

The Price of Air Pollution

Avdovic Lejla, Blasch Fabian, Schachinger Maximilian, Prinz Martin

February 27, 2023



Abstract Air pollution and its consequences for nature and humanity is an omnipresent global issue, where particularly China stands out. **rasch_under_nodate** reports that coal combustion, industrial production and automobile emissions are the main causes for air pollution in China. Coal and oil combustion alone account for around 60 percent of $PM_{2.5}$ pollution in China. This paper analyzes the spatial effects of air pollution on the Chinese health care expenditures. We aim to answer the following research question: "To what extent does air pollution and its spatial spillover effects influence health care expenditures in China? A model with a spatial lag in X (SLX) was chosen based on theoretical justifications and testing procedure, accompanied by a feasible Generalized Least Squares (fGLS) model. The weights matrix chosen to describe the neighborhood structure is contiguity based, since Inner Mongolia, for example, has eight neighbors, whereas Xinjiang has just three. In addition, a robust covariance-variance matrix was included due to heterogeneity of regions and spatial autocorrelation. The results of the SLX model show that Nitrogen Oxid (NO) and Particulate Matter ($PM_{2.5}$) increase public health care spending by 0.006 and 0.032, as well as the lag of the second mentioned variable by 0.117 (where only the lag of Particulate Matter is highly significant). With sulphur dioxide (SO_2), only lag has a significant positive effect on our dependent variable of 0.093. In the fGLS model in comparison, the significant positive effect of Particulate Matter of 0.063 and the negative effect of Sulphur of -0.007 remain. We assume that an ommited variable bias is behind these counterintuitive results.

1 Introduction

Air pollution in China is a persistent issue, but how serious is it? China's recent rise to an economic powerhouse, characterized by the emergence of mega cities and the increasing demand for energy, has resulted in alarmingly high levels of air pollution. According to **chan2008air** the primary drivers of rising levels of air pollutants are China's metropolitan areas. Typically, a large proportion of main direct and indirect emission origins such as power plants, biogenic, industrial and vehicular sources and domestic heating can be allocated to these areas. However, polluting the environment is not only problematic in terms of climate change, but it is also reflected in the health of the population. **maruyama2022air** state that approximately 53 percent of the total population - around 0.765 billion people - are exposed to hazardous concentrations of PM_{2.5} exceeding 35 $\mu\text{g}/\text{m}^3$. **song_air_2017** report that the Global Burden of Diseases (GBD) project could assign around 1.6 million deaths in China to air pollution related causes alone in 2015. In addition, persistent and excessive air pollution can be a huge burden to the economy of China: A sustainable development is endangered because natural and human resources are being exposed to dangerous levels of emissions. With increasing health impairments of the population, costs of the public health system are rising. Due to the outdated structure of the social security system in China, which will be discussed in more detail in section 2.3, it is mainly the rural migrants in the urban areas who are suffering the most, as they have to cover the entire costs of medical treatments in urban institutions and pay "out-of-pocket". From a geographical perspective, China features a variety of regions, from rural, mountainous locations to coastal, heavily industrialized clusters. The impact of air pollution may differ between regions and their adjacent provinces. Inspired by the paper "Does industrial air pollution drive health care expenditures? Spatial evidence from China" by **zeng2019does**, our goal is to conduct an independent analysis, which orients on their research, identifies possible shortcomings and adds value to this research field. In their study the authors employ a spatial lag model in order to disentangle the direct and indirect spatial effects of industrial air pollution on health care expenditures over the period of 2002 to 2014. Their findings reveal that industrial air pollution is a significant contributor to increased health care expenditures not only in the provinces where air pollution is emitted, but also in neighboring provinces. Furthermore, the authors assert that provincial health care expenditures are shaped by a combination of industrial air pollution, health reforms and other socioeconomic factors. These results highlight the importance of considering spatial effects and a range of contextual factors when analyzing the relationship between industrial air pollution and health care expenditures. In this paper we want to adopt the idea of the above mentioned paper and conduct our own spatial analysis using the appropriate methods. Our goal is to answer the following research question:

"To what extent does air pollution and its spatial spillover effects influence health care expenditures in China?"

