The Effect of Long-term Exposure to Air Pollution on Multimorbidity, Functional Disability and Quality of Life[*]

Ye Yuan Runhao Zhao Chong Liu Junjian Yi

July 24, 2023

Abstract

This study investigates the impact of long-term exposure to fine particulate matters (PM2.5) on multimorbidity, functional disability, and quality of life in China. We find that a 10 $\mu g/m^3$ increase in PM2.5 level increases multimorbidity by 5.2 percentage points (95% CI: 1.9-8.5), functional disability by 3.6 percentage points (0.3-6.9), and reduces quality of life by 21.4 years (4.4-38.4) per one thousand population per year. These adverse effects are more pronounced among populations with lower income and education levels. Our analysis leverages quasi-experimental variation in PM2.5 generated by China's winter heating policy, which provides coal-fueled heating to cities north of the Qinling-Huaihe line but not to those to the south. Our findings are derived from a regression discontinuity design based on distance from the Qinling-Huaihe line and are robust to various estimation specifications and adjustments for covariates. Our results highlight the significant health risks of prolonged air pollution exposure on multimorbidity and population quality of life. Our estimates suggest that China's Clean Air Plan's reduction in national PM2.5 levels from 2013-2020 would save 11.3 million life-years per year.

Keywords: Particulate matters, multimorbidity, functional disability, quality of life

JEL Codes: xxx

[*]Yuan: School of Economics, Institute for Global Health and Development, Peking University. Zhao: School of Economics, Peking University. Liu: School of Economics, Peking University. Yi: Corresponding author, China Center of Economic Research, National School of Development, Peking University; email: junjian@nsd.pku.edu.cn.
We thank seminar/conference participants at various universities and institutes for helpful discussions, suggestions, and comments. We thank the National Natural Science Foundation of China (Nos. 72203004) and the Research Seed Fund of the School of Economics, Peking University, for financial support. Any remaining errors are our own.

Significance Statement.

The study investigates the causal impact of prolonged exposure to fine particulate matter (PM2.5) on the occurrence of multimorbidity, defined as the co-occurrence of two or more chronic conditions, and its associated loss in quality of life. Leveraging quasi-experimental variations in PM2.5 levels generated by a historical, location-based policy in China, the research reveals that a 10-$\mu g/m^3$ increase in 30-year average PM2.5 level heightens the prevalence of multimorbidity by 5.2 percentage points and reduces years lost to disability (YLD) by 21.4 years per one thousand population per year. Our results imply that pollution control achieved since China's Clean Air Plan in 2013 has yielded savings of 11.3 million healthy years per year between 2013 and 2020. Our findings highlight the importance of air quality improvement in mitigating multimorbidity and improving overall quality of life and inform public health policies to address the growing challenges of chronic conditions and aging populations globally.

1 Introduction

Chronic conditions are the leading cause of morbidity and mortality globally (Beaglehole et al., 2011). The World Health Organization (WHO) reports that chronic conditions kill 41 million people per year and account for 74% of global deaths. In addition, non-fatal consequences of chronic conditions such as associated disabilities and cognitive-behavioral disorders contribute to as much to the disease burden and loss in quality of life as premature deaths (Zhou et al., 2019).

Multimorbidity, the co-occurrence of two or more chronic conditions, is the most deadly and debilitating form of chronic conditions. It has become increasingly prevalent in recent decades, imposing a growing burden on individuals and healthcare systems worldwide (Chowdhury et al., 2023; Wallace et al., 2015). More than one-third of adults worldwide live with multimorbidity, this number rises to more than half among individuals aged 60 and above (Chowdhury et al., 2023). Multimorbidity significantly elevates the risk of functional decline, disabilities, and impaired physical and mental health, resulting in a substantial reduction in quality of life (Marengoni et al., 2011; Tran et al., 2022) and accounting for a disproportionate share of medical spending (Sum et al., 2018). Despite these significant consequences, well-established causal evidence regarding causes and risk factors of multimorbidity remains limited.

We investigate the effects of long-term exposure to air pollution on multimorbidity. We focus on particulate matters with a diameter less than 2.5 microns (PM2.5) as the primary air pollutant for two reasons: its health damages and global scale. PM2.5 is known to pose the greatest health risks among all ambient pollutants (Peeples, 2020). Short-term exposure to PM2.5 can result in acute cardiovascular and respiratory events and cognitive-behavioral disorders (Ravishankara et al., 2020; Lucas et al., 2006). Even exposure of a few hours can raise blood pressure and trigger brain inflammation (Block et al., 2012). Prolonged exposure to PM2.5 can cause significant health damages and is expected to lead to a host of adverse outcomes such as multimorbidity. Moreover, the threat of PM2.5 is global (Pai et al., 2022; Gakidou et al., 2017; Hammer et al., 2020). According to World Health organization (WHO), 90% of the global population were exposed to annual PM2.5 levels exceeding the recommended limit of 5 $\mu g/m^3$, with 75% exposed to at least double the limit. Furthermore, the current global PM2.5 levels are projected to persist for decades (Pai et al., 2022). Considering the rising prevalence of multimorbidity and the global threat of sustained PM2.5 exposure, a critical assessment on the causal effect of sustained PM2.5 exposure on multimorbidity is of tremendous scientific and policy relevance.

As chronic conditions develop over a prolonged period of time, it is essential to investigate the cumulative impact of long-term exposure to air pollution, rather than focusing solely on short-term exposure. Empirical assessment of the causal impact of sustained air pollution exposure on population health faces two critical challenges. First, reliable measurement of long-term air pollution levels is often unavailable. There exist a few exceptions (Dockery et al., 1993; Pope et al., 2002, 2009; Yazdi et al., 2021), yet they primarily focus on relatively low levels of PM2.5 in the US. Second, existing studies mostly rely on observational variations in PM2.5 exposure, lacking causal inference. Two recent pioneering works by Chen et al. (2013) and Ebenstein et al. (2017) offer quasi-experimental evidence by exploiting a persistent change in air pollution levels in China caused by a decades-long winter heating policy. Both studies reveal a substantial increase in death rate and reduction in life expectancy in cities with sustained higher air pollution levels. Nonetheless, the impact of sustained pollution exposure on chronic morbidity and quality of life remains unexplored. Considering that morbidity-induced disease burden of chronic conditions is as important as premature deaths (Zhou et al., 2019), we prioritize investigating morbidity, specifically multimorbidity, as a critical yet understudied aspect of the link between environmental factors and population health.

We overcome aforementioned two challenges. First, we obtain long-term PM2.5 levels in China from the recently constructed historical PM2.5 series spanning from 1960 to 2020, provided by Zhong et al. (2022). This dataset combines satellite visibility data, ground-based measurements, emissions, and statistical models, providing high-quality measurements of historical PM2.5 levels at a fine geographic resolution and over a long time period. Second, we leverage the decades-old, north-only winter heating policy in China for causal inference (Figure 1, Panel A). China's winter heating policy was implemented since early 1980s. The policy was implemented only in the north of China, with a policy border roughly defined by the Qinglin-Huaihe line (the *line* for short), a major mountain ridge line commonly adopted by geographers to distinguish between Northern and Southern China. This north-only winter heating policy has led to a larger number of coal-powered heating infrastructure being constructed in the north of the line while no equivalent system existed in the south (Almond et al., 2009), resulting in a markedly sharp increase in PM2.5 levels in the north relative to the south (Figure 1, Panel B). This offers a quasi-experimental setting to identify the causal impact of sustained PM exposure on population health. Our empirical setting and methodology is motivated by the pioneering work of Chen et al. (2013) and Ebenstein et al. (2017). We discuss details of the institutional background of China's winter heating policy in Appendix Section A.

We leverage this geographic discontinuity in PM2.5 concentration and adopt the regression discontinuity design (RD) in statistical analysis. Our analysis utilizes data from the nationally representative China Health and Retirement Longitudinal Survey (CHARLS). We primarily examine the population aged 45 to 60 to mitigate survival bias resulting from the impact of sustained pollution exposure on premature mortality. CHARLS provides high-quality information on chronic diseases and health behaviors, enabling us to examine the impact of sustained PM2.5 exposure on the prevalence of various chronic diseases and multimorbidity and investigate individual's behavioral responses to such sustained environmental adversity. Statistical analyses were conducted using STATA 16.0 software (Stata Corp LP, College Station, TX, USA).

We find that China's north-only winter heating policy since the early 1980s created a significant north-south gap in average PM2.5 concentrations. The average gap was estimated to be 26.8 $\mu g/m^3$ (95% CI: 15.7-37.8) for the period of 1980-2010. Notably, this north-south gap in PM2.5 levels emerged in the early 1980s, persisted for three decades from 1980 to 2010, and started to decline only after the 2013 anti-pollution campaign. The north-south gap was primarily observed during the winter heating season (November-March) and was absent during the non-heating season (April-September), confirming the central role of the winter heating policy in driving the long-term north-south gap in air pollution exposure. Other than the gap in sustained PM exposure, our analysis reveals no north-south gap in water pollution, nor statistically significant differences in demographic or socioeconomic characteristics.

The sustained exposure to a higher level of PM2.5 resulted in a greater prevalence of multimorbidity in the north. Among a nationally representative sample of individuals aged 45 to 60, the north-south gap in the prevalence rate of having two or more chronic conditions was estimated at 9.6 pp (95% CI: 3.5-15.7), and for at least three chronic conditions, it was 5.4 pp (95% CI: 0.3-10.5). Per disease, the north-south gap in multimorbidity is mainly driven by significant gaps in the prevalence of common chronic conditions such as hypertension, diabetes, asthma, and COPD. Such co-occurrence of multiple conditions significantly increased the difficulty in performing daily activities. The Katz index and Barthel index, both of which measure functional disability and individual's ability of independent living (Hartigan, 2007), were 13.3 pp (95% CI: 6.4-20.2) and 7.9 pp (95% CI: 5.2-10.6) lower in the north.

We then estimate the effect of sustained PM2.5 exposure on an measure of overall disease burden—Years Lived with Disability (YLD), which estimates the number of healthy years lost due to living with adverse health conditions. YLD quantifies the non-fatal consequences of morbidity and complements measures such as Years of Life Lost (YLL) that primarily focus on mortality (Vos et al., 2016; Naghavi et al., 2017). We find that sustained PM2.5 exposure increased the population YLD by 73.664 (95% CI: 40.377-106.951) years per one thousand population per year. This estimate quantifies and highlights the substantial adverse impact of sustained PM2.5 exposure on multimorbidity and quality of life.

