitation apparent in this study was that the median income is categorized in groups rather than a continuous field. This led to the use of dummy variables and therefor only picking the outer income classes (i.e, low and high) in the regression analysis. This definitely impacts the regression analysis on income in relation to air pollution exposure, since all the other income classes are not considered. For that reason, the ANOVA test is done to further examine the income variable and have a more reliable view of income with respect to air pollution. The ANOVA analysis showed that the mean air pollution differs between high and low income postcodes. In The Hague and Rotterdam, for almost all pollutants, the low income postcodes have an increased mean compared to the high income postcodes. As said before, the reverse is true for Amsterdam, where low income postcodes have a lower mean than the high income postcodes.

Considering the results, it shows how important it is to evaluate the cities individually rather than grouping them together. If the cities were to be analyzed together, the social inequality found in the cities The Hague and Rotterdam would be disregarded by the opposite effects of Amsterdam. Present study therefor provides a better understanding of the social inequality of air pollution exposure in the Netherlands. Policy makers within the (local) government can account for this problem and find approaches to reduce the inequality, to eventually ensure that lower SES groups are not disproportionately burdened by air pollution exposure.

7 Conclusion

This study aimed to identify the relationships between socioeconomic factors and air pollution exposure within the Netherlands. It has become clear that there is indeed an existence of this inequality. Though, some cities show this relationship to a greater extend than others. The Hague showed the most apparent case of inequality, where low income and high non-western ethnicity postcodes are considerably more exposed to air pollution. Rotterdam showed the same relation, but to a lesser extend. In Amsterdam the relationship was reversed, where high income and low non-western ethnicity postcodes are more exposed to air pollution. Further research is needed to establish a better understanding of the social inequality of air pollution exposure in the Netherlands as a whole. Present study focused on the three biggest (urban) cities of the Netherlands, but future studies should take a greater sample of cities and also take rural areas into account. Furthermore, the socioeconomic variables education and employment has not been researched, but would give additional insight.

References

[1] *Health inequities and their causes*, Feb. 2018. [Online]. Available: https://www.who.int/news-room/facts-in-pictures/detail/health-inequities-and-their-causes.

[2] *Air pollution: How it affects our health*, Dec. 2021. [Online]. Available: https://www.eea.europa.eu/themes/air/health-impacts-of-air-pollution.

[3] *Who expert consultation: Available evidence for the future update of the who global air quality guidelines (aqgs) (2016)*, Mar. 2017. [Online]. Available: https://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2016/who-expert-consultation-available-evidence-for-the-future-update-of-the-who-global-air-quality-guidelines-aqgs-2016.

[4] *Air Pollution and Cancer Scientific Publication*, 161st ed., ser. IARC Scientific Publication. International Agency for research on Cancer (IARC). [Online]. Available: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Scientific-Publications/Air-Pollution-And-Cancer-2013.

[5] D. E. Schraufnagel, J. R. Balmes, C. T. Cowl, *et al.*, "Air pollution and noncommunicable diseases," *Chest*, vol. 155, no. 2, pp. 417–426, Feb. 2019. DOI: 10.1016/j.chest.2018.10.041.

[6] R. Fuller, P. J. Landrigan, K. Balakrishnan, *et al.*, "Pollution and health: A progress update," *The Lancet Planetary Health*, 2022. DOI: 10.1016/s2542-5196(22)00090-0.

[7] C. Hausman and S. Stolper, "Inequality, information failures, and air pollution," *Journal of Environmental Economics and Management*, vol. 110, p. 102 552, 2021, ISSN: 0095-0696. DOI: https://doi.org/10.1016/j.jeem.2021.102552. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S009506962100108X.

[8] T. Lockwood, N. T. Coffee, P. Rossini, T. Niyonsenga, and S. McGreal, "Does where you live influence your socio-economic status?" *Land Use Policy*, vol. 72, pp. 152–160, 2018, ISSN: 0264-8377. DOI: https://doi.org/10.1016/j.landusepol.2017.12.045. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264837717304994.