This paper is organized as follows. In Section 2, first health impacts of air pollution on human health are discussed. This is followed by a brief overview of air pollution policies implemented in the observation period and the organization of the Chinese healthcare system. The chapter is intended to shed light on the chain of effects of air pollution on government expenditures. More specifically, exposure to air pollution of individuals has negative health effects on them. As a result of deteriorating public health, provincial health expenditures increase. Section 3 provides a brief description of the data used and an exploratory analysis. Subsequently, in Section 4 the process finding a suitable model and the corresponding testing procedure is described, while Section 5 presents the results. Concluding remarks as well as problems and limitations will be discussed in Section 6.

2 Literature review

2.1 Effects of air pollution on health

Analysing the effect of air pollution on health expenditures implicitly implies the aforementioned causal transition dependencies. Various studies have identified direct effects of exposure to air pollution on health.

franklin2015air and **fiordelisi2017mechanisms** demonstrate that the risk of cardiovascular diseases and the triggering of acute cardiac events is increased by PM air pollution. The pathways through which this occurs include the generation of proinflammatory or oxidative stress mediators in the lung that enter the

systemic circulation, the direct infiltration of certain particles or components into the cardiovascular tissue, or an imbalance of the autonomous nervous system. In that context, **hoek2013long** quantify the effect of PM_{2.5} long-term exposure by conducting a meta-analytic review of previous studies. Their pooled estimates indicate an increase in all-cause mortality and cardiovascular mortality of six percent and eleven percent, respectively, if increments of PM_{2.5} are increased by 10- $\mu\text{g}/\text{m}^3$. Equivalently, an increase in NO concentration of the same magnitude leads to an increased all-cause mortality of five percent. Not only cardiovascular diseases can be enhanced or even caused by air pollution. **Kim, et al. (2018)** found direct impacts of polluted air conditions on lung diseases and pregnancy disorders. According to the authors, even prenatal or perinatal air pollutant exposure can be associated with reduced lung functions of children and can have long-term effects on respiratory health of adults. Additionally, a Swedish study conducted by Nyberg et al. (2000) focussed on the relationship between urban air pollution and lung cancer risk. The authors used data from national health registers and modeled air pollution exposure. The results show a significant correlation between higher NO pollution and increased incidence of lung cancer. TO be specific, they attributed a higher relativ lung cancer risk between 0.8 and 1.6 percent to affic related NO exposure.

A 2010 study from Leliveld et al. estimated that between 1.6 and 4.8 million worldwide premature deaths occur due to outdoor air pollution in 2010 alone, predominately as a result of PM 2.5 pollution. The largest share of these deaths could be attributed to Asia.

Additionally, **zhao2019emerging** suggest that based on recent findings air pollution may be involved in the development of autoimmune diseases such as diabetes mellitus, multiple sclerosis, or rheumatoid arthritis. The authors argue that air pollution can cause imbalances in T cells, the production of proinflammatory cytokines, oxidative stress, local pulmonary inflammation and methylation changes. These effects are involved with initiating or aggravating autoimmune diseases.

To illustrate the magnitude of air pollution-related illnesses, the causes of death by illness patterns provide information. In 2019 according to **who_nodate** the major premature death causes in China were cardiovascular diseases totaling 43 percent, followed by malignant neoplasms (26 percent), respiratory diseases (10 percent), unintentional injuries (6 percent) and neurological conditions (4 percent). Other conditions that accounted for between 3 and 1 percent of premature deaths, in descending order, were digestive diseases, genitourinary diseases, respiratory infections, diabetes mellitus and infectious and parasitic diseases.