Combining the estimated north-south gaps in sustained PM2.5 exposure and disease burden, we find that a ten-unit ($\mu g/m^3$) increase in long-term PM2.5 level is associated with a 5.2 pp (95% CI: 1.9-8.5) increase in the prevalence of having two or more chronic conditions, and a 1.7 pp (95% CI: -0.3-3.7) increase in the prevalence of having three or more conditions. Furthermore, it leads to a decrease of 3.6 pp (95% CI: 0.3-6.9) in the Katz index and 1.5 pp (95% CI: -0.3-3.3) in the Barthel index. Overall, a one-unit increase in long-term PM2.5 level results in a YLD increase of 21.393 (95% CI: 4.392-38.394) years per one thousand population per year.

Our analysis also uncovers significant heterogeneity in the health impacts of long-term PM2.5 exposure among different population groups. The more disadvantaged groups experienced more pronounced north-south gaps in multimorbidity, disabilities, and overall disease burden. Specifically, those with lower education or income levels bore a disproportionately larger burden of disease due to sustained PM exposure. Different levels of pollution avoidance behaviors contribute to the heterogeneous impacts. On average, we observed a significantly higher frequency of physical exercise and lower rates of smoking and alcohol consumption in the north, indicating that individuals adjusted their lifestyles and consumption habits to mitigate the adverse effects of pollution exposure. However, these beneficial adjustments were more prevalent among the more educated and higher-income group, while they were almost nonexistent among the less educated and lower-income group. This disparity helps explain why the latter group experienced a higher burden of chronic diseases due to sustained PM2.5 exposure.

Lastly, we quantify the welfare gains of the nationwide air pollution reduction since 2013. The series of stringent environmental regulations implemented since China's Clean Air Plan in 2013 resulted in a total reduction of the national average PM2.5 level by 20.3 $\mu g/m^3$ from 2013 to 2020. Based on our estimates, this would lead to a 10.5 pp decrease in the average prevalence of multimorbidity and save 11.3 million healthy years per year. This highlights significant welfare gains of pollution control from lowering the disease burden of non-fatal chronic conditions and multimorbidity, which supplement previous estimates based on mortality and hold important policy implications for pollution control programs in developing countries such as India and Indonesia. These findings have implications beyond China and can guide policy decisions and interventions in other countries facing similar challenges related to chronic diseases and air pollution. The research contributes to a more comprehensive understanding of the interactions between environmental factors and health outcomes, fostering evidence-based approaches to address the global health burden posed by chronic conditions.

The rest of this article is organized as follows. *Data* describes the data sources and variables. *Econometric Model* outlines empirical strategies. *Empirical Results* represents the main findings. *Conclusion* concludes.

2 Data

Analytical Sample Our main data source is the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative longitudinal survey of Chinese individuals aged 45 and above. CHARLS provides comprehensive demographic and socioeconomic information along with reliable physician-diagnosed disease records. CHARLS conducted the baseline survey in 2011, followed by biennial surveys. We focus on the CHARLS 2011 baseline survey. We do so because the annual level of PM2.5 remained stable before the 2010s, offering a stable window to analyze the impact of sustained PM exposure. PM2.5 levels sharply declined after the 2013 anti-pollution campaign. Appendix Figure A1 illustrates this pattern. We therefore analyze the impact of the stable long-term PM2.5 levels from 1980-2010 on population health in 2011 and adopt the estimates to infer the welfare gains resulting from the sharp post-2013 reduction in PM2.5 levels.

We restrict to the sample whose residential place is the same as birthplace. This is to lessen the potential selection due to potential compensatory migration in response to pollution exposure. In addition, Ebenstein et al. (2017) has provided robust evidence based on 2005 census that migration was limited during this time period in China and there was no notable difference in migration rates between the north and south of the Qinling-Huaihe line. Therefore, pollution concentrations at an individual's birthplace are a reasonable measure of their lifetime exposure to pollution.

The baseline sample consists of 17,708 respondents from 125 prefecture-level cities in 28 provinces. We aggregate individual-level data to the city level and calculate the population-weighted average prevalence rate of chronic conditions and multimorbidity using the CFPS sampling weights. We combine the long-term PM2.5 levels with the baseline sample, along with relevant city-level demographic and socioeconomic variables such as age structure, GDP per capita, and industrial structure obtained from various city statistical yearbooks.

Main Outcome Variables We analyze three categories of outcome variables: multimorbidity, functional disability, and YLD. To define the incidence of multimorbidity, we obtain the prevalence for the nine most common chronic conditions in each CHARLS city, including hypertension, dyslipidemia, diabetes, asthma, COPD, other chronic respiratory diseases, cardiovascular diseases, liver diseases, and kidney diseases. Stroke and cancer is not included in our analysis due to its extremely low self-reported incidence in our sample. Appendix Table B1 further describes each chronic condition and Appendix Section B provides more details on variable construction.

Multimorbidity and disability We define two measures of multimorbidity: one indicating the co-occurrence of *two* or more chronic conditions (out of aforementioned disease categories), and the other indicating *three* or more chronic conditions. We also analyze functional disability, which is commonly associated with multimorbidity (Marengoni et al., 2011; Tran et al., 2022). We assess functional disability using the widely adopted Katz index and Barthel index, which measure an individual's difficulty in performing basic activities of daily living (ADL) and tasks required for independent living (Hartigan, 2007). We provide further details on definition and construction of ADL, Katz index, and Barthel index in Appendix Table B2, Table B3, and Table B4, respectively.

Years lived with disability We quantify the overall disease burden and its impact on quality of life using the WHO's years lived with disability (YLD). YLD is a composite measure that estimates the number of healthy years lost due to living with adverse health conditions, providing a comprehensive assessment of non-fatal health burdens (Vos et al., 2016; Naghavi et al., 2017). Originally developed to measure the global burden of disease and injury, YLD serves as an important metric for assessing the relative magnitude of health losses from different causes. Additionally, it is valuable for cost-effectiveness analyses of public health interventions, making it a suitable measure for our study on the health impact of the long-standing winter heating policy and subsequent pollution control efforts.

We compute YLD for each city population following the WHO guideline in four steps. First, we obtain the prevalence rate of each chronic condition for a city's population. Second, we assign WHO's disability weights to each condition, representing the relative magnitude of losses of healthy life associated with the condition (WHO, 2020). These disability weights are derived from surveys or expert consensus (Salomon et al., 2015). Appendix Table B5 provides disability weights for chronic conditions reported in the CHARLS survey, ranging from 0 (perfect health) to 1 (equivalent to death). Third, we multiply the prevalence of each condition by its corresponding disability weight to calculate the years lived with disability for each condition. Finally, we obtain the YLD for a population by summing the years lived with disability across all chronic conditions. Additional details on the construction of YLD are presented in Appendix Section B.7.

Long-term PM2.5 level We measure the long-term PM2.5 level in each city as the average annual level of PM2.5 from 1980 to 2010. Our measure is based on a database constructed by Zhong et al. (2022), which provides historical PM2.5 concentrations at a $0.25° \times 0.25°$ grid every 6 hours. The database integrates satellite images, meteorological data, pollution emissions, and elevation using a Light Gradient Boosting Machine (LightGBM) model; and it is the first database of PM2.5 to offer such a long time horizon, high temporal resolution, and fine geographic coverage in China. We validate the accuracy of the data against ground-monitor-based records in Appendix Figure A1.

Meteorological conditions and regional covariates City-level meteorological variables, including temperature, precipitation, relative humidity, wind speed, and sunshine duration, are sourced from the Chinese daily surface meteorological dataset (V3.0) provided by the National Meteorological Science Data Center. City-level socioeconomic variables are obtained from China Urban Statistical Yearbook, including GDP per capita, public investment in education and medical resources, population size, population growth rate and share of non-farm population, the ratio of value-added from primary and secondary industries, as well as annual measurements of wastewater discharge and soot emissions.

Health behaviors We gather health-related behavior measures from CHARLS, including physical exercise frequency per week, years of smoking (from the initial age of smoking to either the year of survey or the year of cessation), years of alcohol drinking (from the initial age of drinking to either the year of survey or the year of cessation), and nightly sleep duration. Demographic and socioeconomic variables such as education years and family health expenditure are also collected from CHARLS. Furthermore, nutritional intakes variables (calories, carbohydrates, protein, and fat) are obtained from the China Health and Nutrition Survey (CHNS). Additional information on variable definitions can be found in Appendix Section B.

Descriptive statistics Table 1 provides descriptive statistics for the main outcome variables, showing the average values separately for cities located north (Column 1) and south (Column 2) of the Qinling-Huaihe line, with the corresponding mean differences shown in Column 3. The data reveal significant disparities in PM2.5 concentrations and population health outcomes between the north and south regions. In the subsequent statistical analyses, we aim to assess the extent to which the differences in health outcomes can be causally attributed to the disparity in long-term air pollution exposure, while controlling for other meteorological and socioeconomic differences. To achieve this, we include a set of socioeconomic, meteorological, and demographic city characteristics as controls in the regression analyses. Appendix Table B6 provides descriptive statistics for these covariates.

3 Econometric Model

We employ a regression discontinuity design (RD) in our statistical analysis, to addresses concerns of unobserved confounding factors when estimating the effect of air pollution on health outcomes. RD is widely adopted across various fields, including psychology, education, biostatistics, environmental science, and economics (Imbens and Lemieux, 2008; Lee and Lemieux, 2010; Cook, 2008). Building on the approach used by Chen et al. (2013) and Ebenstein et al. (2017), we leverage the public winter heating policy implemented only in the north of the Qinglin-Huaihe line as a quasi-experimental setting to establish causal inference. We investigate the north-south gap in long-term PM2.5 concentrations and in population health, and explore any potential north-south gap in predetermined demographic and socioeconomic characteristics that could potential correlate with PM2.5 concentration and population health.