[9] A. Hajat, C. Hsia, and M. S. O'Neill, "Socioeconomic disparities and air pollution exposure: A global review," *Current Environmental Health Reports*, vol. 2, no. 4, pp. 440–450, 2015. DOI: 10.1007/s40572-015-0069-5.

[10] J. D. Marshall, "Environmental inequality: Air pollution exposures in california's south coast air basin," *Atmospheric Environment*, vol. 42, no. 21, pp. 5499–5503, 2008, ISSN: 1352-2310. DOI: https://doi.org/10.1016/j.atmosenv.2008.02.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1352231008001350.

[11] H. Zhao, G. Geng, Q. Zhang, *et al.*, "Inequality of household consumption and air pollution-related deaths in china," *Nature Communications*, vol. 10, no. 1, 2019. DOI: 10.1038/s41467-019-12254-x.

[12] T. Yang and W. Liu, "Does air pollution affect public health and health inequality? empirical evidence from china," *Journal of Cleaner Production*, vol. 203, pp. 43–52, 2018, ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2018.08.242. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652618325915.

[13] A. Rosofsky, J. I. Levy, A. Zanobetti, P. Janulewicz, and M. P. Fabian, "Temporal trends in air pollution exposure inequality in massachusetts," *Environmental Research*, vol. 161, pp. 76–86, 2018, ISSN: 0013-9351. DOI: https://doi.org/10.1016/j.envres.2017.10.028. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001393511731054X.

[14] A. Hajat, A. V. Diez-Roux, S. D. Adar, *et al.*, "Air pollution and individual and neighborhood socioeconomic status: Evidence from the multi-ethnic study of atherosclerosis (mesa)," *Environmental Health Perspectives*, vol. 121, no. 11-12, pp. 1325–1333, Sep. 2013. DOI: 10.1289/ehp.1206337.

[15] E. A. Richardson, J. Pearce, H. Tunstall, R. Mitchell, and N. K. Shortt, "Particulate air pollution and health inequalities: A europe-wide ecological analysis," *International Journal of Health Geographics*, vol. 12, no. 1, 2013. DOI: 10.1186/1476-072x-12-34.

[16] M. Branis and M. Linhartova, "Association between unemployment, income, education level, population size and air pollution in czech cities: Evidence for environmental inequality? a pilot national scale analysis," *Health Place*, vol. 18, no. 5, pp. 1110–1114, 2012, ISSN: 1353-8292. DOI: https://doi.org/10.1016/j.healthplace.2012.

04.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1353829212000731.

[17] D. Fecht, P. Fischer, L. Fortunato, *et al.*, "Associations between air pollution and socioeconomic characteristics, ethnicity and age profile of neighbourhoods in england and the netherlands," *Environmental Pollution*, vol. 198, pp. 201–210, 2015, ISSN: 0269-7491. DOI: https://doi.org/10.1016/j.envpol.2014.12.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0269749114005144.

[18] M. Kampa and E. Castanas, "Human health effects of air pollution," *Environmental Pollution*, vol. 151, no. 2, pp. 362–367, 2008, Proceedings of the 4th International Workshop on Biomonitoring of Atmospheric Pollution (With Emphasis on Trace Elements), ISSN: 0269-7491. DOI: https://doi.org/10.1016/j.envpol.2007.06.012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0269749107002849.

[19] K. Katsouyanni, "Ambient air pollution and health," *British Medical Bulletin*, vol. 68, no. 1, pp. 143–156, Dec. 2003, ISSN: 0007-1420. DOI: 10.1093/bmb/ldg028. [Online]. Available: https://doi.org/10.1093/bmb/ldg028.

[20] O. Schmitz, R. Beelen, M. Strak, *et al.*, "High resolution annual average air pollution concentration maps for the netherlands," *Scientific Data*, vol. 6, no. 1, Mar. 2019. DOI: 10.1038/sdata.2019.35.

[21] J. K. Nelson and C. A. Brewer, "Evaluating data stability in aggregation structures across spatial scales: Revisiting the modifiable areal unit problem," *Cartography and Geographic Information Science*, vol. 44, no. 1, pp. 35–50, 2017. DOI: 10.1080/15230406.2015.1093431. [Online]. Available: https://doi.org/10.1080/15230406.2015.1093431.