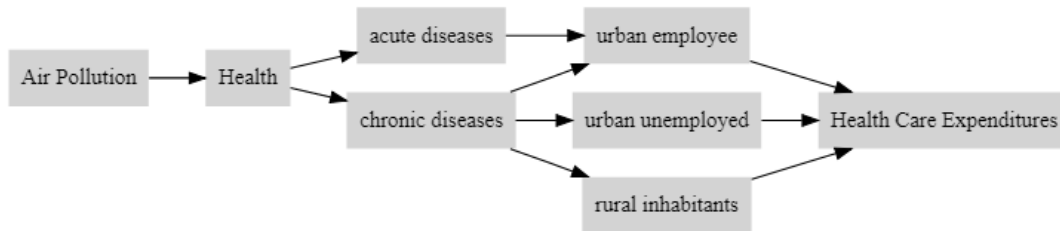


Figure 2.1 illustrates the channel between Air pollution and Health Care Expenditures as a directed acyclic graph (DAG). As already mentioned in detail above, air pollution - namely nitrogen oxide (NO), sulfur dioxide (SO_2) and Particulate Matter ($PM_{2.5}$) have a direct negative impact on health, as the harmful pollutants penetrate the blood circulation through the lungs and damage other organs. Depending on the duration or concentration of exposure to pollutants, acute diseases such as asthma, bronchitis, respiratory infections, and others can occur. Those can result in long-term consequences, i.e., chronic diseases such as cardiovascular disorders, lung cancer, etc. Depending on which social insurance system a patient is covered -

more on this in the section 2.3 - outpatient stays are partially covered. In general chronic diseases are partly covered by public expenditures. And so environmental pollution affects the public budget of health care.

2.2 Air pollutions policies from 2011 to 2018

Air pollution in China has been a major public health problem for many years. However in recent years the government has taken various measures to address the issue. One of the main focuses of these measures has been to reduce emissions of particulate matter, sulfur dioxide and nitrogen oxides. In 2013, prompted by a period of heavy smog in eastern China, the government introduced an "Air Pollution Control Action Plan" to combat air pollution, which included specific targets for reducing particulate matter, sulfur dioxide, and nitrogen oxide emissions. More specifically, targets included, among others, reducing PM10 concentrations in cities by more than 10 percent and reducing PM2.5 concentrations in the Beijing-Tianjin-Hebei, Yangtze River Delta and Pearl River Delta regions of around 25, 20 and 15 percent, respectively.

Furthermore, the 13th Five-Year Plan (2016-2020) also set a target to reduce PM2.5 concentrations in areas heavily affected by air pollution by 18 percent by 2020. Targets have also been set for reducing sulfur dioxide and nitrogen oxide emissions by 15 percent and 15 percent, respectively, compared to 2015 levels. Huang, Pan, Guo, Li (2018) indicate in their analysis, in which they map the national air quality in 74 cities that the issued Air Pollution Prevention and Control Action Plan (APPCAP) in 2013 has shown effect. Between 2013 and 2017 PM2.5 concentrations have reduced in average by about 33 percent, and PM10 concentrations by about 28 percent. Sulfur dioxide concentrations have reduced as well with an average reduction of about 54 percent. Nitrogen Oxides emissions, however, have not significantly decreased.

In addition, new air quality standards were set in February 2012. These were initially enforced in large cities and their surrounding areas in the following years until they were implemented nationally in 2016. The new standards set stricter limits for air pollutants such as particulate matter, sulfur dioxide, nitrogen oxides and ozone to better protect public health. Chinese authorities have also introduced new monitoring and reporting technology to ensure that air quality is measured and monitored in every part of the country. Companies that violate the new standards can be fined or even shut down.

In summary, a number of measures have been taken by the Chinese government from 2011 to 2018 to reduce air pollution, including stricter emission standards for vehicles, power plants and industrial facilities, and the closure or upgrading of older, heavily polluting factories, which were partially effective.