Our regression models are specified as follows:

$$PM_i = \alpha_0 + \alpha_1 North_i + \alpha_2 f(Lat_i) + \alpha_3 North_i \times f(Lat_i) + Long_i + W_i\alpha_4 + \varepsilon_i; \quad (1)$$

$$Y_i = \beta_0 + \beta_1 North_i + \beta_2 f(Lat_i) + \beta_3 North_i \times f(Lat_i) + Long_i + W_i\beta_4 + u_i. \quad (2)$$

PM_i refers to the average PM2.5 concentration from 1980 to 2010 in city i, while Y_i represents three categories of health outcomes for the population aged 45-60 in city i in CHARLS 2011 wave: the prevalence rate of multimorbidity, the average Katz index and Barthel index measuring functional disability, and the population YLD reflecting overall disease burden. The function $f(Lat_i)$ relates to the distance Lat_i, measured in units of latitude to the Qinling-Huaihe line. The variable $North_i$ indicates whether city i is located north of the line. The variable $Long_i$ represents longitude-region fixed effects, dividing the area along the Qinling-Huaihe line into three regions. This accounts for the differences in local demographic and socioeconomic characteristics between different longitude regions. The vector W_i includes city-level demographic, socioeconomic, and meteorological covariates (see the descriptive statistics of included covariates in Appendix Table B6). All regressions are weighted by the city population and robust standard errors are clustered at the city level.

Causal inference in this model relies on the (weak) assumption that unobserved factors influencing PM2.5 and morbidity change smoothly with distance to the line (Lat_i). We validate this assumption through a range of RD graphical analyses and statistical tests (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). Please refer to Appendix Section C for details of these tests.

The coefficients of interest, α_1 and β_1, estimate the causal impact of the north-only winter heating policy on long-term PM2.5 concentrations and chronic conditions, respectively. To ensure consistent estimation, the choice of polynomial function $f(\cdot)$ and the bandwidth (i.e., the distance range used in estimation) are crucial. Recent advancements in the RD methodology emphasize the sensitivity of results to the polynomial order when using parametric specifications (Gelman and Imbens, 2019). Therefore, we adopt a non-parametric, local linear approach with a narrow bandwidth. The optimal bandwidth is determined using a data-driven approach (Imbens and Kalyanaraman, 2012; Calonico et al., 2014, 2020). We present estimation results for two kernel functions in the tables below to assess the stability of results. Sensitivity checks based on quadratic functional forms of $f(\cdot)$ are also conducted (Gelman and Imbens, 2019).

4 Empirical Results

4.1 Long-term PM2.5 concentration

We present a visual representation of the impact of the winter heating policy on long-term PM2.5 concentration through a RD figure. Figure 2 illustrates the distribution of PM2.5 concentrations in latitude bins relative to the Qinling-Huaihe line. Each circle represents the average PM2.5 concentration (1980-2010) in cities within a 0.5-degree latitude bin. The figure demonstrates a distinct and abrupt rise increase in PM2.5 concentrations at the northern border of the line.

Building on this empirical evidence, we conduct statistical estimation on the north-south gap of PM2.5 and restrict to an optimal, roughly 3-latitude band of region along the Qinling-Huaihe line. This optimal bandwidth is determined by a data-driven algorithm (Imbens and Kalyanaraman, 2012; Calonico et al., 2014, 2020). Table 2 (A) reports the estimated north-south gap in long-term PM2.5 level is 26.8 $\mu g/m^3$ (95% CI: 15.7-37.8), which corresponds to 42.1% of the national average. This sizable gap in PM2.5 is consistent with the documented PM10 gap of 41.7 $\mu g/m^3$ as reported by Ebenstein et al. (2017), considering that PM2.5 is a primary component of PM10 (accounting for approximately 70% of PM10 in our dataset).

We further examine the dynamic pattern of the north-south gap of PM2.5 concentrations from 1960 to 2020. Figure 3 illustrates that the gap was small from 1960 to 1975, increased sharply since the early 1980s, reached its peak and stabilized between 1985 and 2000. In the late 2000s, following the implementation of the nationwide anti-pollution campaign, the gap began to decline. By 2015, the PM2.5 gap had returned to pre-1980 levels. It is also noteworthy that the north-south gap in PM2.5 predominantly manifests during the winter heating season (November-March), and are much smaller during the non-heating season (April-September). This observation highlights the pivotal role played by the winter heating policy in exacerbating the north-south gap in PM2.5 concentrations.

4.2 Health Outcomes

Prevalence of multimorbidity and specific chronic diseases The sustained higher level of PM2.5 from 1980 to 2010 is anticipated to contribute to a higher prevalence of chronic conditions and increase the co-occurrence of multiple conditions (i.e., multimorbidity). Figure 4 (A) and (B) plots the city-level prevalence rates of multimorbidity in latitude bins relative to the Qinling-Huaihe line. Notably, there is a sharp increase in the average prevalence rate of multimorbidity at the northern border of the line.

Table 2 (B) shows the estimated north-south gap in multimorbidity obtained from RD regressions. The findings indicate that individuals residing in the north of the line had a significantly higher rate of multimorbidity, with an increase of 9.6 pp (95% CI: 3.5-15.7, 41.2% of the baseline rate in 2011 in the north) for the multimorbidity rate for two or more chronic conditions, and 5.4 pp (95% CI: 0.3-10.5, 52.4% of the baseline rate in 2011 in the north) for the multimorbidity rate for three or more chronic conditions. The RD regressions based on different kernel specifications and corresponding optimal bandwidths yield similar estimates. It is important to note that these estimates likely underestimate the north-south gap in multimorbidity, considering that our analytical sample only includes the middle-aged cohorts (age 45-60).

Appendix Table B7 presents estimates of the north-south gap in the prevalence of specific chronic condition. The sustained exposure to PM2.5 has resulted in increased prevalence rates for nine major chronic conditions. These findings highlight the rising occurrence of various chronic conditions, which contributes to the increase in multimorbidity prevalence.

Prevalence of functional disability Having established a causal link between long-term PM2.5 exposure and an increased prevalence rate of multimorbidity, we now investigate whether prolonged pollution exposure further elevates the incidence of functional disability—a well-known consequence of multimorbidity (Marengoni et al., 2011). Disability not only undermines one's ability to live independently but also significantly impairs overall quality of life. Moreover, the physical and cognitive impairments resulting from disability hinder the effective management of multiple conditions, exacerbating the severity of multimorbidity and perpetuating a detrimental cycle (Sousa et al., 2009; Sheridan et al., 2019). We adopt two instrument-based measures, namely the Katz index and Barthel index, to assess functional disability and limitations in activities of daily living (ADL).

Figure 4 (C) and (D) clearly depict a pronounced decrease in the Katz index and Barthel index, respectively, in the north of the Qinling-Huaihe line. Our regression analysis, as presented in Table 2 (C), reveals a statistically significant north-south gap in the Katz index of 13.3 pp (95% CI: 6.4-20.2, 18.5% of the baseline rate in 2011 in the north), and a north-south gap in the Barthel index of 7.9 pp (95% CI: 5.2-10.6, 8.9% of the baseline rate in 2011 in the north). These findings consistently indicate a higher prevalence of functional disabilities among individuals exposed to higher PM2.5 levels in the north.

Quality of life The increase in multimorbidity rate and functional disabilities both indicate a significant decline in quality of life. We adopt the WHO's Years Lived With Disability (YLD) as a composite measure for the overall disease burden and loss in quality of life. YLD quantifies the impact of non-fatal health conditions on populations by estimating the number of healthy years lost due to living with adverse health conditions (Vos et al., 2016; Naghavi et al., 2017). Details on constructing the YLD in each city are provided in Appendix Section B.7.

Figure 4 (E) depicts a marked increase in the YLD in the north of the Qinling–Huaihe line, indicating a significant decline in quality of life attributed to sustained exposure to elevated levels of PM2.5. The regression results, as presented in Table 2 (D), confirms a statistically significant north-south gap in YLD by 73.7 years (95% CI: 40.4-107.0) for every one thousand population per year, which accounts for 60.0% of the baseline rate in 2011 in the north. Overall, Figure 4 presents robust evidence of the adverse impacts of sustained PM2.5 exposure on the burden of chronic diseases and the quality of life.

Marginal effect of sustained PM2.5 exposure Table 3 reports on the estimates of the marginal effect of additional 10 $\mu g/m^3$ PM2.5 on multimorbidity, functional disabilities and YLD from the instrumental variable (IV) estimation. The estimated marginal effect can be understood as the estimated north-south gap in various health outcomes divided by the estimated north-south gap in long-term PM2.5 concentrations. Panel A shows that an 10 $\mu g/m^3$ PM2.5 increase the prevalence of suffering two and three chronic conditions by 5.2 pp (95% CI: 1.9-8.5) and 1.7 pp (95% CI: -0.3-3.7) respectively, which is significant at the 5% level. Meanwhile, Panel B shows that an additional 10 $\mu g/m^3$ PM2.5 lower the degree of independence measured by Katz index and Barthel index by 3.6 pp (95% CI: 0.3-6.9) and 1.5 pp (95% CI: -0.3-3.3) respectively. Finally, we find an additional 10 $\mu g/m^3$ PM2.5 increase the YLD for every one thousand population per year by 21.4 (95% CI: 4.4-38.4) years.

Different impacts across socioeconomic groups We proceed to examine the heterogeneity in the results across different socioeconomic groups. Table 4 reveals that individuals with lower education levels (illiterate or below primary school) and lower income (below median income) experienced more significant adverse impacts when residing north of the line. The north-south disparity in multimorbidity ($\geq$2 conditions) was 13.2 pp compared to 10.1 pp for the less educated and more educated groups, respectively. Similarly, the north-south gap in the Katz index was 19.4 pp versus 11.9 pp, and the gap in YLD was 122.050 years versus 48.174 years. When comparing the low-income group to the high-income group, the estimated north-south gap in multimorbidity ($\geq$2 conditions) was 18.3 pp versus 7.3 pp, the gap in Katz index was 38.9 pp versus 12.8 pp, and the gap in YLD was 101.045 years versus 46.169 years. Overall, our findings indicate that more disadvantaged groups (less educated or lower income) bear a greater disease burden due to prolonged exposure to air pollution. The observed discrepancy in disease burden across socioeconomic groups may stem from variations in health knowledge, awareness of air pollution, and differential adjustments in lifestyle behaviors as a response to pollution exposure.

