Department of Economics, University of Warwick

Monash Business School, Monash University

as part of

Monash Warwick Alliance

Does Bad Air Quality Contribute to Obesity? Evidence from China's Central Heating System.

Yuxuan Ma

Warwick-Monash Economics Student Papers

December 2021 No: 2021-18

ISSN 2754-3129 (Online)

https://warwick.ac.uk/fac/soc/economics/research/wmesp/

The Warwick Monash Economics Student Papers (WM-ESP) gather the best Undergraduate and Masters dissertations by Economics students from the University of Warwick and Monash University. This bi-annual paper series showcases research undertaken by our students on a varied range of topics. Papers range in length from 5,000 to 8,000 words depending on whether the student is an undergraduate or postgraduate, and the university they attend. The papers included in the series are carefully selected based on their quality and originality. WM-ESP aims to disseminate research in Economics as well as acknowledge the students for their exemplary work, contributing to the research environment in both departments.

"*We are very happy to introduce the Warwick Monash Economics Student Papers (WM-ESP). The Department of Economics of the University of Warwick and the Economics Department at Monash University are very proud of their long history of collaboration with international partner universities, and the Monash Warwick Alliance reflects the belief in both Universities that the future will rely on strong links between peer Universities, reflected in faculty, student, and research linkages. This paper series reflects the first step in allowing our Undergraduate, Honours, and Masters students to learn from and interact with peers within the Alliance.*"

Jeremy Smith (Head of the Department of Economics, University of Warwick) and Michael Ward (Head of the Department of Economics, Monash University)

Recommended citation: Ma, Y. (2021) Does Bad Air Quality Contribute to Obesity? Evidence from China's Central Heating System. Warwick Monash Economics Student Papers 2021/18.

[1] Warwick Economics would like to thank Lory Barile, Gianna Boero, and Caroline Elliott for their contributions towards the selection process.

Does Bad Air Quality Contribute to Obesity?

Evidence from China's Central Heating System

Ma, Yuxuan (Sylvia) *

Abstract

This study finds that individuals exposed to an additional 1 $\mu g/m^3$ airborne particulate matter smaller than 2.5 lead to a statistically significant 0.121 kg/m^2 rising of body mass index. This positive relationship is identified by two-stage least square regression using a regression discontinuity estimator of air pollution generated by China's coal-burning winter heating policy, which only heats for northerners but not for southerners, as the instrument variable. This identification utilizing the quasi-experimental method of regression discontinuity design based on the difference of county's latitude from both parametric and nonparametric approaches, using different kernel types and bandwidth sizes, with 6000 observations in 2008. Further, the result shows that heating policy caused airborne particulate matter smaller than 2.5 and body mass index significantly increasing in the north and south divided line. These findings not only contribute to the identification of causality between air pollution and obesity but help guide social and environmental policy as well.

JEL Classifications: C54, I10, Q53

Keywords: Airborne particulate matter, Body mass index, China, Central heating policy, Regression discontinuity

---

* Contact: mayuxuan98@outlook.com
Link to online appendix (includes full variable descriptions, regression tables, robustness checks, and do file): https://github.com/sylvm98/Does-Bad-Air-Quality-Contribute-to-Obesity
Acknowledgements: The author would like to extend the sincerest gratitude to Thijs Van Rens for his supervision, unwavering support and insightful guidance on this dissertation.

1. Introduction

It is well known that obesity has become a worldwide issue in those years, which has nearly tripled since 1975. In 2016, more than 1.9 billion adults were overweight, and over 650 million of those were obese[1]. To respond to the issue, scientists are committed to discovering the causes of obesity. The fundamental cause of obesity is generally considered as an imbalance between calories[2] consumed - eating too much and expended - exercising too little. For the former, energy-dense foods that are high in fat and sugars, and hyper-processed foods that are several ingredients that are not often used in cooking has become increasingly popular with modern people; for the latter, physical activity levels are decreasing (An et al., 2018), because of technology development, for instance, the changing of transportation, and the changing sedentary mode of many works especially in the post-COVID-19 era. Besides, the consequence of obesity and overweight is the risk of having noncommunicable diseases (WHO, 2021), such as type two diabetes, heart disease and stroke[3], musculoskeletal disorders[4], and some cancers.

For the other potential causes of obesity, this research is curious about how air pollution contributed to body weight. We aim to identify the causal effect of air pollution generated by China's coal-burning heating policy, on obesity. The health science researchers have already proved the connections between air quality and body weight, which we will discuss in detail in the later section.

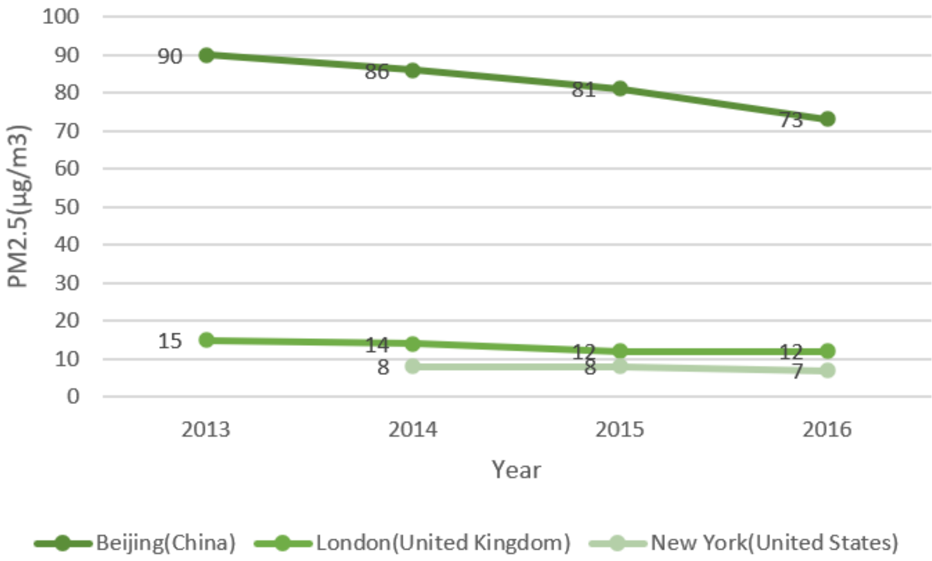
To first state the importance of study air pollution as a determinant, it has also become a major issue and poses a threat to public health. It kills an estimated 7 million people worldwide every year[5]. Figures 1 and 2 below are based on the database of WHO[6], which shows that air quality is notoriously terrible in, the world's most populous country, China. Ambient concentrations of particulate matter of less than 2.5 µm diameter ($PM_{2.5}$) or particulate matter of less than 10 µm diameter (PM_{10}) from 2013

---

[1] Obesity and overweight – World Health Organization (WHO), 2021
[2] The energy value of food is measured in units called calories.
[3] They were the leading cause of death in 2012 due to WHO.
[4] Especially osteoarthritis – a highly disabling degenerative disease of the joints.
[5] Air pollution - WHO, 2021
[6] Data source link: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/ambient-air-pollution

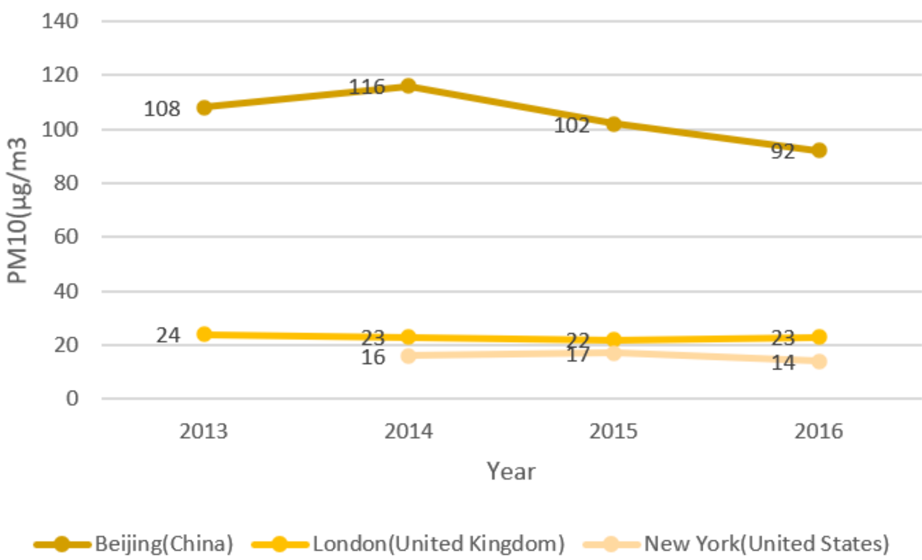
to 2016 in Beijing (China) are separately about 6 times or 4.5 times the level that measured in London (United Kingdom) and are even 10 times or 6 times compared to the New York (United States). For this reason, it is a good idea to focus on a study related to air pollution in China.