2.3 Organization of the Chinese Healthcare System

Since the late 1990s, which are also considered to be the years of the opening and globalization of China according to **kanbur_fifty_2005**, in which the country developed from a centralized to an open market economy, there have been several economic reforms in the country, including in the Health Care System. **hougaard_chinese_2011** state that the Chinese Central Government was heavily reduced from healthcare meaning the provinces are now responsible for financing and administering the Health Care System. Financing the system based upon local taxation, which led to substantial inequalities between rich coastal regions and poor rural regions.

Before we delve deeper into the organization of the Chinese Health Care System, it is worth mentioning how social benefits are allocated in China. The hukou system is the official residence registration system on household level in China, where each individual is assigned a permanent residence. **noauthor_oecd_nodate** report that workers from rural regions who migrate to urban areas face discrimination in accessing services due to not having an urban residency registration. **liu_institution_2005** describes that originally the hukou system was established in Chinese cities to "maintain social peace and order, ensure the security of the people, and protect their freedom of residence and movement." However, this changed in 1955. There was a strong influx of farmers into the cities. In response to the government, the hukou was expanded to also include rural regions. The registration system was now considered a government instrument not only control but also mitigate migration flows between urban and rural areas. After that, any change of residence required an administrative authorization. It was essential from the point of view of the centrally planned government to

distribute human resources across various geographic locations. Therefore, from a contemporary perspective, the Hukou system could also be seen as a relic of the centrally planned economy.

Because rural areas can further absorb and make use of surplus labor, the central government thought that the majority of the population should live there. Domestic migration of people was also viewed as dangerous because it could result in city overpopulation and endanger farm production. Domestic migration was likewise discouraged by the federal government. The central government closely regulated migration flows, and only recently have these limitations been loosened, in the era of Mao Zhedong. This doesn't imply that the hukou system is no longer in place; rather, it only means that, for instance, rural dwellers can move to a city but won't be eligible for any social benefits there and will have to make due with their own resources.

In the late 1990s, the years of the opening and globalization as **kanbur_fifty_2005empty citation** describes it, China developed from a centralized to an open market economy, there have been several reforms in the country, including in the health care system.

According to **hougaard_chinese_2011empty citation** no longer the central Chinese government but rather the different Provinces are responsible for the financing and administrating the health care system, instead of a . This, since funding is provided through taxation, in turn led to strong inequalities between the rich coastal regions and the impoverished rural regions in the east of the country. The Chinese health care system is characterized by the division of the population between two groups: The proportion of the population living in the urban and rural areas.

3 Data

For our analysis, we use data retrieved from the National Bureau of Statistics of China. The data was compiled from an extensive database. It provides data on government spending, air pollution, and various socioeconomic factors on a provincial level. In total we use yearly data from 2011-2019 from 31 provinces. In advance, we excluded autonomous regions such as Hong Kong, Macau or Taiwan, for which data were also available. The dependent variable of our analysis "Local Governments Expenditure on Medical and Health Care" includes all conceivable kinds of government spending, including for example health care management and services expenditures, disease prevention and motherhood expenditures as well as rural health spending. The variable is measured in 100 million Yuan. Additionally, we extracted three different pollution variables, namely particulate matter, sulfur dioxide and nitrogen oxides emissions volumes, which are measured in 10000 tons. However, no information was provided on whether the data was smoothed regionally. Furthermore, include various socioeconomic control variables in our analysis. ariables such as the population shares of different age groups (0-14, 15-65, 65-), which were collected through random sampling, the average disposable household income for each urban and rural population in the province, the coverage rate by forest in the region, and the proportions of urban and rural population in the region serve as important control variables in our analysis. In addition, we included other variables in our initial models, but at a certain point we no longer considered them important because the influences of this variable were either already covered by another variable or they offered very little explanatory value for our analysis. These include the consumer price index, the gross regional product, total local government expenditures und various key figures relating to the health system.