Given the intuitive understanding of disparities in knowledge and awareness, we delve into exploring diverse behavioral adaptations to prolonged exposure to pollution. Our investigation centers on key health behaviors including frequency of physical exercise, smoking and drinking habits, and nutrition intakes. Previous research has underscored the significant role of these behaviors in the onset and management of chronic diseases (Liu et al., 2013; Paudel et al., 2019; Ng et al., 2020). Indeed, Figure 5 presents evidence that residents in the north exhibit higher frequency of physical exercise, lower alcohol consumption, and fewer years of smoking in comparison to their southern peers (also see Appendix Table B8). We do not observe notable north-south differences in nutrition intakes (Appendix Figure A2 and Table B9). Overall, the lifestyle adaptations in exercise, smoking and drinking contribute to mitigating the detrimental impacts of prolonged exposure to PM2.5.

However, Table 5 reveals a substantial disparity in the adoption of these lifestyle adjustments among different socioeconomic groups. Individuals with lower education levels and lower income exhibit lower adaptations in these lifestyle behaviors, while those with higher education and higher income demonstrate a more pronounced pattern of such adaptations. These findings help explain why more disadvantaged groups bear a disproportionate disease burden from prolonged exposure to PM2.5 (as evidenced in Table 4).

Robustness Lastly, we conduct a comprehensive series of diagnostic tests and robustness checks to ensure the validity and reliability of our findings. First, we perform the RD density test (McCrary, 2008; Cattaneo et al., 2020, 2021), which aims to detect any potential discontinuity in the sample density distribution in an RD design. Appendix Figure A3 demonstrates that there is no evidence of such a discontinuity in the sample density. Second, we examined the appropriateness of selecting the Qinglin-Huaihe line as the geographic reference line for the winter heating policy. To do this, we construct various placebo RD reference lines by shifting the Qinglin-Huaihe line to the north or south. Appendix Figure A4 shows a sizable and statistically significant north-south gap in long-term PM2.5 concentrations only with respect to the real Qinling-Huaihe line, but not the placebo lines. Third, we conduct balance tests in Appendix Figure A5, Figure A6, and Appendix Table B10 and ensure that there are no noticeable north-south differences in socioeconomic and meteorological factors. Lastly, we assess the sensitivity of our RD estimates to several factors. These include: (1) alternative lengths of bandwidth in the local linear RD design (Appendix Figure A7), (2) the adoption of the optimal bandwidth determined by the coverage error rate (CER) optimal bandwidth algorithm instead of MSE (Appendix Tables B11 and B12), (3) the use of a quadratic distance function rather than local linear distance function (Appendix Tables B13 and B14), and (4) the implementation of a bias correction procedure on RD estimates with a robust variance estimator (Appendix Tables B15 and B16). All of these analyses consistently support our baseline findings and provide further validation for our RD specification.

Social welfare Our back-of-the-envelope calculations demonstrate the significant welfare gains achieved through China's nationwide efforts in pollution control and environmental protection since 2013. We find that each 10-unit reduction in long-term PM2.5 levels corresponds to a 5.2 percentage point reduction in multimorbidity prevalence and a saving of 21.4 healthy years per one thousand population per year. Considering that China's Clean Air Plan has successfully reduced the average national PM2.5 level by 20.3 $\mu g/m^3$ from 2013 to 2020, this translates into a reduction in multimorbidity by 10.5 percentage points and a total saving of 11.3 million healthy years per year (=20.3/10*21.4*260/1000, based on a national population aged 45 to 60 of 260 million in 2010).

To put these numbers into perspective, previous estimates by Gruber et al. (2023) show that China's implementation of the new rural medical scheme, the world's largest provision of public health insurance, would save about 1.01 million lives per year since its national rollout. In comparison, the estimated saving in life years from the 2013 national anti-pollution campaign is about an order of magnitude larger than the corresponding saving in life years from the national provision of public health insurance. This highlights the substantial overall welfare gains achieved through environmental protection in China.

5 Conclusion

Our study contributes to the literature by being one of the first to examine the causal impact of prolonged exposure to air pollution on multimorbidity and quality of life. To achieve this, we utilize a unique dataset of 30-year long panel series of PM2.5 concentrations derived from a high-resolution remote sensing database. Leveraging the implementation of China's winter heating policy as a natural experiment, we exploit the geographic variation in PM2.5 levels to make causal inferences.

Our findings reveal a sharp increase in PM2.5 concentrations in the northern regions of the Qinglin-Huaihe line following the implementation of the winter heating policy, and this elevated pollution level persisted significantly higher than the southern regions over the three-decade period from 1980 to 2010. This sustained higher level of PM2.5 exposure led to a significantly higher rate of multimorbidity and disability, and a substantial reduction in quality life. Furthermore, our research assesses potential heterogeneous effects across different socioeconomic groups within the population and reveals that the disadvantaged group of society, such as those less educated and with lower income, suffered a greater burden of multimorbidity and disability. This finding sheds light on the distributional consequences of long-term PM exposure, which can inform targeted and equitable public health interventions.

Overall, our study fills a crucial knowledge gap and underscores the long-term implications of air pollution on public health and quality of life. By utilizing robust empirical methods and comprehensive data, we provide valuable insights into the far-reaching effects of air pollution on population health and the importance of addressing this pressing environmental issue for the well-being of global populations.

In addition, our findings underscore the significant welfare gains from China's Clean Air Plan, which effectively reduced the average PM2.5 exposure for the entire population. Our results suggest that the reduction in national average PM2.5 levels would have resulted in savings of 11.3 million healthy life years per year for the population aged 45 to 60 alone, and the positive impact on health is expected to be even greater for older cohorts. Moreover, the policy implications of our research extend beyond China. With over 75% of the global population living in areas with PM2.5 concentrations exceeding the WHO's safe limit, there is a pressing need for effective environmental protection policies worldwide. By reducing population exposure to PMs, such policies have the potential to significantly decrease the prevalence of chronic conditions, multimorbidity, and associated functional decline, ultimately leading to substantial improvements in the quality of life for billions of people globally.

References

Almond, D., Y. Chen, M. Greenstone, and H. Li (2009). Winter heating or clean air? unintended impacts of china's huai river policy. *American Economic Review 99*(2), 184–90.

Beaglehole, R., R. Bonita, R. Horton, C. Adams, G. Alleyne, P. Asaria, V. Baugh, H. Bekedam, N. Billo, S. Casswell, et al. (2011). Priority actions for the non-communicable disease crisis. *The lancet 377*(9775), 1438–1447.

Block, M. L., A. Elder, R. L. Auten, S. D. Bilbo, H. Chen, J.-C. Chen, D. A. Cory-Slechta, D. Costa, D. Diaz-Sanchez, D. C. Dorman, et al. (2012). The outdoor air pollution and brain health workshop. *Neurotoxicology 33*(5), 972–984.

Braga, A. L., A. Zanobetti, and J. Schwartz (2002). The effect of weather on respiratory and cardiovascular deaths in 12 us cities. *Environmental health perspectives 110*(9), 859–863.

Bu, X., Z. Xie, J. Liu, L. Wei, X. Wang, M. Chen, and H. Ren (2021). Global pm2. 5-attributable health burden from 1990 to 2017: Estimates from the global burden of disease study 2017. *Environmental Research 197*, 111123.

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal 23*(2), 192–210.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal 14*(4), 909–946.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*(531), 1449–1455.

Cattaneo, M. D., M. Jansson, and X. Ma (2021). Local regression distribution estimators. *Journal of econometrics*.

Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (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), 12936–12941.

Chowdhury, S. R., D. C. Das, T. C. Sunna, J. Beyene, and A. Hossain (2023). Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. *EClinicalMedicine 57*.

Cook, T. D. (2008). "waiting for life to arrive": a history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics 142*(2), 636–654.

Ćwirlej-Sozańska, A., A. Wiśniowska-Szurlej, A. Wilmowska-Pietruszyńska, and B. Sozański (2019). Determinants of adl and iadl disability in older adults in southeastern poland. *BMC geriatrics 19*(1), 1–13.

Ding, L., M. Fan, and P. Nie (2021). The long-term effect of air pollution on human cognition: Evidence from china. *Available at SSRN 3792594*.

Dockery, D. W., C. A. Pope, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris Jr, and F. E. Speizer (1993). An association between air pollution and mortality in six us cities. *New England journal of medicine 329*(24), 1753–1759.

Ebenstein, A., M. Fan, M. Greenstone, G. He, and M. Zhou (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), 10384–10389.

Fan, M., G. He, and M. Zhou (2020). The winter choke: Coal-fired heating, air pollution, and mortality in china. *Journal of Health Economics 71*(C), S0167629619311257.

Gakidou, E., A. Afshin, A. A. Abajobir, K. H. Abate, C. Abbafati, K. M. Abbas, F. Abd-Allah, A. M. Abdulle, S. F. Abera, V. Aboyans, et al. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet 390*(10100), 1345–1422.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Goggins, W. B. and E. Y. Chan (2017). A study of the short-term associations between hospital admissions and mortality from heart failure and meteorological variables in hong kong: Weather and heart failure in hong kong. *International journal of cardiology 228*, 537–542.

Gruber, J., M. Lin, and J. Yi (2023). The largest insurance expansion in history: Saving one million lives per year in china. Technical report, National Bureau of Economic Research.

Hammer, M. S., A. van Donkelaar, C. Li, A. Lyapustin, A. M. Sayer, N. C. Hsu, R. C. Levy, M. J. Garay, O. V. Kalashnikova, R. A. Kahn, et al. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology 54*(13), 7879–7890.

Hartigan, I. (2007). A comparative review of the katz adl and the barthel index in assessing the activities of daily living of older people. *International journal of older people nursing 2*(3), 204–212.

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

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

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

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

Liu, J., B. M. Varghese, A. Hansen, M. A. Borg, Y. Zhang, T. Driscoll, G. Morgan, K. Dear, M. Gourley, A. Capon, et al. (2021). Hot weather as a risk factor for kidney disease outcomes: A systematic review and meta-analysis of epidemiological evidence. *Science of The Total Environment 801*, 149806.

Liu, Y., J. B. Croft, A. G. Wheaton, G. S. Perry, D. P. Chapman, T. W. Strine, L. R. McKnight-Eily, and L. Presley-Cantrell (2013). Association between perceived insufficient sleep, frequent mental distress, obesity and chronic diseases among us adults, 2009 behavioral risk factor surveillance system. *BMC public health 13*(1), 1–8.