[22] L. P. Clark, D. B. Millet, and J. D. Marshall, "National patterns in environmental injustice and inequality: Outdoor no2 air pollution in the united states," *PLOS ONE*, vol. 9, no. 4, pp. 1–8, Apr. 2014. DOI: 10.1371/journal.pone.0094431. [Online]. Available: https://doi.org/10.1371/journal.pone.0094431.

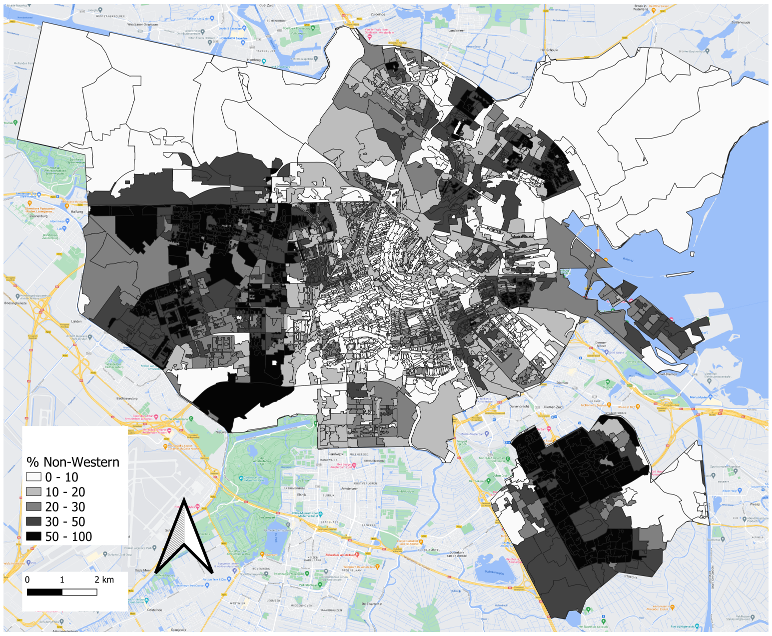
[23] N. Fann, H. A. Roman, C. M. Fulcher, *et al.*, "Maximizing health benefits and minimizing inequality: Incorporating local-scale data in the design and evaluation of air quality policies," *Risk Analysis*, vol. 31, no. 6, pp. 908–922, 2011. DOI: https://doi.org/10.1111/j.1539-

6924.2011.01629.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2011.01629.x.

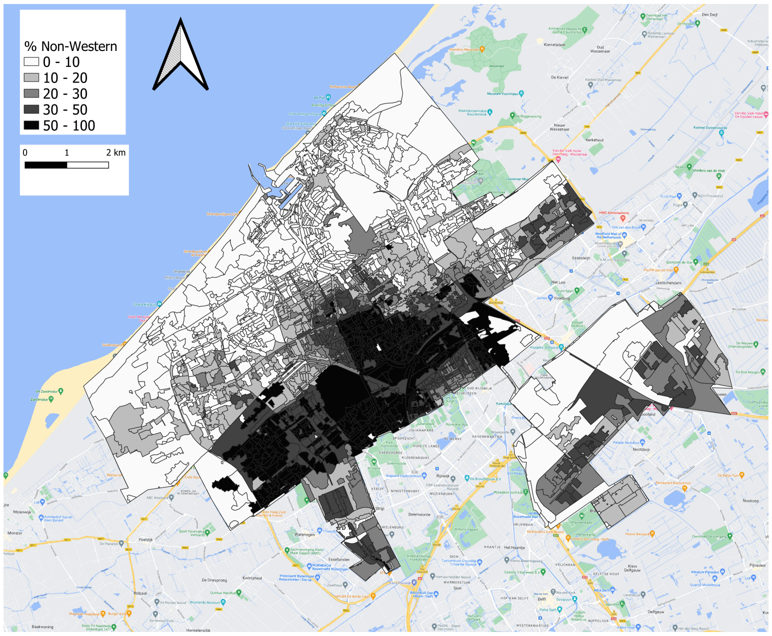
[24] F. Forastiere, M. Stafoggia, C. Tasco, *et al.*, "Socioeconomic status, particulate air pollution, and daily mortality: Differential exposure or differential susceptibility," *American Journal of Industrial Medicine*, vol. 50, no. 3, pp. 208–216, 2007. DOI: https://doi.org/10.1002/ajim.20368. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ajim.20368.