3.1 Exploratory Analysis

Now that we have established the source as well as the meta information. We will concisely examine the variables of highest relevance. The nature of our dataset does unfortunately not allow for a static representation that could display information of all provinces across time. Accordingly, we will focus on the values of our dependent variable as well as all emissions variables based on the beginning of our time series in 2011 and the end in 2019. In particular, the focus will be on the values in the different provinces at those points in time as well as their spatial interaction measured via the local Morans's I. Starting with our dependent variable we observe that the health care expenditures across all provinces increased across the eight years in our sample, this is not very surprising as china's population grew quite drastically in this time frame and accordingly the health care expenditures grew as well.

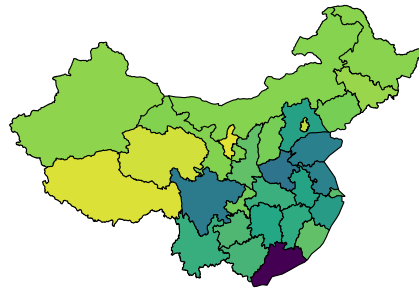
Health Care Expenditure in: 2011



Expenditure in
100MM of Yuan

1200
800
400

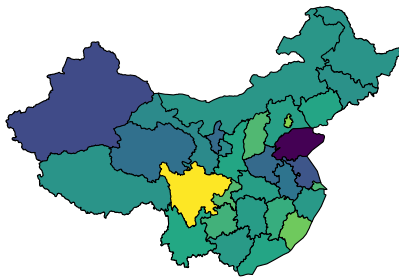
Health Care Expenditure in: 2019



Expenditure in
100MM of Yuan

1200
800
400

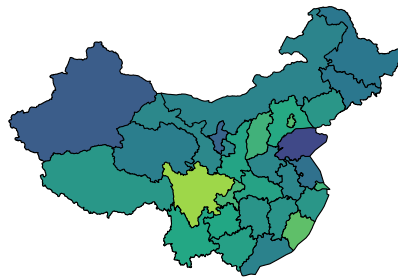
Local Moran's I for Health Care Expenditure in: 2011



Local Moran's I
Queen Contiguity

1
0
-1

Local Moran's I for Health Care Expenditure in: 2019



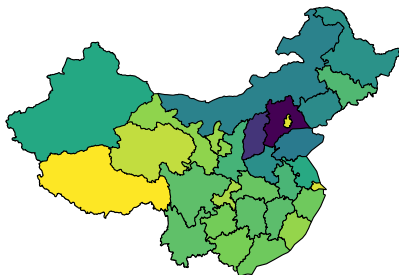
Local Moran's I
Queen Contiguity

1
0
-1

Further, when examining the spatial interaction we see that it changed quite drastically from 2011 to 2019. This is at least to some degree caused by the increased heterogeneity across provinces when it comes to health care expenditure at the end of our time frame.

As a representative for our pollution variables, Particulate Matter emission developments are displayed below. We can immediately tell that Particulate Matter emissions for most provinces and especially for Hebei, the province that surrounds Beijing, decreased quite substantially from 2011 to 2019. However, this does not seem to be the case for two special cases located in the north-east of china. Both in Jilin and in Inner Mongolia, the amount of Particulate matter emission increased.