Figure 1: Changing of ambient $PM_{2.5}$ from 2013 to 2016



Note: Figure 1 plots $PM_{2.5}$, mean annual Particulate matter of less than 2.5 µm diameter for Beijing (China), London (United Kingdom), and New York (United States) the U.S. from 2013 to 2016. Their populations are separately 20383994, 10313307, and 8174962.

Figure 2: Changing of ambient PM_{10} from 2013 to 2016



Note: Figure 2 plots PM_{10}, mean annual Particulate matter of less than 10 µm diameter for Beijing (China), London (United Kingdom), and New York (United States) the U.S. from 2013 to 2016.

Furthermore, there is an interesting fact that air quality is especially poor in northern Chinese cities, such as Beijing[7], Tianjin. It is also frequently claimed that more people are overweight in northern than southern China[8]. Meanwhile, when it comes to the differences between north and south, the coal-fired heating systems of China are the poster child. It is because of the budget constraints, the Chinese government decided to provide centralized heating to the northern cities only in 1958. Northern and southern China are divided by the Huai River and Qinling Mountains because the average January temperature is roughly 0° Celsius along the line[9]. This arbitrary policy has caused plenty of particulate matter released every year in north China, which is the main reason for the more pollution of the north. Thus, it would be significant to estimate the effect of China's central heating policy on air pollution of the northern and southern areas, and even on obesity for the north and the south people.

We utilize the geographic regression discontinuity design, based on the difference of latitude from the dividend line of northern and southern China, to estimate the local average treatment effect of the central heating policy on air pollution and obesity. As the government provision of the heating north border is arbitrary, the quasi-experimental research can be set based on the policy, and both parametric and non-parametric approaches are applied for estimating in this quasi-random design. The fitted value for air pollution releasing by coal-burning of winter heating is then used as an instrumental variable identifying the causal effect of air pollution on body weight. By applying the two-stage least squares (2SLS) regression, we could better complete this causal inference and avoid the potential omitted variable bias in the ordinary least squares (OLS) approach.

The research uses the individual as the unit of the analysis. The data we use for body weight is the weight and height from the 2008 China General and Social Survey (CGSS), which is the earliest national, comprehensive, and continuous academic

[7] For example, Prime Minister Zhu Rongji quipped after following a career in Shanghai (the southern city) in 1999: "If I work in your Beijing , I will shorten my life at least five years" (The Economist, 2004, pp. 55-57).
[8] The great weight divide - Global Times, 2021
[9] This line is not a border for any other administrative purposes. So, it satisfies the important assumption of regression discontinuity design for this study. See more details in the section two.

survey in China. The 2008 CGSS provided detailed information on health and demographic data along with the information of county administrative code for 6,000 individuals from 100 different counties of China. We use the National Aeronautics and Space Administration (NASA) satellite-based annual average $PM_{2.5}$ measurements of each county in 2008 on air pollution. The final data set are a cross-sectional dataset with 6000 observations, matched the individual-level 2008 CGSS data with the count-level air pollution, the geographical data, and county demographic data from China County Statistical Yearbooks.

By combine RD design and 2SLS estimator, this study finds the Qinling Mountains and Huai River central heating system, which polluted air by burning fossil fuels like coal, had statically significant contributed to human body weight. Specifically, the heating policy caused $PM_{2.5}$ a significantly 3.163 $\mu g/m^3$ increase in the north of the border and a 0.461 kg/m^2 marginal significant improvement of body mass index (BMI) at the north and south divided line. What is more, the main result indicates that being exposed to an additional 1 $\mu g/m^3$ $PM_{2.5}$ contributes to a 0.121 kg/m^2 rising in BMI for people's obesity level.

Our study provides three primary contributions to the literature. Principally, the study contributes to literature identifying the causality between air pollution and obesity, which is quite little. The difficulties of identification are the endogeneity and the potential omitted variable bias, for both subjects are potentially correlated with economic indicators. To be more specific, air pollution generated by manufacturing factories, power plants, and automobile exhaust is the subsidiary product of the human economic activity, so it is highly related to the economic indicators, for example, food prices, income level, and so on, but they are the determinants of obesity as well. There is some health science literature, attempting to identify the causality between air quality and body weight by using regression models or other statistical methods. For example, Madrigano et al.(2010) examine the association of $PM_{2.5}$ concentrations and obesity or diabetes status, using the mixed regression models based on a longitudinal U.S. Normative Ageing Study with 809 participants and a total of 1819 observations. They finally find a positive correlation between air pollution and obesity. However, when associating air quality to obesity, these previous health science studies do not have clear identification strategies and especially have the potential of causing omitted variable bias. So, identifying this causality has been hard until Deschenes et al. (2019) first provided a study utilizing an instrument variable, which is thermal inversions[10], to estimate the impact of air pollution on obesity and overweight. Thermal inversion is a good instrument in that it is not correlated with any human activities and is deemed to be an exogenous variable with controlling the weather, seasonality in environmental and economic conditions. Their result shows an additional 1 $\mu g/m^3$ increase in the average annual $PM_{2.5}$ would increase the body mass index by 0.27 percentage points. As their study provides us a general test on whether air pollution is causally related to obesity, and test of few possible behavioral mechanisms, our goal is focusing on if being exposed to the air pollution generated by China's central heating policy contributed to an individual's level of obesity. In other words, our study focuses on the local average treatment effect of the heating policy on air pollution, rather than the average treatment effect of $PM_{2.5}$ on obesity.

While generally, most previous literature studied the causes of obesity from the economic variables. For example, the growing price disparity between healthy and unhealthy foods (Drewnowski and Specter, 2004) is considered as a determinant, and Currie et al. (2010) find that if a fast-food restaurant is located within 0.1 miles of school would result in a 5.2 percent increase in obesity rates among the ninth graders. Education is also considered as one of the determinants of obesity. Brunello et al. (2013) adopt a multi-country setup to estimate if education level has a causal protective effect on body mass index in nine European countries. They find no such effect for males but stronger among overweight females. Moreover, another determinant is income inequality (Monteiro et al, 2004). Popkin and Slining (2013) using quantile regression with a sample of eight countries suggest that levels of overweight and obesity in low- and middle-income countries significantly approach the levels found in higher-income countries.

On the other hand, by treating China's heating policy as an exogenous shock, this study will help policymakers linking air pollution to obesity. The China heating

[10] Thermal inversion is a deviation from the normal decrease in air temperature with increasing altitude. The air above the ground is warmer than the air below, so that the near side cannot convection and pollutants are accumulated.

system policy-related research started from Almond, et al. (2009), who evaluate the significant role of China winter heating in generating air pollution using a cross-sectional RD estimation approach. This dramatically higher air pollution level estimated in the north led by the heating policy is roughly 5-8 times the levels in the US. Then some researchers found that sustained exposure to coal-fired air pollution decreases life expectancy (Chen et al., 2013; Ebenstein et al., 2017; Fan et al., 2020). Ebenstein et al. (2017) find that a 10-$\mu g/m^3$ increase in PM_{10} reduces life expectancy by 0.64 years. They further suggest that bringing all of China into compliance with its Class I standards for PM_{10} would save 3.7 billion life-years. Moreover, Ito and Zhang (2020) have developed a standard logit model to estimate people's willingness of cleaning air which is measured by the demand for Air Purifiers, and an RD design to estimate the air pollution generated by central heating policy. Their results indicate that people would be happy to pay $32.7 every year for eliminating the policy-induced air pollution.