Lu, F., D. Xu, Y. Cheng, S. Dong, C. Guo, X. Jiang, and X. Zheng (2015). Systematic review and meta-analysis of the adverse health effects of ambient pm2. 5 and pm10 pollution in the chinese population. *Environmental research 136*, 196–204.

Lucas, S.-M., N. J. Rothwell, and R. M. Gibson (2006). The role of inflammation in cns injury and disease. *British journal of pharmacology 147*(S1), S232–S240.

Marengoni, A., S. Angleman, R. Melis, F. Mangialasche, A. Karp, A. Garmen, B. Meinow, and L. Fratiglioni (2011). Aging with multimorbidity: a systematic review of the literature. *Ageing research reviews 10*(4), 430–439.

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

Muller, N. Z., R. Mendelsohn, and W. Nordhaus (2011). Environmental accounting for pollution in the united states economy. *American Economic Review 101*(5), 1649–75.

Naghavi, M., A. A. Abajobir, C. Abbafati, K. M. Abbas, F. Abd-Allah, S. F. Abera, V. Aboyans, O. Adetokunboh, A. Afshin, A. Agrawal, et al. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the global burden of disease study 2016. *The lancet 390*(10100), 1151–1210.

Ng, R., R. Sutradhar, Z. Yao, W. P. Wodchis, and L. C. Rosella (2020). Smoking, drinking, diet and physical activity—modifiable lifestyle risk factors and their associations with age to first chronic disease. *International Journal of Epidemiology 49*(1), 113.

Noelker, L. S. and R. Browdie (2014). Sidney katz, md: A new paradigm for chronic illness and long-term care. *The Gerontologist 54*(1), 13–20.

Pai, S. J., T. S. Carter, C. L. Heald, and J. H. Kroll (2022). Updated world health organization air quality guidelines highlight the importance of non-anthropogenic pm2. 5. *Environmental Science & Technology Letters.*

Paudel, S., A. J. Owen, E. Owusu-Addo, and B. J. Smith (2019). Physical activity participation and the risk of chronic diseases among south asian adults: a systematic review and meta-analysis. *Scientific reports 9*(1), 1–12.

Peeples, L. (2020). How air pollution threatens brain health. *Proceedings of the National Academy of Sciences 117*(25), 13856–13860.

Pope, C. A., R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama 287*(9), 1132–1141.

Pope, C. A., M. Ezzati, and D. W. Dockery (2009). Fine-particulate air pollution and life expectancy in the united states. *New England Journal of Medicine 360*(4), 376–386.

Ravishankara, A. R., L. M. David, J. R. Pierce, and C. Venkataraman (2020). Outdoor air pollution in india is not only an urban problem. *Proceedings of the National Academy of Sciences 117*(46), 28640–28644.

Salomon, J. A., J. A. Haagsma, A. Davis, C. M. de Noordhout, S. Polinder, A. H. Havelaar, A. Cassini, B. Devleesschauwer, M. Kretzschmar, N. Speybroeck, et al. (2015). Disability weights for the global burden of disease 2013 study. *The Lancet Global Health 3*(11), e712–e723.

Schinasi, L. H., C. C. Kenyon, K. Moore, S. Melly, Y. Zhao, R. Hubbard, M. Maltenfort, A. D. Roux, C. B. Forrest, and A. J. De Roos (2020). Heavy precipitation and asthma exacerbation risk among children: a case-crossover study using electronic health records linked with geospatial data. *Environmental Research 188*, 109714.

Sheridan, P., C. Mair, and A. Quiñones (2019). Associations between prevalent multimorbidity combinations and prospective disability and self-rated health among older adults in europe. *BMC Geriatrics 19*(1), 198.

Song, C., J. He, L. Wu, T. Jin, X. Chen, R. Li, P. Ren, L. Zhang, and H. Mao (2017). Health burden attributable to ambient pm2. 5 in china. *Environmental pollution 223*, 575–586.

Sousa, R. M., C. P. Ferri, D. Acosta, E. Albanese, M. Guerra, Y. Huang, K. Jacob, A. Jotheeswaran, J. J. L. Rodriguez, G. R. Pichardo, M. C. Rodriguez, A. Salas, A. L. Sosa, J. Williams, T. Zuniga, and M. Prince (2009). Contribution of chronic diseases to disability in elderly people in countries with low and middle incomes: A 10/66 dementia research group population-based survey. *The Lancet 374*(9704), 1821–30.

Sum, G., T. Hone, R. Atun, C. Millett, M. Suhrcke, A. Mahal, G. C.-H. Koh, and J. T. Lee (2018). Multimorbidity and out-of-pocket expenditure on medicines: a systematic review. *BMJ global health 3*(1), e000505.

Tas, Ü., A. P. Verhagen, S. M. Bierma-Zeinstra, E. Odding, and B. W. Koes (2007). Prognostic factors of disability in older people: a systematic review. *British Journal of General Practice 57*(537), 319–323.

Tran, P. B., J. Kazibwe, G. F. Nikolaidis, I. Linnosmaa, M. Rijken, and J. van Olmen (2022). Costs of multimorbidity: a systematic review and meta-analyses. *BMC medicine 20*(1), 234.

Vos, T., C. Allen, M. Arora, R. M. Barber, Z. A. Bhutta, A. Brown, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen, et al. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The lancet 388*(10053), 1545–1602.

Wallace, E., C. Salisbury, B. Guthrie, C. Lewis, T. Fahey, and S. M. Smith (2015). Managing patients with multimorbidity in primary care. *Bmj 350*.

Weller, R. B. (2020). Beneficial effects of sunlight may account for the correlation between serum vitamin d levels and cardiovascular health. *JAMA cardiology 5*(1), 109–109.

WHO, A. (2020). Who methods and data sources for global burden of disease estimates: 2000-2019.

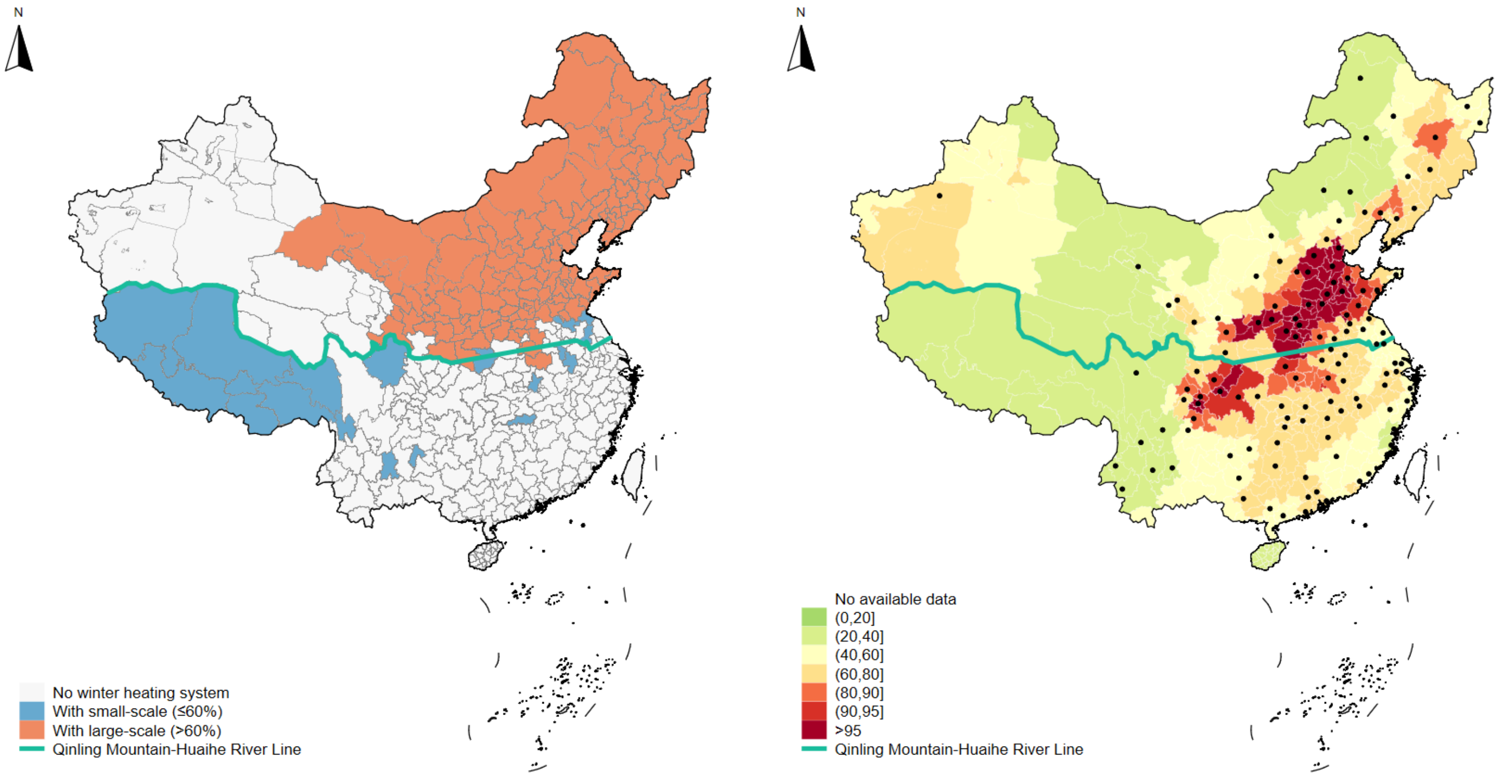
Yazdi, M. D., Y. Wang, Q. Di, W. J. Requia, Y. Wei, L. Shi, M. B. Sabath, F. Dominici, B. Coull, J. S. Evans, et al. (2021). Long-term effect of exposure to lower concentrations of air pollution on mortality among us medicare participants and vulnerable subgroups: a doubly-robust approach. *The Lancet Planetary Health 5*(10), e689–e697.

Yin, P., M. Brauer, A. J. Cohen, H. Wang, J. Li, R. T. Burnett, J. D. Stanaway, K. Causey, S. Larson, W. Godwin, et al. (2020). The effect of air pollution on deaths, disease burden, and life expectancy across china and its provinces, 1990–2017: an analysis for the global burden of disease study 2017. *The Lancet Planetary Health 4*(9), e386–e398.