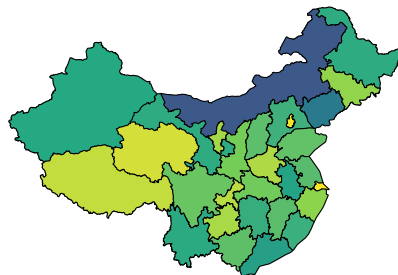
Particulate Matter Emissions in: 2011



Particulate Matter
in 10K of Tons

125
100
75
50
25

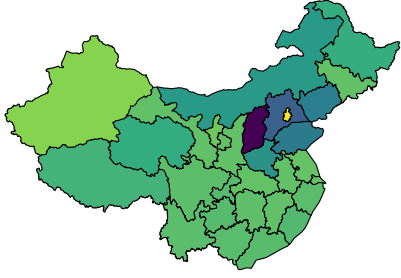
Particulate Matter Emissions in: 2019



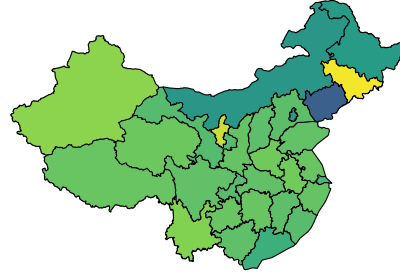
Particulate Matter
in 10K of Tons

125
100
75
50
25

Local Moran's I for Particulate Matter in: 2011



Local Moran's I for Particulate Matter in: 2019



In regard to the spacial interaction, we can observe that the Moran's I also changed quite significantly, again a large part of those changes may be attributed to stark decreases in emissions in provinces around Beijing. The maps for the remaining variables have been added to the appendix.

4 Methodology

As already discussed in section 2, the possibility of a spatial autoregressive process (SAR) is rather unlikely considering the Hukou system. We therefore excluded the possibility of a SAR component before using statistical testing to determine the final model.

We started the testing procedure by using the (Spatial) Hausmann Test to decide whether fixed or random effects suit the data better. The test strongly indicated the usage of a fixed effects model over time, which is in line with the intuition as China's provinces are heterogeneous and eliminating time constant features such as geographic differences seems like a reasonable approach.

In the following a Lagrange Multiplier Test to decide whether a spatial error model (SEM) should be used. We could not reject the H_0 of a OLS model. Therefore, we continued to use a spatial lag in X (SLX) model, whose lag components can directly be tested by t-tests. The inclusion of a SLX component comes from the fact that air pollution of one province is likely to influence air pollution of neighbouring provinces. But, as already discussed in section 3, it is not known whether the emission variables are only total emission in a provinces that were also created in this province or if some smoothing of the data was already applied such that some distance travelled by means of kernel distribution or similar is used in order to account the effect on neighbouring provinces. The introduction of spatial lags concerning the air pollution will not cause any harm other than potential multicollinearity issues if they are in fact redundant.

The next step was to check to possible presence of heteroscedastic and serial correlated errors. Firstly, a Breusch-Pagan test was used to test for heteroscedasticity. Here we could reject the Null-Hypothesis of homoscedastic errors. Secondly, a Durbin-Watson test was conducted to test for serial correlation. The test strongly suggested serial correlation. As a result, we used a robust variance covariance matrix in order to account for this error structure and improve the reliability of our results

For the identification of the model, a general-to-specific approach was utilized, where we started with all reasonable variables and potential lags and eliminated them iteratively with the use of t-tests.

As a robustness check of our results, we used a feasible generalized least square model. Very similar to the SLX model we use a robust variance covavariance matrix that is corrected for heteroscedasticity and serial correlation. Also, the most general model is assumed again and variables are eliminated one-by-one.

Some variables of interest are in monetary terms. To account for the mere growth in these variables due to increased prices we adjust every monetary variables (X) by the general consumer price index (CPI)

$$\tilde{X} = \frac{X}{CPI/100}.$$

The CPI is divided by 100 as it is reported with baseline value 100.

Additionally, on every variable a logarithmic transformation was used. The reason for this is twofold. To

begin with it might further help with the heteroscedastic and serially correlated errors and therefore provide more reliant estimates. And it also helps to make the interpretation of the estimates more intuitively. One exception to this procedure is the share of working population, as this variable already presents a share of the whole population, a further transformation would make its interpretation meaningless.