2. Body Weight, Air Quality, and Winer Heating in China

In this section, we provide the background information on the relationship between air pollution and obesity, and China's central heating policy, which are vital to our empirical framework.

2.1 Air pollution and Obesity

To further ascertain the meaning of this causality study, the health science study has proved two types of mechanisms between air pollution and obesity. One is the biological approach. Researchers find air pollution may cause metabolic disorders and chronic diseases, which are highly related to obesity. For instance, air pollutants are significantly associated with the risk of type 2 diabetes mellitus (Rajagopalan and Brook, 2012; Andersen et al., 2012). Being long-term exposure to air pollutants induces an inflammatory response in the lung and visceral adipose tissue, insulin resistance, and ultimately causes type 2 diabetes mellitus (Meo et al., 2015). Diabetic people generally have higher BMI, waist-to-hip ratio, and alcohol and fat intake. Schwartz and Porte (2005) find the connection between obesity and type 2 diabetes controlled by the neuronal systems. Similarly, $PM_{2.5}$ exposure contributes to cardiovascular according to the American Heart Association scientific statement on "Air Pollution and Cardiovascular Disease" in 2004 and strengthen by Brook et al. (2010). Cardiovascular is also suggested associated with body weight (Romero-Corral et al., 2006).

The other access is the behavior channels, such as sleep disorder (Heyes and Zhu, 2019; Keith et al., 2006) caused by breath in pollutants and sedentary lifestyle caused by air pollution would reduce the net calories expended, which increases obesity risk (WHO, 2021). Air pollution could even lead to a direct increase in calories consumed. It has the potential of getting people depression or anxiety, and these mental illnesses release the appetite-related hormone in the human body (Chen et al., 2018), which gives people an extremely enormous appetite and causes them to load up on fats.

2.2 The Central Heating Policy and its Development

The Chinese government established the free heating entitlement for households and units by providing boilers with free coal fuel in winter during the central planning period from 1950 to 1980. The burning of coal in boilers contributes to the release of air pollutants, especially the total suspended particulates, which measure the mass concentration of particulate matter in the air. But not for long, the government decide to provide city-wide centralized heating to northern cities only because of the budget constraints (Almond et al., 2009). Especially, northern and southern China are divided by a line formed by the Qinling Mountains and Huai River in central China as shown in Figure 3, because the average January temperature is roughly 0° Celsius along the line, and the line is not a border for other administrative purposes.

Figure 3: Arbitrary Border for Provision of China Central Heating System[11]

Note: Figure 3 shows China is divided by the Qinling Mountains and Huai River, which is used as the arbitrary border for the central heating policy.

From then on, cities in the north of the river have received centralized heating relies on coal-fired heating systems, of which two-thirds of heat is generated by heat-only hot water boilers for one or several buildings in an apartment complex, and the

[11] Graph source: Qinling Mountains: The central park of China – CGTN News, 2021.(Link: https://news.cgtn.com/news/2020-04-21/Qingling-Mountains-The-central-park-of-China-PRUrshCKWs/index.html)

remaining one-third is generated by combined heat and power generators for the larger areas of each city (Ito and Zhang, 2020), in every winter.

For the following years, China has transited to a market economy, thus the heating reform took place in 2003. The payment system of winter heating is changing from free provision to flat-rate billing, according to the World Bank (2005)[12]. The so-called flat-rate billing is a fixed charge per square meter of floor area, regardless of actual heating usage. From then on, households in the north were responsible for the payment of their heating bills, but the government would provide subsidies directly to the heating companies in the northern area.

[12] World Bank (2005). China heat reform and building energy efficiency project. Technical report, World Bank, Washington, DC.

3. Estimation Framework

In this section, we first provide the data source, describe the measurement of the variables, and summary the statistics. After that, we detailed introduce the empirical methodology of this study.

3.1 Data Sources and Measuring

The cross-sectional dataset of our research is integrated from four data files, which are county-level air pollution data, individuals' health and education data, geographic data, and demographic information for each county. This subsection describes each data source, variable measuring and provides the summary statistics.

3.1.1 Air Pollution

To ascertain the valid China air pollution data for studying, there are mainly two sources. One is the ground-based Air Pollution Index (API)[13] and Air Quality Index (AQI)[14] data from the China National Environmental Monitoring Center (CNEMC), but most archived observations are not publicly available and the available only cover a few cities. To compensate, real-time data is self-reported by Chinese cities on the local websites of the Ministry of Environmental Protection (MEP), but Ghanem and Zhang(2014) proved that the data is manipulated. Thus, we do not use this ground-based pollution data.

The other source of air pollution data is the satellite-based Aerosol Optical Depth (AOD)[15] retrievals. AOD data can take more sampling units than the MEP, which is closer to the data of the CNEMC (Chen et al, 2013). So, we derived the county-level annual mean particulate matter less than 2.5 micrometers ($PM_{2.5}$) from the Atmospheric Composition Analysis Group at Dalhousie University, which estimates[16]

---

[13] For China, the reference standard for API classification calculation is GB 3095-1996 "Ambient Air Quality Standard" (obsolete). The only pollutants evaluated are SO_2, NO_2 and PM_{10}, which are published once a day.

[14] The reference standard for AQI grading calculation is GB 3095-2012 "Ambient Air Quality Standard" (current). The pollutants involved in the evaluation are SO_2, NO_2, PM_{10}, $PM_{2.5}$, O_3, CO, etc., which are published every hour.

[15] Aerosol Optical Depth refers to a dispersion system formed by solid or liquid particles stably suspended in a gas medium. AOD can be converted into particulate matter for air pollution research.

[16] The annual mean $PM_{2.5}$ concentrations is estimated at $0.01° \times 0.01°$ (approximately 1 km × 1 km) spatial resolution globally using a Geographically Weight Regression with an out-of-sample cross-validated R^2 of 0.81. The regression incorporated satellite data, simulated aerosol composition, and

$PM_{2.5}$ by the AOD data obtained from the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) of the National Aeronautics and Space Administration (NASA).

3.1.2 Obesity

The data on obesity is from the 2008 Chinese General Social Survey (CGSS), which is publicly available on China National Survey Data Archive (CNSDA).[17] The 2008 CGSS is based on a sample with 6000 observations from 100 different counties extracted. In other words, the survey questionnaires 60 individuals in each sampling county. It provides interviewee self-reported body weight and height which shall be used to calculate body mass index(BMI)[18] through the following formula to measure human body obesity level,

$$BMI = \frac{Weight(kg)}{[Height(m)]^2}$$

The samples of CGSS were selected using a random cluster, and stratified probability proportional to size (PPS) sampling method, but the detailed design varies each year. The county identifier[19] of individuals are disclosed only in 2008[20], which is vital for this research design to match with the county-level database of air pollution. The location of each county in the map is reported in Figure 4 below.

3.1.3 The geographic data

We make four geographic variables based on the county locations and north-south dividing line in our empirical analysis. The line of Huai River and Qinling Mountains is approximately 33°N. The latitude and longitude of each sample unit's centroid shall

land use information. The population-weighted annual mean $PM_{2.5}$ concentrations were assigned to each postcode area.

[17] The CGSS is the earliest national, comprehensive, and continuous academic survey project in China mainland. It is jointly conducted by the Hong Kong University of Science and Technology and Renmin University since 2003, over the two periods of 2003–2008 and 2010-2019.

[18] Body mass index (BMI) is a simple index of weight-for-height that is commonly used to classify overweight and obesity in adults.

[19] It is a six-digit code. China has three administrative levels, namely, provinces/municipal cities (two-digit code), prefectures/cities (four-digit code), and counties/districts (six-digit code).

[20] Since opening for academic use, CGSS decides not to disclose the names, geographic locations, and zip codes of counties to protect the privacy of its subjects. However, in 2008, not like the other years using province as primary sample unit (PSU), it adopts counties as PSUs. So, the county identifier can be disclosed in 2008.

be obtained from Google Map manually, and they arrange as shown in Figure 4 below. We calculate the vertical difference of latitude (Degree North) between counties centroid and the Qinling Mountains and Huai River. Northern counties which received central heating are marked in red, while southern counties are marked in blue, and the location of the north-and-south divided line is also marked in Figure 4. We also derived the area of each county from the China County Statistical Yearbooks.