[25] H. S. Shawki, "Gender-related differences in housing preferences: A qualitative approach," Ph.D. dissertation, The Ohio State University, 2007.

A Appendix

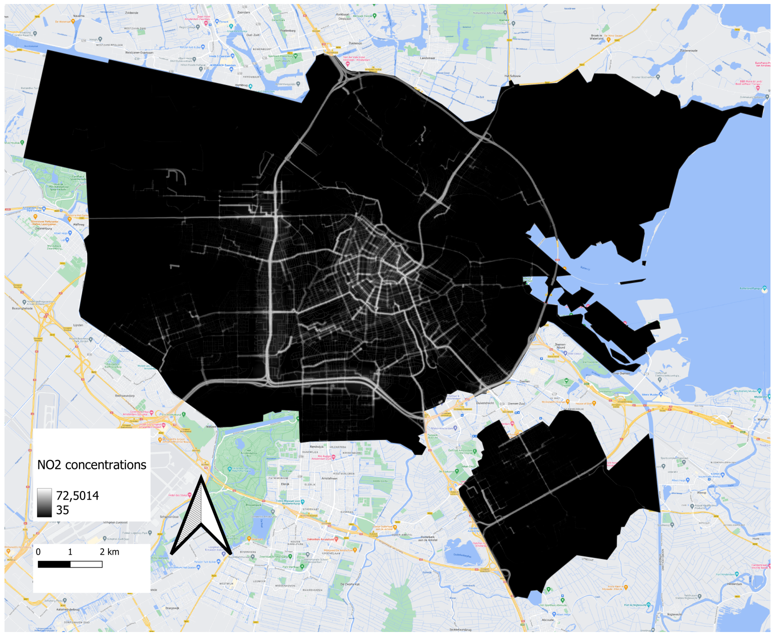


(a) Amsterdam

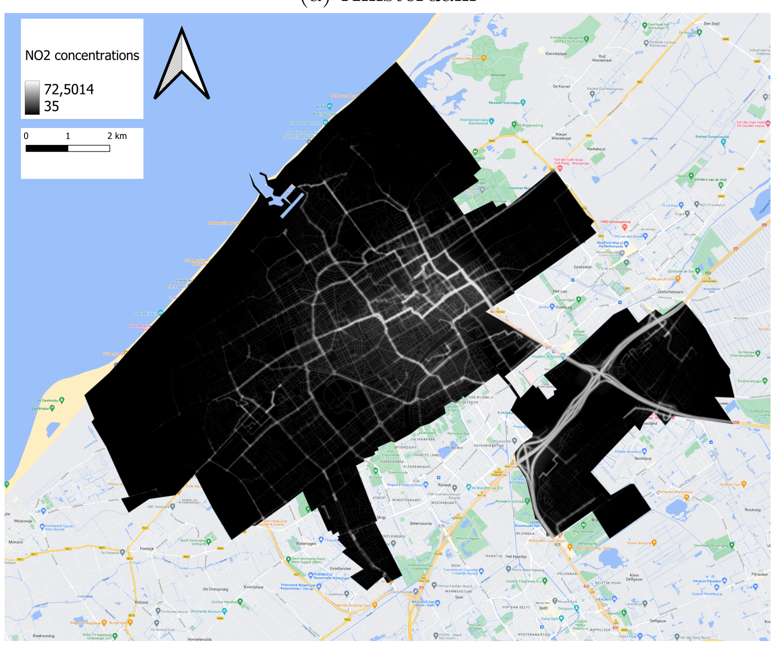


(b) The Hague

Figure 3: For the cities Amsterdam and The Hague, postcodes showing the percentages of non-western ethnicity for the postcodes.



(a) Amsterdam



(b) The Hague

Figure 4: For the cities Amsterdam and The Hague, NO2 exposure above $> 35\,\mu g/m^3$.