Zhong, J., X. Zhang, K. Gui, J. Liao, Y. Fei, L. Jiang, L. Guo, L. Liu, H. Che, Y. Wang, et al. (2022). Reconstructing 6-hourly pm 2.5 datasets from 1960 to 2020 in china. *Earth System Science Data 14*(7), 3197–3211.

Zhou, M., H. Wang, X. Zeng, P. Yin, J. Zhu, W. Chen, X. Li, L. Wang, L. Wang, Y. Liu, et al. (2019). Mortality, morbidity, and risk factors in china and its provinces, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet 394*(10204), 1145–1158.

6 Figures



(A) **Cities with public winter heating**

(B) **Long-term PM2.5 concentration (1980-2010)**

FIGURE 1
China's Winter Heating Policy and Long-term PM2.5 Concentration

Notes: In panel (A), The red areas represent cities completely coverd by winter heating system, the blue areas refer to cities partly coverd by winter heating system, and cities in the white zone do not have winter heating system. In panel (B), cities in yellower areas have higher levels of PM2.5 concentration, and white areas lack corresponding pollution data. The green line in the map represents the Qinling-Huaihe Line. The cities marked with black dots are sampled in CHARLS.

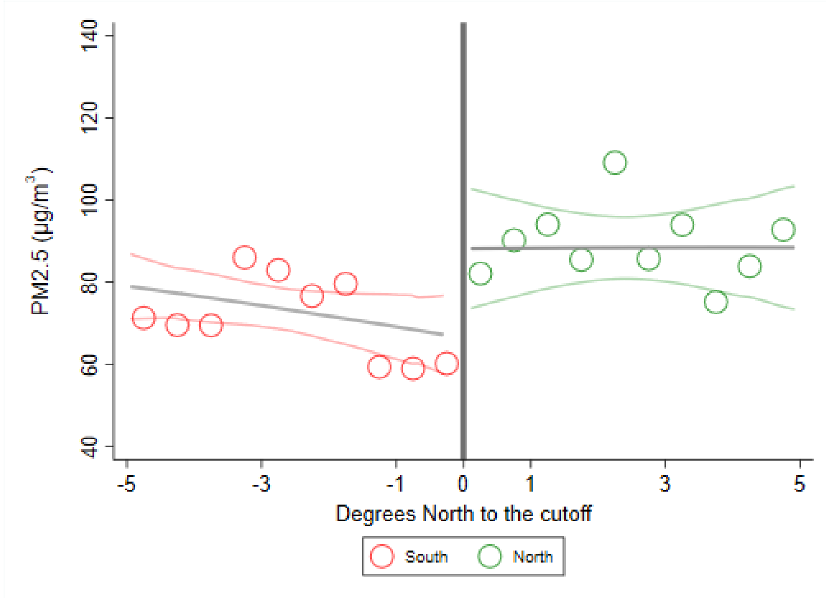


FIGURE 2
Average PM2.5 concentrations (1980-2010) in the north and south of the Qinling-Huaihe line

Notes: This figure display the average PM2.5 concentrations from 1980 to 2010 in 0.5 latitude bins. The sample encompasses 121 cities located within 16 latitudes to the Qinling-Huaihe line, which serves as the reference line. Each circle represents a 0.5-degree latitude bin. We plot means in bins and include estimated local polynomial fit lines (with a 95% confidence interval band) on each side of the reference line. Data source for PM2.5 concentrations is Zhong et al. (2022).

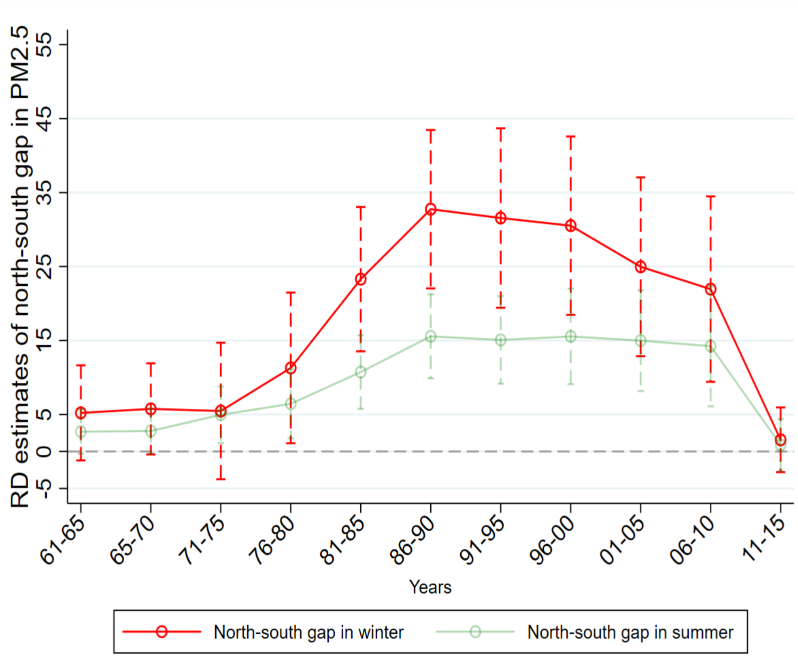
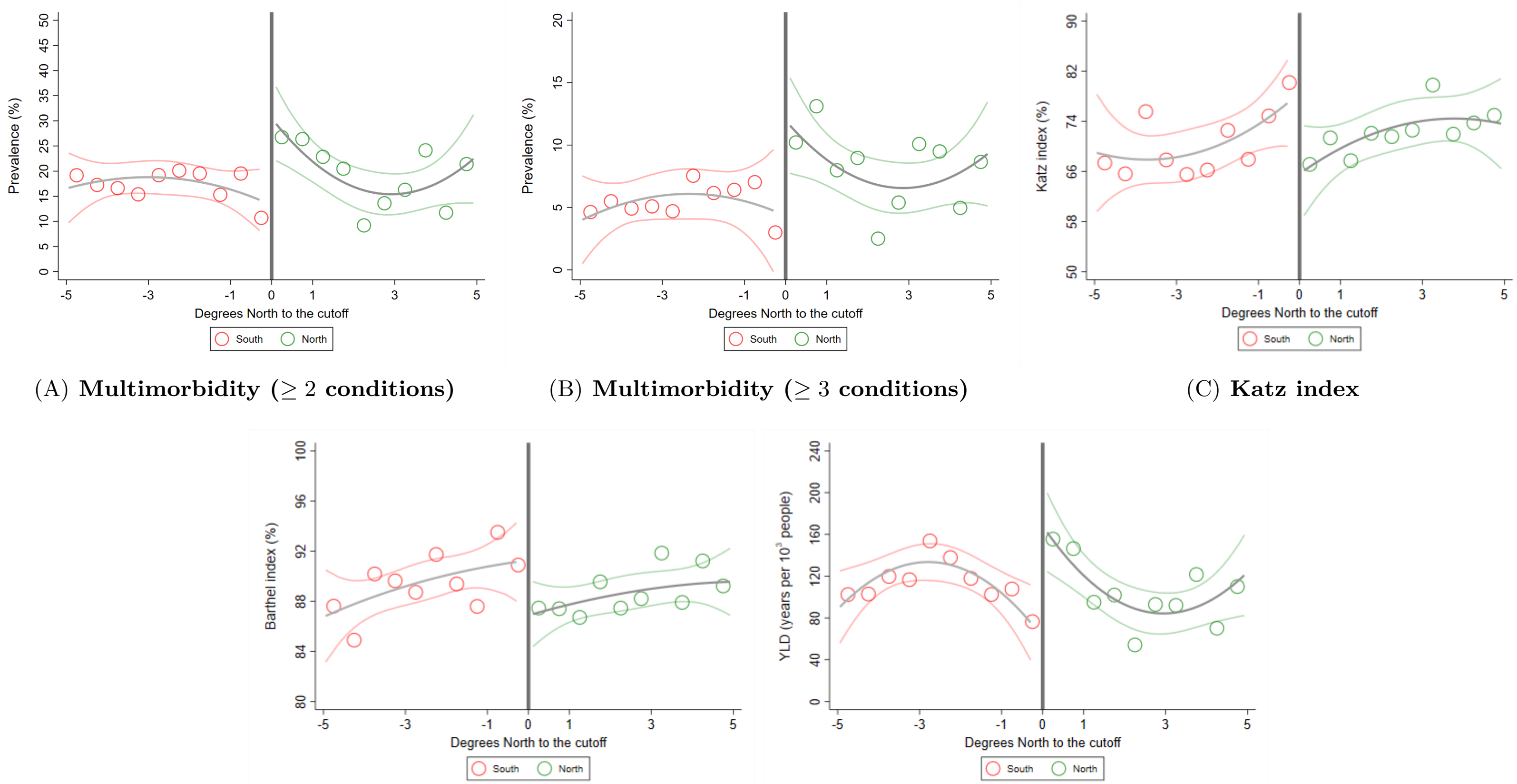


FIGURE 3
RD estimates of the north-south gap in PM2.5 from 1960 to 2015

Notes: The figure displays the dynamic pattern of the north-south gap of PM2.5 concentrations in 5-year intervals from 1960 to 2015. Each circle represents the RD estimate of the north-south gap in PM2.5 over a 5-year interval, using the regression equation 1 in Section 3. Each regression adopts the mean square error (MSE) optimal bandwidth (Imbens and Kalyanaraman, 2012; Calonico et al., 2014, 2020). Each vertical line represents the 95% confidence interval of the corresponding RD estimate. The red lines represents the estimates in winter heating season (November-March), while the green line for the non-heating season (April-September). Data source for PM2.5 concentrations is Zhong et al. (2022).