Figure 4: The County Sample Units Distribution of 2008 CGSS

Note: Figure 4 plot the distribution of sampling counties.[21] The first period CGSS uses the map drawing method for field sampling. It might be the reason for Qinghai Province, Tibet Autonomous Region, Hong Kong Special Administrative Region, Macau Special Administrative Region, and Taiwan Special Administrative Region are not concluded in PPUs, that they have either difficult geographical conditions or huge cultural and economic differences, although there is no official explanation in the supplementary document.

[21] Map data ©2021 Google, SK telecom

3.1.4 Covariate Variables

The demographic and county characteristics data shall be derived from the 2008 China County Statistical Yearbooks. And individual heterogeneity controls are from the 2008 CGSS survey. For more details, see the summary statistics and the variable description of A. Table 1 and A. Table 2 in the Appendix.

3.1.5 Final Merged Dataset

The heart of the matching process is to merge the individual-level obesity data with the county-level air pollution. We use regression models estimated at the individual level to ensure maximizing the number of observations and use of individual local average treatment effects (LATE). Our final datasets are merged in the following manner. First, checking for county administrative codes in the air pollution dataset and obesity dataset are duplicates for each county. Second, modifying the changed identifier and drop the data of the canceled county division. Third, the obesity dataset and the air pollution dataset are merged by individual county codes in 2008. Finally, repeat the above process for adding county controls. The final dataset is a cross-sectional dataset for 6000 individuals across 100 counties in 2008. For the details, see the summary statistics of A. Table 1 in the Appendix.

3.2 Methodology

This study utilizes two methods to identify the relationship between air pollution and obesity. One is the regular way of using ordinary least square (OLS) regression, the other is applying a sharp[22] regression discontinuity (RD) design to estimate the impact of the heating policy on air pollution and then using a two-stage least square (2SLS) regression for the casual identification.

At the first stage, we simply run the following OLS equation:

$$Y_{\mathrm{ic}} = \alpha + \beta PM_c + X_{\mathrm{ic}}\mu + Z_c\omega + \varepsilon_{\mathrm{ic}} \text{ , (1)}$$

[22] We do not consider counties in the south near the border, receiving the central heating. Because study will transfer to a fuzzy RD design (as the cut-off point will no longer be 33° N). Besides, in 2008 there is few southern counties receive winter heating, especially for our observations.

where i refers to individuals, c refers to counties. Y_{ic} represents the BMI of individual *i* in county *c* in 2008, and PM_c is the annual mean $PM_{2.5}$ of county *c* in 2008. X_{ic} and Z_i are vectors of the observable characteristics of the individual *i* in county *c* and county *c* that might influence health outcomes other than air quality separately. ε_{ic} is the error term. Standard errors are clustered at the county level.

Secondly, we built geographic RD designs for obesity and air pollution to estimate the effect of the heating policy. RD design is a quasi-natural experiment, which is good at counteracting the endogeneity in OLS regression and allowing estimation of the LATE. Dell (2010) pioneered the introduction of geographic boundary into the RD problem, that is, setting the geographic division as a breakout point and the geographic distance as the assignment variable[23]. Similar to our paper, the difference of latitude of each county to the boundary line is the forcing variable. The longitude is adding to the covariates to allow counties in the same latitude level could vary from each other.

We set D_c as a treatment for individuals receiving central heating(in the north), which following:

$$D_c = \mathrm{f}(x_c) = \begin{cases} 1, & x_c\text{-}33°N > 0 \\ 0, & x_c\text{-}33°N < 0 \end{cases},$$

where $(x_c\text{-}33°N) \in [-10.47, 16.19]$ is the difference to the border, x_c is the latitude of the county *c*, and 33°N is the latitude of the North-south dividing line.

Assume any unobserved determinants of air pollution or bodyweight change smoothly as they cross the north-south dividing line, and the treatment will have a discontinuous effect on air pollution and body weight for individuals. To estimate the effect of heating policy on air pollution and obesity, we built the first approach of parametric estimation[24] as follows,

[23] Also called forcing variable or running variable.

[24] The linear and polynomial fit of order n (n⩾1).

$$PM_c = \alpha + \delta_{PM} D_c + \sum_{n=1}^{n} \beta (x_c - 33°N)^n + \sum_{n=1}^{n} \gamma (x_c - 33°N)^n D_c + Z_c \omega + \varepsilon_c \text{, (2.1)}$$

$$Y_{ic} = \alpha + \delta_{\text{Fat}} D_i + \sum_{n=1}^{n} \beta (x_i - 33°N)^n + \sum_{n=1}^{n} \gamma (x_i - 33°N)^n D_i + X_{ic} \mu + Z_c \omega + \varepsilon_{ic} \text{, (3.1)}$$

where δ_{PM} and δ_{Fat} represents Local average treatment effect (LATE) at the cut-off in air pollution and obesity separately due to the coal-burning of heating system.

Another approach we apply is the non-parametric estimation. The parametric strategy allows all the observations to be used in the analysis, however, for RD design only the observations near the threshold are useful for causal inference. The local linear regression is applied then (Imbens and Lemieux, 2010), which sets the bandwidth to change the observations used to those around the threshold and has the benefit that no functional form needs to be imposed on the data. The non-parametric RD regression equations are built as:

$$PM_c = \alpha + \delta_{PM} D_c + \beta (x_c - 33°N) + \gamma (x_c - 33°N) D_c + Z_c \omega + \varepsilon_{\text{c}} \text{, (2.2)}$$

$$Y_{ic} = \alpha + \delta_{\text{Fat}} D_c + \beta (x_c - 33°N) + \gamma (x_c - 33°N) D_c + X_{ic} \mu + Z_c \omega + \varepsilon_{ic} \text{, (3.2)}$$

Finally, considering if the heating policy only influences obesity through its impact on air pollution, treat Eq. (2) as the 1st stage and treat LATE of $PM_{2.5}$ as an instrument variable (IV). Use the quasi-experimental fitted value of $PM_{2.5}$ generated by the winter heating policy from Eq. (2) to run the 2nd stage equation,

$$Y_{ic} = \alpha + \beta PM_{\text{c}} + \beta (x_c - 33°N) + \gamma (x_c - 33°N) D_c + X_{\text{ic}} \mu + Z_c \omega + \varepsilon_{ic} \text{, (4)}$$

where PM_{c} is the fitting value from Eq. (2).[25] Comparing the 2SLS approach of Eq. (4) with OLS of Eq. (1), it has the advantage of solving the omitted variables bias associated with identifying the causality between air pollution and obesity. Moreover, it controls the other threats to identification, for instance, the potential measurement error for $PM_{2.5}$[26].

[25] Choosing the linear fit (2.1) or the optimal polynomial in degrees (2.2).

[26] Like the attenuation bias associated with the mismeasurement of $PM_{2.5}$ because we are using the satellite retrieval data, not the direct ground-based pollution data.

4. Empirical Analysis and Results

We first focus on the air pollution generated by China's central heating, which is measured by a sharp regression discontinuity design using the difference of latitude from county centroid to the arbitrary north and south divided line as the instrument variable. Then we use two-stage least square regression to identify how the air quality contributed to human body weight. The main results are reported and discussed in this section. Further, we do test the robustness of our empirical framework.

4.1 The Validity of the Central Heating RD Design

To borrow greater support to emphasizing the application of the RD design and helping us better trust its results, this research checks the validity of applying central heating policy RD design on air pollution and obesity at first. Besides, we will discuss more from several other ways of robust check in the later sub-section.

4.1.1 visualizing the discontinuity

We first test if the central heating policy caused discontinuous changes in $PM_{2.5}$ and BMI to the north of the border. We graph the scatters of $PM_{2.5}$ and BMI on the forcing variable, but they are not intuitive enough to see the jumps. See the A. Figure 1 and A. Figure 2 in the Appendix. Then we draw the histogram-style conditional mean graphs of $PM_{2.5}$ or BMI exposing or locating on the degrees north from the boundary. As shown in the linear fitted Figure 5 and Figure 6 below, now it is clear to see that the jumps do exist for both two main outcome variables, air pollution, and obesity. Apart from the line of the best-fit graphs, see the graphs of quadratic of best fit with the confidence interval of $PM_{2.5}$ and BMI conditional on degree north from the Huai River the A. Figure 3 and A. Figure 4 in the Appendix.