5 Results

As previously outlined, the models presented in the table below were obtained using the general to specific model selection procedure. Starting from the model with all controls, statistically insignificant explanatory control variables were subsequently removed one at a time. During this process, the sign as well as the significance of our pollution variables were very sensitive to small changes in the model, i.e., we observed large changes in obtained estimates over a single iteration of the general to specific selection procedure. Accordingly, the presented results are to be interpreted cautiously. The unstable coefficient estimates are most likely caused by a combination of highly collinear explanatory variables and the absence of important variables that drive health care expenditures in our model.

Nevertheless, some key takeaways can be made. We observe that in our model, which does not contain spatially lagged variables, that Particulate Matter Emissions have a significant effect while in the SLX model this is not the case. This might be due to the mentioned multicollinearity problems, as all pollution variables exhibit a positive spatial correlation. The negative and significant effect of Sulphur Gas emissions, is most likely also a result of the multicollinearity in our model. Especially, given that the spatial lag features a positive and significant coefficient, there is no sensible interpretation to be made. The key takeaway should thus be, even though we find a positive and significant effect of Particulate Matter Emission (fGLS) on Health Care Expenditure, the result is both unstable over iterations in our model selection procedure and also across model types. Further, given our modelling and identification strategy, we may not claim that the results represent a causal relationship between the examined variables.

Table 1: Results fGLS & SLX

	<i>Dependent variable:</i>	
	fGLS	SLX
log(Disposable Income per Capita Rural)	1.072*** (0.067)	
log(Disposable Income per Capita Urban)		1.180*** (0.081)
log(Forest Coverage Rate)	0.490*** (0.102)	0.287** (0.144)
log(Urban Population)	0.308** (0.145)	0.504*** (0.135)
log(Waste Gas Emissions Nitrogen)	-0.045 (0.035)	0.006 (0.041)
log(Waste Gas Emissions Particulate Matter)	0.063*** (0.024)	0.032 (0.023)
log(Waste Gas Emissions Sulphur)	-0.007 (0.016)	-0.063** (0.028)
log(Waste Gas Emissions Nitrogen Lag)		-0.111 (0.088)
log(Waste Gas Emissions Particulate Matter Lag)		0.117*** (0.040)
log(Waste Gas Emissions Sulphur Lag)		0.093** (0.041)
<i>Note:</i>		
*p<0.1; **p<0.05; ***p<0.01		

6 Conclusion

This short research note on air pollution in china offers a concise overview of the relevant current literature as well as our analysis carried out utilizing an SLX panel model as well as a standard fixed effects panel model. Due to inconsistencies in the model selection procedure the obtained results are to be interpreted cautiously. The underlying cause of those inconsistencies are not entirely straight forward to identify, however, it is quite likely that they may be explained by a large part by two factors. First, our pollution variables measure emissions and not actual pollution, depending on the difference to the actual measured pollution that could be obtained by measurements on the ground, this inaccuracy might weaken the link between pollution and increases in Health Care Expenditure. Further, even though we included quite a few variables as controls, it is still quite likely that we are suffering to an omitted variable bias. Missing some important explanatory variables would also explain the large changes in estimates across a single iteration of our model selection procedure. When compared to the original paper that we had intended to replicate by **zeng2019** we noticed a few issues afterwards.

Die haben einfach ein SAR model durchgezogen obwohl es nicht machen sollen

Variable die drinnen sind doctors, hospital bed als x und es kann sein dass das hce auch beschreiben (sprich ein simultatety bias)

govenment expenditures -> HCE

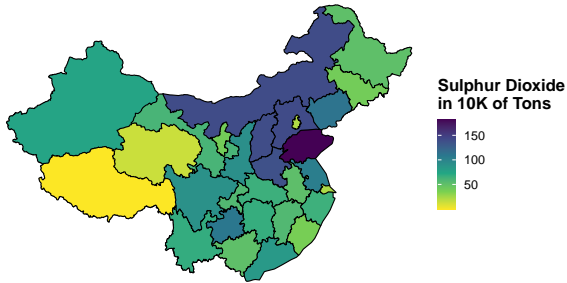
[compare and contrast results to literature]