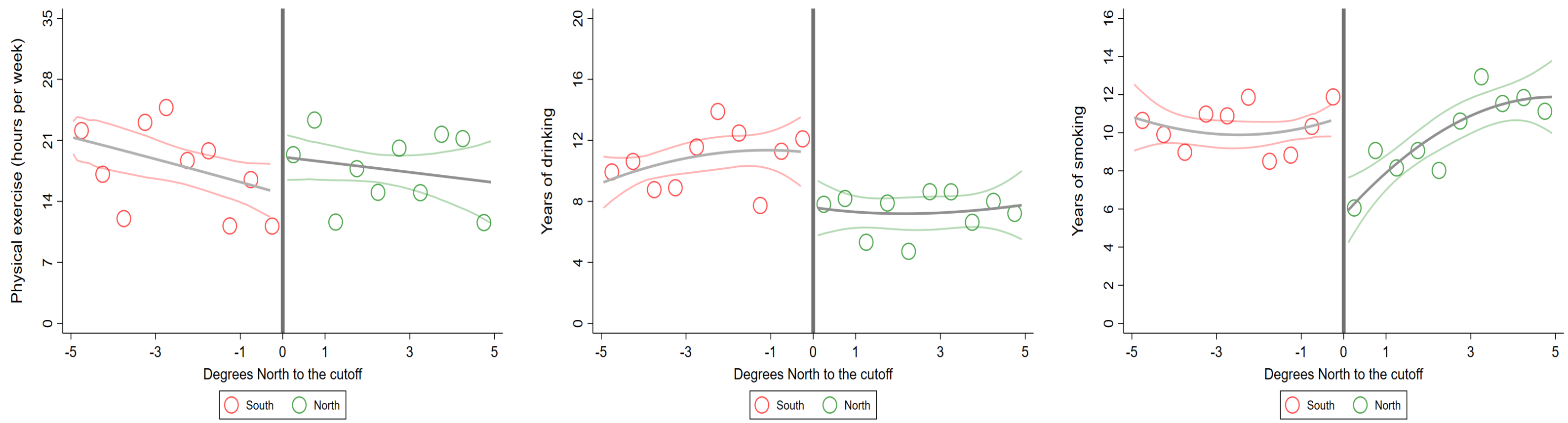
(A) **Multimorbidity ($\geq$ 2 conditions)** (B) **Multimorbidity ($\geq$ 3 conditions)** (C) **Katz index**

(D) **Barthel index** (E) **Years lived with disability (YLD)**

FIGURE 4

Average Health Outcomes in the North and South of the Qinling-Huaihe Line

Notes: This figure displays the average prevalence rates of multimorbidity (Panels A and B), the rates of functional disabilities (Panels C and D), and the average years lived with disability (YLD) in 0.5-latitude bins. The analysis encompasses 121 cities located within 16 latitudes to the Qinling-Huaihe line, which serves as the reference line. The figure presents bin means and includes local polynomial fit lines on both sides of the reference line. Data source for all health measures is CHARLS 2011 baseline survey. See Section 2 and Appendix Section B for further details on variable definition and constructions.



(A) **Physical exercises** (B) **Alcohol drinking** (C) **Smoking**

FIGURE 5
Average Health Behaviors in the North and South of the Qinling-Huaihe Line

Notes: This figure displays the average frequency of physical exercises (hours per week) in Panel (A), the average years of drinking in Panel (B), and the average years of smoking in Panel (C) in 0.5-latitude bins. The analysis encompasses 121 cities located within 16 latitudes to the Qinling-Huaihe line, which serves as the reference line. The figure presents bin means and includes local polynomial fit lines on both sides of the reference line. Data source for all health outcome measures is 2011 baseline survey. See Section 2 and Appendix Section B for further details on variable definition and constructions.

7 Tables

TABLE 1 Summary Statistics of Outcome Variables

Outcomes	North	South	Raw Diff (1)-(2)
PM25 ($\mu g/m^3$)	77.658 (25.561)	63.380 (17.906)	14.278 (3.979)
Multimorbidity ($\geq$ 2 conditions)	0.233 (0.119)	0.159 (0.074)	0.074 (0.018)
Multimorbidity ($\geq$ 3 conditions)	0.103 (0.068)	0.052 (0.041)	0.051 (0.010)
Katz index	0.719 (0.084)	0.695 (0.081)	0.025 (0.015)
Barthel index	0.886 (0.031)	0.887 (0.032)	-0.002 (0.006)
YLD (per 10^3 people per year)	122.759 (53.025)	111.062 (38.561)	11.697 (8.366)
Observations	57	64	121

Notes: This table presents the average values of PM2.5 concentration, health outcomes of cities in the north and south of the reference line. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line, of which 57 are in the north and 64 are in the south. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. Standard deviations are reported in parentheses.

TABLE 2 RD Estimated North-south Gap in PM2.5 Levels, Multimorbidity, Disability, and YLD

Outcomes	(1)	(2)
Panel A: Long-term PM2.5 concentration (1980-2010)		
PM25 ($\mu g/m^3$)	26.770***	26.701***
	(5.625)	(5.206)
Bandwidth (North)	2.140	2.407
Bandwidth (South)	2.526	2.508
Panel B: Multimorbidity		
Multimorbidity ($\geq$ 2 conditions)	0.096***	0.096***
	(0.031)	(0.027)
Bandwidth (North)	2.561	3.327
Bandwidth (South)	3.283	3.214
Multimorbidity ($\geq$ 3 conditions)	0.054**	0.045**
	(0.026)	(0.022)
Bandwidth (North)	2.646	4.138
Bandwidth (South)	3.150	2.976
Panel C: Disability		
Katz index	-0.133***	-0.131***
	(0.035)	(0.031)
Bandwidth (North)	4.128	4.916
Bandwidth (South)	2.307	2.249
Barthel index	-0.079***	-0.080***
	(0.014)	(0.014)
Bandwidth (North)	4.178	4.490
Bandwidth (South)	2.406	2.316
Panel D: Quality of life		
YLD (per 10^3 people per year)	73.664***	73.400***
	(16.983)	(15.978)
Bandwidth (North)	2.082	2.266
Bandwidth (South)	2.784	2.772
Observations	121	121
Longitude-region FEs	Yes	Yes
Control Variables	Yes	Yes
Bandwidth selection	MSE	MSE
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in long-term PM2.5 concentrations and population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 1 and 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE 3 IV Estimated Marginal Effect of PM2.5 on Multimorbidity, Disability, and YLD

	(1)	(2)	(3)	(4)
Panel A: Marginal effects on multimorbidity				
Outcomes:	Multimorbidity ($\geq$ 2 conditions)		Multimorbidity ($\geq$ 3 conditions)	
PM2.5 (10 $\mu g/m^3$)	0.052***	0.051***	0.017*	0.018*
	(0.017)	(0.016)	(0.010)	(0.010)
Bandwidth (North)	2.201	2.496	2.236	2.556
Bandwidth (South)	2.410	2.387	4.049	3.775
Panel B: Marginal effects on disability				
Outcomes:	Katz Index		Barthel Index	
PM2.5 (10 $\mu g/m^3$)	-0.036**	-0.039**	-0.015*	-0.016*
	(0.017)	(0.016)	(0.009)	(0.009)
Bandwidth (North)	2.133	2.604	1.896	2.146
Bandwidth (South)	3.394	3.281	2.943	3.095
Panel C: Marginal effects on quality of life				
Outcomes:	YLD			
PM2.5 (10 $\mu g/m^3$)	21.393**	26.039***		
	(8.674)	(8.828)		
Bandwidth (North)	1.991	2.249		
Bandwidth (South)	3.532	2.928		
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth selection	MSE	MSE	MSE	MSE
Kernel Function	Epa	Tri	Epa	Tri

Notes: This table reports the estimated marginal effect of ten-unit increase (10 $\mu g/m^3$) in long-term PM2.5 concentrations on population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and that on health outcomes are from CHARLS 2011 survey. We estimate the impact of PM2.5 on the listed health outcomes using local linear regression, treating distance from the Qinling-Huaihe line as the forcing variable and PM2.5 as the treatment variable, with the Qinling-Huaihe line representing a "fuzzy" discontinuity in the level of long-term PM2.5 exposure. We also adopt local linear distance function, epanechnikov kernel (Columns 1 and 3) or triangular kernel (Columns 2 and 4), optimal minimal standard error (MSE) bandwidth, and control for a set of city-level variables in regression analysis. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE 4 RD Estimated North-south Gap in Health Outcomes across Different Socioeconomic Groups

Socioeconomic Group	**Education Attainment**		**Income level**	
	Low	High	Low	High
Panel A: Multimorbidity				
Multimorbidity ($\geq$ 2 conditions)	0.132***	0.101***	0.183***	0.073*
	(0.047)	(0.030)	(0.027)	(0.039)
Bandwidth (North)	2.020	2.302	2.208	2.465
Bandwidth (South)	2.623	3.196	2.062	2.033
Multimorbidity ($\geq$ 3 conditions)	0.094***	0.029	0.067**	0.044
	(0.030)	(0.025)	(0.030)	(0.033)
Bandwidth (North)	2.162	2.125	2.021	1.948
Bandwidth (South)	2.493	2.826	2.041	2.489
Panel B: Disability				
Katz index	-0.194***	-0.119***	-0.389***	-0.128**
	(0.055)	(0.043)	(0.081)	(0.052)
Bandwidth (North)	3.474	4.334	4.546	3.288
Bandwidth (South)	2.812	2.662	1.735	2.361
Barthel index	-0.108***	-0.079***	-0.143***	-0.063**
	(0.035)	(0.020)	(0.020)	(0.026)
Bandwidth (North)	2.173	4.519	3.676	3.290
Bandwidth (South)	2.217	1.720	2.245	2.315
Panel C: Quality of life				
YLD (per 10^3 people per year)	122.050***	48.174***	101.045***	46.169*
	(22.697)	(16.209)	(18.164)	(23.662)
Bandwidth (North)	1.500	2.658	1.866	2.822
Bandwidth (South)	2.394	2.820	2.196	2.311
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth Selection	MSE	MSE	MSE	MSE
Kernel Function	Epa	Epa	Epa	Epa

Notes: This table reports the estimated north-south gap in population health outcomes across different socioeconomic groups. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on health outcomes are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 2 based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE 5 RD Estimated North-south Gap in Life-style Adjustments across Different Socioeconomic Groups

Socioeconomic Group	Educational background		Income level	
	Low	High	Low	High
Frequency of exercise (hrs/week)	8.373***	13.664*	9.399	12.648*
	(3.171)	(7.714)	(7.071)	(6.926)
Bandwidth (North)	2.445	3.428	2.903	3.300
Bandwidth (South)	2.642	2.418	2.488	2.206
Years of alcohol drinking	-1.227	-4.864*	-3.374	-6.722*
	(2.464)	(2.775)	(2.299)	(4.043)
Bandwidth (North)	2.415	2.921	2.386	2.833
Bandwidth (South)	2.159	2.177	2.626	2.075
Years of smoking	-2.177	-5.179***	-2.053*	-6.571*
	(1.846)	(1.376)	(1.186)	(3.492)
Bandwidth (North)	2.847	2.832	4.602	2.593
Bandwidth (South)	2.501	2.455	2.061	2.427
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth Selection	MSE	MSE	MSE	MSE
Kernel Function	Epa	Epa	Epa	Epa

Notes: This table reports the estimated north-south gap in life-style Adjustments across different socioeconomic groups. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on life-style are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 2 based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

Appendices

(For Online Publication Only)

This appendix contains online supplementary material. In this appendix, we collect the analyses, discussions, figures, and tables omitted from the main text.