Figure 5: Fitted values from a linear regression of $PM_{2.5}$ on Degrees North from the North-South line

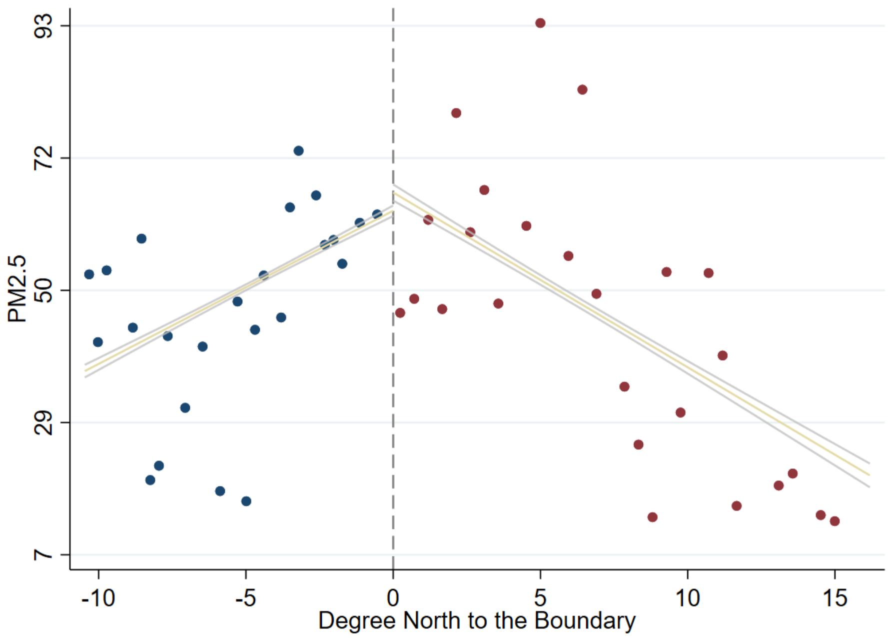
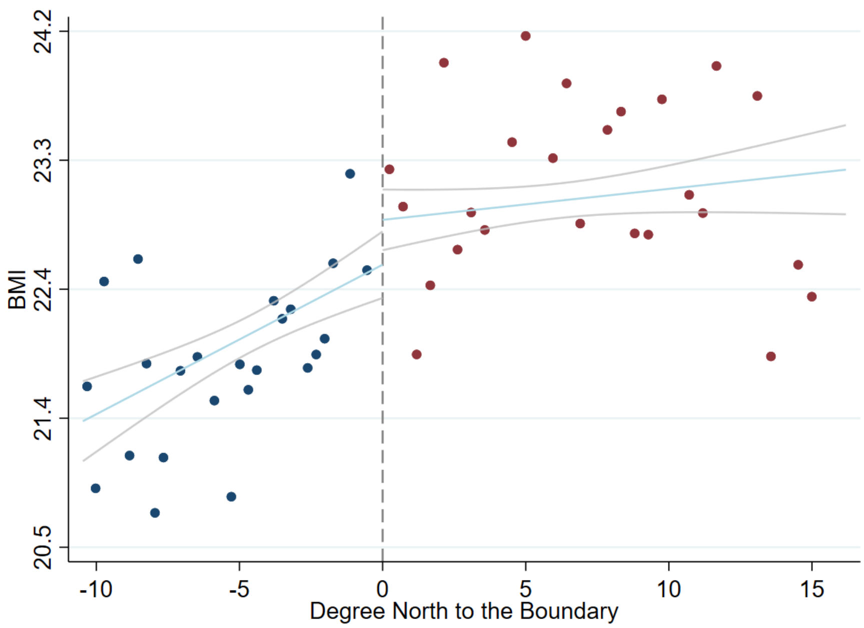


Figure 6: Fitted values from a linear regression of BMI on Degrees North from the North-South line



Note: The discontinuity is defined when the distance from the location of observations to the boundary is 0 (when the latitude is 33 degrees north). The linear fitted value of $PM_{2.5}$ or BMI on the difference of latitude from the boundary is estimated respectively on each side of the division line with 34 bins and confidence interval.

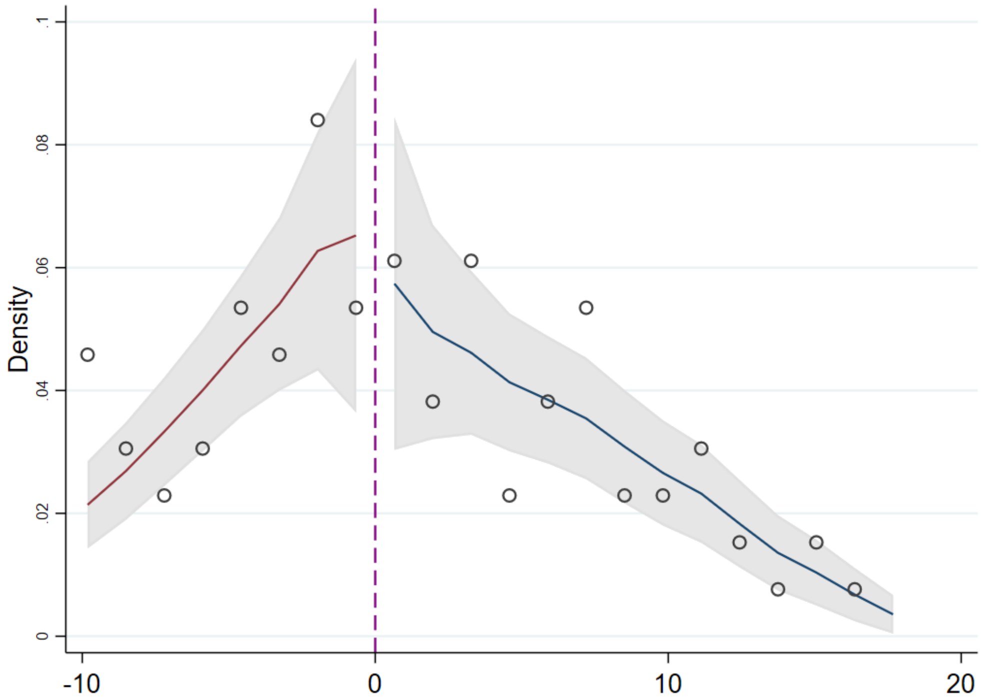
4.1.2 Covariates are not discontinuous at the threshold

The most vital identifying assumption of the RD design, that any unobservable determinants change smoothly at the boundary, is difficult to test. Nevertheless, as we mentioned above, China government provision of central heating north of the Qing-ling Mountains and Huai River, the precise location of which is arbitrary. We could roughly assume this assumption is satisfied. While we can still test whether the observable determinants are changing smoothly at the north-south dividing line. This is analogous to the placebo test using covariates as placebo outcomes and building RD regressions to see if the covariates are discontinuous at the threshold. In the Appendix, A. Table 3 reports the results of the non-parametric LLR approaches to summarize the differences in the covariates between the north and south of the river, whereas A. Table 1 provides the full set of individual and county covariates. As the table shows, the robust bias correlated point estimates tend to indicate that individuals to the north have the higher year of education, and the smaller year of birth, sex ratio than the southerners; also, counties to the north have higher area, general expenditure, number of employees, and smaller population, deposits, fixed, and longitude assets than the south area. But none of these tests could be judged statistically significant by 0.01 significant level. To conclude, the null hypothesis of the placebo outcome test is not rejected, indicating there are no discontinuities in these control variables on each side of the river.

4.1.3 No Manipulation over the Assignment Variable

Generally, RD can be invalid if the assignment variable can be precisely manipulated. While our forcing variable, the latitude of county centroids, is depending on the sampling methods of CGSS. That is saying if the number of observations near the cut-off is similar, there is no manipulation. We draw the histogram of the forcing variable(latitude difference) with a bin width of one. See the histogram of A. Figure 5 in the Appendix, while not intuitive enough, the amounts of observation are about the same on both sides of the border. Another method is testing whether the distribution of it is discontinuous at the border (McCrary, 2008). Figure 7 below shows the density plot.