A Appendix

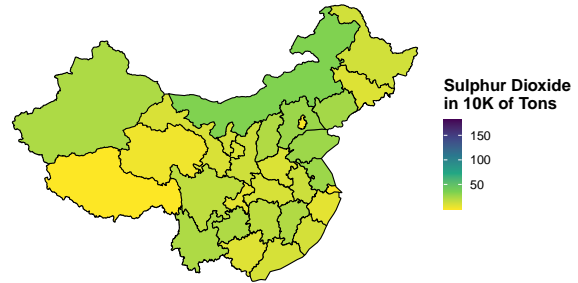
A.1 Further Plots

A.1.1 Sulphur Emissions

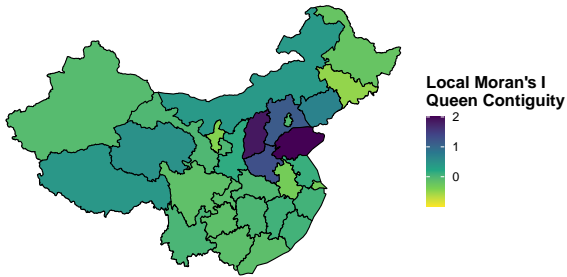
Sulphur Dioxide Emissions in: 2011



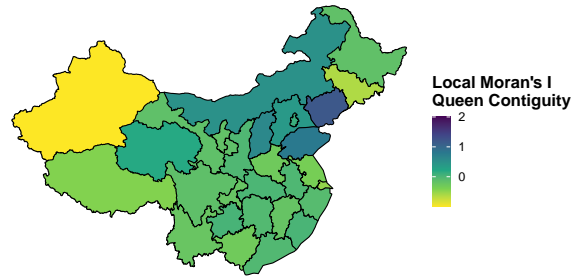
Sulphur Dioxide Emissions in: 2019



Local Moran's I for Sulphur Dioxide in: 2011

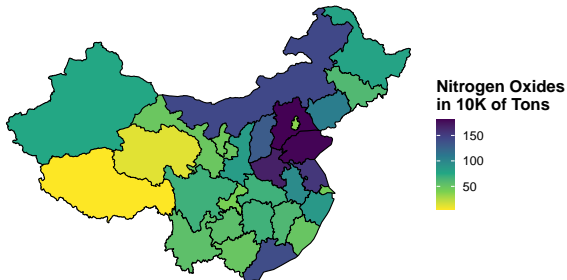


Local Moran's I for Sulphur Dioxide in: 2019

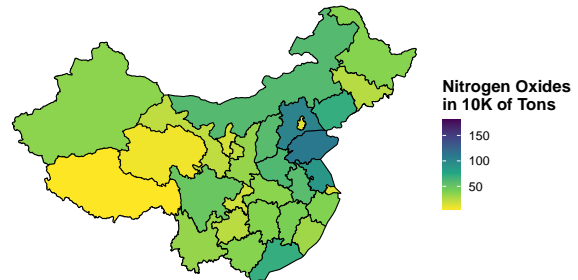


A.1.2 Nitrogen Emissions

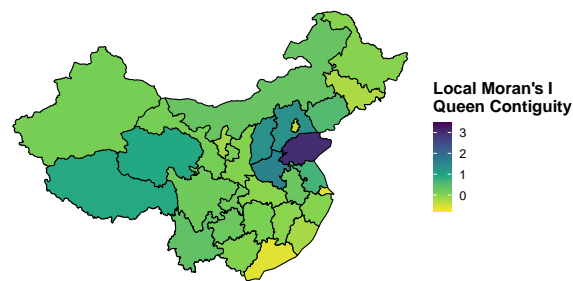
Nitrogen Oxide Emissions in: 2011



Nitrogen Oxide Emissions in: 2019



Local Moran's I for Nitrogen Oxides in: 2011



Local Moran's I for Nitrogen Oxides in: 2019