Figure 7: McCrary Density Plot of the Assignment Variable



Although there is a jump in the density function on the threshold, the confidence interval overlaps, indicating that the distribution of the assignment variable on both sides of the cut-off is a continuous function. And the discontinuity estimate on the threshold [p=0.86] is not significant[27], which proves that there is no manipulation over the differences in the degree north of latitude to the border.

4.2 The Effect of the Heating Policy on $PM_{2.5}$ and BMI

The main RD results of how the heating policy impacts air pollution and obesity are summarized separately in table 1 below. We mainly focus on the LATE of $PM_{2.5}$, because it will be used as IV in the later casual inference process. In table 1, Columns (1) and (2) report the linear estimates of parametric RD design with the full sample size, following Equation (2.1) and Equation (3.1). We report both regression results including or excluding the available covariates and the standard errors are clustered at

[27] The estimated log difference in height [theta] is -0.08, standard error is 0.48.

the county level. Details of controls estimates and results for more functional forms can be found in the Appendix, A. Table 4. As the results show, the Qinling Mountains and Huai River policy increases $PM_{2.5}$ concentrations by 3.163 μg/m^3 (95% CI = 1.5 - 4.8) or 2.964 μg/m^3 (95% CI = 1.2 - 4.7) at the border with or without the covariates. Both results are significant at the 1% level. White Columns (3) and (4) represents the local linear regression of non-parametric RD design, following Equation (2.2) and Equation (3.2) with a uniform kernel and the optimal bandwidth choice (Imbens and Kalyanaraman, 2012) with the county-level observations. The local linear regression $PM_{2.5}$ estimate with covariates shows an unexpected decline of 18.94 μg/m^3 to the north of the border, but it is only a marginal significance at the 10% level with the covariates. Thus, we decide to use the statistically significant linear fit of parametric RD estimates with covariates[28], which is a striking discrete jump of 3.163 μg/m^3 to the north of the mountains and the river, as IV for the later 2SLS regression.

Table 1: RD Estimates of Central Heating Policy

Outcome	(1)	(2)	(3)	(4)
$PM_{2.5}$	2.964***	3.163***	-24.98***	-18.94*
	(0.8867)	(0.8155)	(9.0195)	(10.3495)
N	5940	5940	99	99
adj. R^2	0.253	0.420		
BMI	0.321	0.461*	-0.543*	0.297
	(0.2603)	(0.2667)	(0.3176)	(0.3669)
N	6000	5491	6000	5491
adj. R^2	0.021	0.066		
RD Type	Polynomial	Polynomial	LLR	LLR
Function	linear	linear		
Kernel			Uniform	Uniform
Controls	No	Yes	No	Yes

Standard errors in parentheses: * $p < .1$, ** $p < .05$, *** $p < .01$

Note: In columns (1) and (2), we report RD estimates of the coefficient on $PM_{2.5}$ and BMI, using different degree latitude from the Qinling Mountains and Huai River and a linear polynomial in degrees latitude from the border interacted with a policy dummy variable, following the Equation (2.1) and Equation (3.1) with the full sample. All standard errors in columns (1) and (2) are clustered at the county level. Columns (3) and (4) presents estimates from local linear regression (LLR), following Equation (2.2) and Equation (3.2) with uniform kernel and bandwidth selected by the method proposed by Imbens and Kalyanaraman (2012). Results are based on the full sample and (2) and (4) include the covariates listed in A. Table 1. Note that, for $PM_{2.5}$ columns (3)and (4) are based on 99 counties observations, and (1) and (2) are based on 5940 individual observations, so estimates of (3) and (4) are higher than (1) and (2).

To avoid repetition, RD estimates of BMI are similar to $PM_{2.5}$. It reports a 0.461

[28] Although RD design do not rely on covariates, we still add the controls in the regression to better compare the results with the OLS approach.

kg/m^2 rising of BMI for individuals to the north of the boundary for the full sample parametric approach, which is significant at the 10% level. In general, the effect of the coal-burning heating policy on body weight is not so significant as the local effect on air pollution. However, this will not affect our subsequent study on the causality between air pollution and obesity. For more functional forms for the parametric different kernel forms and optimal bandwidths for the non-parametric approach, they are reported at the table in A. Table 5 in the Appendix.

4.3 The Effect of Air Pollution on Body Weight

The estimates of the effect of $PM_{2.5}$ on obesity are reported in Table 2 below. Column (1) and (2) reports the OLS estimates results and Column (3) and (4) provides the parametric linear fitted RD estimates of $PM_{2.5}$'s effects (2SLS) estimates. Note that we've matched the individual-level dataset with the county-level dataset, which would cause the regression errors correlated within groups. In other words, the assumption of independent errors will be violated, and the unadjusted OLS and 2SLS standard errors would have a substantial downward bias. So, the standard errors are clustered at the county level to solve this issue and allow a valid inference.

Table 2: Comparing OLS and RD estimates of $PM_{2.5}$'s impact on BMI

BMI	(1)	(2)	(3)	(4)
$PM_{2.5}$	0.112*	0.121**	0.00995**	0.0138***
	(0.0616)	(0.0517)	(0.0046)	(0.0036)
Method	OLS	OLS	2SLS	2SLS
RD Type			Polynomial	Polynomial
Controls	No	Yes	No	Yes
N	5940	5431	5940	5431
adj. R^2	.	.	0.003	0.069

Standard errors in parentheses: * $p < .1$, ** $p < .05$, *** $p < .01$

Note: In columns (1) and (2), we report OLS estimates of the association between $PM_{2.5}$ and BMI, following the equation (1). In columns (3) and (4), we report the 2SLS estimates using the difference of degree latitude from the Qinling Mountains and Huai River as the IV and a linear polynomial in degrees latitude from the border interacted with a policy dummy variable, following the equation (4). Results are based on the full sample and (2) and (4) include the covariates listed in A. Table 1. All standard errors are clustered at the county level.

The OLS approach infers that an additional 1 μg/m^3 $PM_{2.5}$ increases BMI by 0.014 kg/m^2, which is significant at the 1% level. The result of the RD estimates of PM's effects shows a substantially larger increase of 0.121 kg/m^2 in BMI than the OLS estimate, which is significant at the 5% level. The full set results of the OLS and 2SLS approach which conclude the covariates are shown A. Table 6 and A Table 7 in the Appendix.

To compare OLS and 2SLS results, the larger number of the RD estimates of the $PM_{2.5}$ effects infers that the OLS approach has some omitted variables bias or measurement error as we mentioned before when inferring the causality. The 2SLS result is more valid than OLS, and the 2SLS estimator is considered as the closest one to the true effect of being exposed to $PM_{2.5}$ from the heating policy on the BMI changing.

4.4 The Robustness Checks

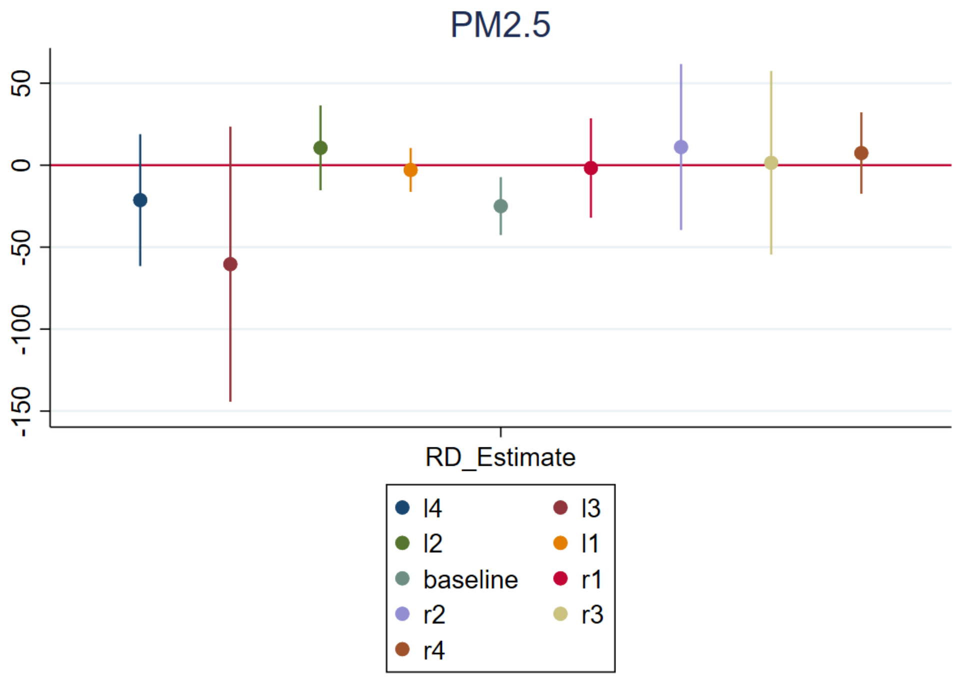
In this sub-section, we continued to take several tests checking the robustness of RD design for our instrument (air pollution) and causal inference process (2SLS).

4.4.1 Placebo Different Threshold Level

One of the RD design tests checking the robustness of the local randomization assumption is doing the threshold placebo test. We set 8 quintiles[29] of the different ranges on both left and right of the threshold as the new cut-off points and use the non-parametric approach to test whether the discontinuity exists significantly. We only discuss the validity of the $PM_{2.5}$ estimator here, because not only do we use it as IV in the causal inference process, but BMI non-parametric results are marginally significant. Figure 8 below integrates the new cutoffs of the RD estimates results on $PM_{2.5}$ and their confidence intervals. We still report A. Figure 6 of the BMI placebo test and discuss it in the Appendix.

[29] 20%, 40%, 60%, and 80% of the minimum (-10.47)and maximum(16.19) of the difference from county centroids to the border $\left(x_c\text{-}33°N\right)$.

Figure 8: Placebo Test for Different Thresholds



Note: The baseline point is the real cut-off of zero, l or r indicates the new cut-off is on the left or right of the threshold, and the following number is the number of the quintile. For instance, 'l4' indicates the new cut-off is the 80 percent point on the left of the minimum difference from county centroids to the border.

As shown in the graph below, only RD estimate with real cut-off point is significant from zero. The placebo cut-offs' RD estimates are not significant, for their confident intervals all go through zero. To conclude, by changing the cut-off point, we don't find any local treatment effect for air pollution. The RD design of air pollution is valid.

4.4.2 Week Instrument Test

To ensure the robustness of the 2SLS approach, we check the F-statistics to test the first stage least square regression for the significance of excluded the instruments variable, the difference to the border. If the first-stage F-statistics is smaller than 10, this detects the presence of a weak instrument (Olea and Pflueger, 2013).

Our F-statistics at the first stage regression with the full set of available covariates and clustered standard errors at county-level is $F(13, 5417) = 617.11$, which is much bigger than 10. Moreover, the coefficient on the difference to the border is significantly different from zero. We might be concerned that the weak instrument test is rejected.

4.5 Further Discussion

To further discuss our results, we provide the policy implications. On the one hand, keep intensifying efforts to prevent and control $PM_{2.5}$ air pollution is vitally important. In 2013, the State Council of China promulgated the Air Pollution Prevention and Control Action Plan, which set specific requirements for the prevention and control of air pollution in various regions based on the concentration of particulate matter. China's $PM_{2.5}$ concentration has continued to decline in recent years, and air quality has improved significantly. However, $PM_{2.5}$ concentrations of China from 2013 to 2016 are still roughly five times the value of the other developed countries as shown in Figure 1 and Figure 2. Due to the extremely low threshold for the impact of $PM_{2.5}$ on health, the WHO provides a guideline value of 10 $\mu g/m^3$ for the average annual concentration of $PM_{2.5}$, and China's current $PM_{2.5}$ standard has still a long way to go from the WHO guidance value. In other words, with the deepening of the pollution control process, the reduction of pollutants has gradually narrowed, and the task of reducing emissions remains arduous. Compared with transportation and power, the implementation of emission reduction policies for industrial sources faces more difficulties. This is because, firstly, coal-fired boilers and Industrial production equipment are more dispersed; secondly, the profitability of heavy-polluting industries, such as steel and cement, is low and the ability to pay for environmental protection equipment is limited; thirdly, regulatory policies have been frequently introduced in 2013 that companies and government lacking capability to make the coal shift towards cleaner sources of energy. Cheng, et al. (2021) suggest increasing the intensity of source control, and the proportion of renewable energy, and accelerating the clean replacement process of coal is also important.

On the other hand, the increasing prevalence of overweight and obesity and the associated economic burden needs to be resolved urgently. China has implemented many policies prevention and controls obesity. Wang et al. (2019) suggest government to responsible for enhancing cross-sector collaboration, which includes integrating obesity prevention and control into government mandates and the work of the relevant authorities, improving the nutrition policy system, and providing individual-level counseling and guidance on obesity treatment. With the development of national or local surveillance systems, the obesity issue can be better controlled.

To sum up, our research suggests that reducing air pollution may be an effective strategy to reduce overweight and obesity in China and could have large benefits in terms of avoided health expenditure on overweight and obesity.

5. Concluding Remarks

To conclude, our study suggests that the Qinling Mountains and Huai River central heating system, which polluted air by burning fossil fuels like coal, had contributed to human body weight. For these estimated local average treatment effects, the heating policy caused $PM_{2.5}$ a significantly 3.163 μg/m^3 increase in the north of the border, and a 0.461 kg/m^2 marginal significant improvement of body mass index (BMI) at the north and south divided line. What is more, the main result indicates that being exposed to an additional 1 μg/m^3 $PM_{2.5}$ contributes to a 0.121 kg/m^2 (95% CI = 0.01 - 0.23) rising in BMI for people's obesity level.

We follow Deschenes et al. (2019) using two-stage least square regression to identify how the air quality contributed to human body weight, but we focus on the air pollution generated by China's central heating, which is measured by a sharp regression discontinuity design using the difference of latitude from county centroid to the arbitrary north and south divided line as the instrument variable. Our contributions can be further discussed as follows,

Primarily, we identify the statistically significant positive association between air pollution and obesity. Our casual inference strategy utilizing the geographic regression discontinuity regression and two-stage least square regression is close to the natural experiment and helps us avoid the potential omitted variable bias. The OLS results also suggest a statistically significant positive causal relationship between air pollution and obesity, but because of endogeneity concerns, the magnitude of these results cannot be interpreted as real. By applying the individuals' self-reported body weights and heights data from the 2008 CGSS and merged with the county-level $PM_{2.5}$ data derived from the Atmospheric Composition Analysis Group at Dalhousie University, we get a final database which is a one-year, individual-level and cross-sectional database with six thousand individual observations in 2008. Both parametric and non-parametric of our RD approach estimators of the air pollution are robust to various econometric specifications, and the instrument variable we used in the 2SLS process rejects the weak instrument test. Thus, the results are reliable with minor errors which enrich the literature of considering air pollution as the determinant of obesity.

Further, by treating China's heating policy as an exogenous shock, this study provides evidence for the policymaker. Developing countries generally have seriously poor air quality. This is often regarded as one of the primary obstacles to economic development. China has promulgated many acts and regulations to reduce air pollution, such as the most recent "Three-year plan of action for winning the war to protect blue skies" implemented by the Ministry of Ecology and Environment in 2018. However, the reduction of pollutants has gradually narrowed, and the task of reducing emissions remains arduous with the deepening of the pollution control process so that increasing the intensity of source control, the proportion of renewable energy, and accelerating the clean replacement process of coal are urgent. On the other hand, the increasing prevalence of obesity and the economic burden caused by it also urgently need to be resolved. In Short, our results provide an effective strategy to reduce overweight and obesity, which is to decrease air pollution. This will also help generate huge benefits in terms of health expenditures for obesity.

Notably, the data on body weight and height, which we use to define BMI, are being self-reported by individuals from 2008 CGSS, instead of recorded by survey enumerators. Thus, the accuracy may be doubted and may raise concerns about measurement error bias. At the same time, the potential measurement error bias also exists in measuring the air pollution as we have discussed in the above data section that we can only access the valid satellite-return $PM_{2.5}$ above the atmosphere data from NASA, instead of the ground-base data which is more accurate and approach to those human being's are exposing to. Besides, due to the difficulties of tracking the survey, we only have the one-year obesity dataset in 2008 and have not set an exposure window for air pollution to introduce more years. This can be improved in the future by expanding the sample size and time dimension to make this study more reliable.

In addition, there is a precondition of our empirical framework which is negligible but needs to mention. That is if the central heating policy has caused behavioral responses that amplify or mitigate the estimated body weight. For instance, the winter heating is likely to cause northerners to spend more time at home for the indoor temperature is higher, and this would lead to a sedentary lifestyle which will let them accumulate fat. Further, flat-rate billing of central heating policy with government subsidies can increase the disposable income of the northerners. This may cause them to alter their consumption patterns in ways that increase the obesity risk, such as buying foods, or decrease their body weight, like pay for a gymnasium. When people’s body weights are affected by these cases, the 2SLS estimates of the effect of heating policy on obesity would be biased, while the estimated local average treatment effects of the heating policy on air pollution and obesity would still be valid.

All in all, our estimated results are limited in those terms. Nevertheless, this study provides a link between air pollution and human body weight for the policymaker to rethink the balance between the environment's sustainable development and economic development. Particularly, for those days air quality and obesity have become the most serious issues globally and the population is overloading as the economy growing rapidly causing the growing contradictions between humans and the environment.

References

Almond, D., Chen, Y., Greenstone, M. and Li, H., 2009. Winter heating or clean air? Unintended impacts of China's Huai river policy. *American Economic Review*, 99(2), pp.184-90.

Andersen, Z.J., Raaschou-Nielsen, O., Ketzel, M., Jensen, S.S., Hvidberg, M., Loft, S., Tjønneland, A., Overvad, K. and Sørensen, M., 2012. Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes care*, *35*(1), pp.92-98.

An, R., Zhang, S., Ji, M., and Guan, C., 2018. Impact of ambient air pollution on physical activity among adults: a systematic review and meta-analysis. *Perspectives in public health*, *138*(2), pp.111-121.

Brook, R.D., Rajagopalan, S., Pope III, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A. and Peters, A., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*, *121*(21), pp.2331-2378.

Brunello, G., Fabbri, D. and Fort, M., 2013. The causal effect of education on body mass: Evidence from Europe. *Journal of Labor Economics*, 31(1), pp.195-223.

Cheng, J., Tong, D., Zhang, Q., Liu, Y., Lei, Y., Yan, G., Yan, L., Yu, S., Cui, R.Y., Clarke, L. and Geng, G., 2021. Pathways of China's PM2. 5 air quality 2015–2060 in the context of carbon neutrality. *National Science Review*.

Chen, S., Oliva, P. and Zhang, P., 2018. *Air pollution and mental health: evidence from China* (No. w24686). National Bureau of Economic Research.

Chen, Y., Ebenstein, A., Greenstone, M. and Li, H., 2013. Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32), pp.12936-12941.

Currie, J., DellaVigna, S., Moretti, E., & Pathania, V., 2010. The effect of fast food restaurants on obesity and weight gain. *American Economic Journal: Economic Policy*, 2(3), 32-63.

Dell, M., 2010. The persistent effects of Peru's mining mita. *Econometrica*, *78*(6), pp.1863-1903.

Deschenes, O., Wang, H., Wang, S. and Zhang, P., 2020. The effect of air pollution on body weight and obesity: Evidence from China. *Journal of Development Economics*, 145, p.102461.

Drewnowski, A. and Specter, S.E., 2004. Poverty and obesity: the role of energy density and energy costs. *The American journal of clinical nutrition*, 79(1), pp.6-16.

Ebenstein, A., Fan, M., Greenstone, M., He, G. and Zhou, M., 2017. New evidence on the impact of sustained exposure to air pollution on life expectancy from

China's Huai River Policy. *Proceedings of the National Academy of Sciences*, 114(39), pp.10384-10389.

Fan, M., He, G. and Zhou, M., 2020. The winter choke: coal-fired heating, air pollution, and mortality in China. *Journal of health economics*, *71*, p.102316.

Ghanem, D. and Zhang, J., 2014. 'Effortless Perfection:'Do Chinese cities manipulate air pollution data?. *Journal of Environmental Economics and Management*, *68*(2), pp.203-225.

Global Air Quality Database App: App for exploring air quality in countries. WHO Global Air Quality Database (update 2018) edition. Version 1.0. Geneva, World Health Organization, 2018.

Globaltimes.cn. 2021. *The great weight divide - Global Times*. [online] Available at: <https://www.globaltimes.cn/content/949282.shtml> [Accessed 15 June 2021].

Heyes, A. and Zhu, M., 2019. Air pollution as a cause of sleeplessness: Social media evidence from a panel of Chinese cities. *Journal of Environmental Economics and Management*, *98*, p.102247.

Imbens, G. and Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), pp.933-959.

Imbens, G.W. and Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), pp.615-635.

Ito, K. and Zhang, S., 2020. Willingness to pay for clean air: Evidence from air purifier markets in China. *Journal of Political Economy*, *128*(5), pp.1627-1672.

Keith, S.W., Redden, D.T., Katzmarzyk, P.T., Boggiano, M.M., Hanlon, E.C., Benca, R.M., Ruden, D., Pietrobelli, A., Barger, J.L., Fontaine, K. and Wang, C., 2006. Putative contributors to the secular increase in obesity: exploring the roads less traveled. *International journal of obesity*, *30*(11), pp.1585-1594.

Lee, D.S. and Lemieux, T., 2010. Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), pp.281-355.

Madrigano, J., Baccarelli, A., Wright, R.O., Suh, H., Sparrow, D., Vokonas, P.S. and Schwartz, J., 2010. Air pollution, obesity, genes and cellular adhesion molecules. *Occupational and environmental medicine*, *67*(5), pp.312-317.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), pp.698-714.

Meo, S.A., Memon, A.N., Sheikh, S.A., Rouq, F.A., Usmani, A.M., Hassan, A. and Arian, S.A., 2015. Effect of environmental air pollution on type 2 diabetes mellitus. *Eur Rev Med Pharmacol Sci*, *19*(1), pp.123-128.

Monteiro, C.A., Conde, W.L., Lu, B. and Popkin, B.M., 2004. Obesity and inequities in health in the developing world. *International journal of obesity*, 28(9), pp.1181-1186.

News.cgtn.com. 2021. *Qinling Mountains: The central park of China*. [online] Available at: <https://news.cgtn.com/news/2020-04-21/Qingling-Mountains-The-central-park-of-China-PRUrshCKWs/index.html> [Accessed 15 June 2021].

Olea, J.L.M. and Pflueger, C., 2013. A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), pp.358-369.

Popkin, B.M. and Slining, M.M., 2013. New dynamics in global obesity facing low- and middle-income countries. *Obesity reviews*, 14, pp.11-20.

Rajagopalan, S. and Brook, R.D., 2012. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes*, *61*(12), pp.3037-3045.

Romero-Corral, A., Montori, V.M., Somers, V.K., Korinek, J., Thomas, R.J., Allison, T.G., Mookadam, F. and Lopez-Jimenez, F., 2006. Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: a systematic review of cohort studies. *The Lancet*, *368*(9536), pp.666-678.

Schwartz, M.W. and Porte, D., 2005. Diabetes, obesity, and the brain. *Science*, *307*(5708), pp.375-379.

Stock, J.H. and W Watson, M., 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, *41*(3), pp.788-829.

The World Health Organization (WHO). 2021. *Air pollution*. [online] Available at: <https://www.who.int/health-topics/air-pollution> [Accessed 15 June 2021].

The World Health Organization (WHO). 2021. *Body Mass Index (BMI)*. [online] Available at: <https://www.who.int/data/gho/data/themes/theme-details/GHO/body-mass-index-(bmi)> [Accessed 15 June 2021].

The World Health Organization (WHO). 2021. *Obesity and overweight*. [online] Available at: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight> [Accessed 15 June 2021].

Wang, Y., Xue, H., Sun, M., Zhu, X., Zhao, L. and Yang, Y., 2019. Prevention and control of obesity in China. *The Lancet Global Health*, *7*(9), pp.e1166-e1167.