Appendix A Policy and Geographical Contexts for the Empirical Strategy

A.1 China's Winter Heating Policy and the Choice of Pollutants

China's winter heating policy began in the 1950s. The Soviet Union was the country with the largest central heating scale in the world at that time. China followed the Soviet model and initially established a residential boiler heating system. Based on geography, climate, budget, energy shortage, and some social factors, the central government roughly divides China into the south and the north with the Qinling-Huaihe line as the boundary, and only provide district heating for northern cities in winter. But it wasn't until the 1980s that district heating became widespread in northern China. During the period from November 15th to March 15th of the following year, most cities in northern China can enjoy free or highly subsidized heating. For some northern cities with very cold winters (such as Harbin in Heilongjiang), the heating season will be extended to more than 6 months, from October of this year to April of the following year (Fan et al., 2020).

At that time, the main fuel used in China's wintering heating system was coal, which was utilized very inefficiently. The incomplete combustion of coal will produce a large amount of air pollution, such as particulate matter (especially PM2.5) and sulfur dioxide (SO2) (Almond et al., 2009; Muller et al., 2011). Therefore, whenever winter comes and the central heating system starts to work, air pollution levels in northern Chinese cities will increase rapidly, much higher than in southern cities during the same period. This provides a quasi-experimental setting for us to use the discontinuity in air pollution induced by China's winter heating policy to estimate the health effects of long-term air pollution exposure.

We chose PM2.5 as an air pollution indicator for the following two reasons. First, as early as 1998, under the background of serious environmental pollution, the Chinese government proposed the "two-control zone policy". That is, according to natural conditions such as meteorology, topography, soil, etc., the areas where acid rain had occurred or may occur or areas with severe sulfur dioxide pollution would be designated as acid rain control areas or sulfur dioxide pollution control areas, which would be heavily regulated to reduce pollution levels. Ebenstein et al. (2017) shows that China's winter heating policy does not cause a significant difference in the concentration of sulfur dioxide and nitrogen oxide (NOx) between the north and the south. In contrast, the Chinese government first proposed to control PM2.5 pollution in 2013, and our calculations show that the annual average PM2.5 concentration from 1980 to 2010 does have a significant discontinuous rise at the cutoff. Second, extensive researches clearly point to the health effects of rising PM2.5 concentrations (Lu et al., 2015; Song et al., 2017; Bu et al., 2021). And Although control of PM2.5 pollution is already underway in current China, we find the PM2.5 concentration is much higher that the recommended level (Pai et al., 2022), indicating that PM2.5 pollution is still an important threat to residents' health that cannot be ignored (Yin et al., 2020).

A.2 The Qinglin-Huaihe line

The Qinglin-Huaihe line is located at about 32 to 35-degree north latitude, which basically coincides with the 800 mm annual precipitation line and the 0 degree centigrade isotherm in January [1]. It is also the dividing line between the subtropical monsoon climate and the temperate monsoon climate. There are significant geographic differences in the northern and southern regions of China. A large number of studies have shown that weather factors such as temperature, precipitation, and sunshine are related to the prevalence of chronic conditions and the resulting mortality (e.g., Braga et al. (2002); Goggins and Chan (2017); Schinasi et al. (2020); Weller (2020); Liu et al. (2021)). Therefore, in order to ensure the validity of our RD design and obtain more accurate result, we must take these meteorological conditions into consideration.

Appendix B Variable Construction and Description

B.1 Sample Restriction

The baseline sample contains 17,708 respondents from 125 prefecture-level cities in 28 provinces. We aggregate individual-level data to the city level and calculate the population-weighted, average prevalence rate of multimorbidity and other health out-

1. The 800 mm isotherm is the demarcation index between humid and semi-humid areas, and the meaning of the zero-degree isotherm determines whether the river freezes in winter, that is, whether the average temperature is above zero.

comes. We match the city-average PM2.5 levels and other pollution measures to the CHARLS cities, and obtain a set of relevant city-level demographic and socioeconomic variables such as age structure, GDP per capita, and industrial structure from various city statistical yearbooks. We drop 3 cities with missing values on PM2.5 levels or key covariates (mostly minority autonomous regions). We also drop Shanghai, which has the highest GDP per capita during the sample period, respectively, to avoid unobserved socioeconomic factors from confounding our analysis. Finally, we obtain an analytical sample of 121 cities.

B.2 Chronic Conditions

We use the data on the prevalence chronic conditions from the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is hosted by the National Development Institute of Peking University and implemented by the China Social Science Survey Center. CHARLS selects a total of 125 prefecture-level cities for the survey, and its questionnaires cover basic personal information, family structure and financial support, health status, anthropocentric measurements, medical service utilization and medical insurance, work, retirement and pensions, income, consumption, assets, and the basic situation of the community, etc. The data quality of this survey is particularly high and has been widely used and recognized in academic world.

A total of 14 categories of chronic conditions are considered in the CHARLS questionnaire. Since this paper does not investigate the impact of air pollution on mental and memory diseases, there is no literature showing that air pollution has direct influence on arthritis or rheumatism, and limited data for the prevalence of cancer and stroke, this paper only examine how long-term exposure to PM2.5 affect the multimorbidity composed of 9 categories of chronic conditions, including hypertension, diabetes, dyslipidemia, asthma, heart diseases, lung diseases, kidney diseases, liver diseases and stomach diseases. Since several types of these conditions encompass a wide variety of specific diseases, we do not use disease-specific terminology in this paper. A detailed description of these conditions is given in Table B1.

B.3 Calculation of average PM2.5 concentration

Our measure is based on a database constructed by Zhong et al. (2022). The historical PM2.5 concentrations are constructed at a $0.25° \times 0.25°$ longitude-latitude grid every 6 hours. The authors adopt a Light Gradient Boosting Machine (LightGBM) model and combine long-term visibility from satellite images, conventional meteorological observations, ground monitored pollution emissions, and elevation. This is the first historical database of PM2.5 at such long time horizon, high temporal resolution, and fine geographic resolution. We verify in Appendix Table A1 that the annual PM2.5 levels constructed from Zhong et al. (2022) have strong correlations and good accuracy with ground-monitor-based records after 2014, when the latter became available after the nationwide construction of modern air pollution ground monitoring stations.

B.4 Meteorological Conditions and Socioeconomic Covariates

In order to verify the validity of our empirical strategy and adjust the estimates for potential confounders, we collect annual climate panel data and socioeconomic characteristics at the city level from the Chinese daily surface meteorological data set (V3.0) of the National Meteorological Science Data Center and China Urban Statistical Yearbook respectively. The data for the meteorological conditions are their annual means from 1980 to 2010, while the data on socioeconomic factors are their annual means from 1990 to 2011. Table B6 gives their summary statistics.

B.5 Other Health-related Variables

Demographic variables, including share of married population, age structure of sampled population, share of sampled population with at least 6 years of education and share of males of targeted population are all collected from CHARLS 2011 survey.

Meanwhile, we construct several channel variables based on individual's health-related behaviors. In terms of physical exercises, CHARLS survey in 2011 gathered data for individuals' frequency of three types of physical activities, including vigorous activities, moderate-intensity activities, and walking. Interviewers would asked respondents how many days each week they participated in these activities, and the day should be counted only if the aggregate time for certain type of activities in that day is at least 10 minutes. Meanwhile, interviewers count whether the respondents participate in such activities per day for less than 30 minutes, or greater than or equal to 30 minutes and less than two hours, or greater than or equal to two hours and less than 4 hours, or greater than or equal to 4 hours. In order to obtain the total time of each type of exercise per week for each respondent, we set the longest time for each type of activities per day as 4 hours, and take the midpoint value of each time interval as the representative time for activities per day. For example, when the respondents done certain type of activities for more than or equal to 30 minutes and less than two hours, we considered their representative exercise time to be 1.25 hours. Further the average time for physical exercises at the city level could be calculated.

For the two behaviors of smoking and drinking, CHARLS collected the age at which the respondents started smoking and alcohol drinking, and the age when they quit smoking and drinking, respectively. Based on it, we obtain the number of years of smoking and alcohol drinking of the respondents, and calculate the mean at the city level like before.

With regard to sleeping, CHARLS survey counted the time each respondent actually fell asleep during the night, and we calculate the city-level average actual sleeping time from that. Finally, since the CHARLS questionnaire only counts the number of meals the respondents ate per day which could measure the intakes of various nutrients well, we use another nationally representative survey data, the China Health and Nutrition Survey (CHNS), to construct the nutrient intakes variables. The China Health and Nutrition Survey (CHNS) is an ongoing open-cohort international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health (NINH) of the Chinese Center for Disease Control and Prevention (CCDC), which was designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments, and to see how the social and economic transformation of Chinese society affects the health and nutritional status of its population. Its rigorous sampling method and questionnaire design ensure the reliability of data quality. CHNS counted the 3-day average intakes of 4 types of nutrients in total, including carbohydrates, fats, proteins and calories. We use the data from CHNS 2011 survey. The construction of variables is the same as that described above.

B.6 Performance of Activities of Daily Living (ADLs)

In order to estimate the effect of long-term PM2.5 exposure on onset of disability, we construct two types of indicators to measure the degree of disability in an individual. We measure an individual's degree of disability from the perspective of daily living behaviors. Disability is commonly defined as a difficulty in performing activities necessary for independent living, such as basic activities of daily living (ADLs) and complex instrumental activities of daily living (IADLs) (Tas et al., 2007; Ćwirlej-Sozańska et al., 2019). The concept of ADLs, which was originally proposed in the 1950s by Sidney Katz and his team at the Benjamin Rose Hospital in Cleveland, Ohio, includes the daily activities we perform for self-care, like feeding oneself, bathing, dressing, grooming, cleaning oneself after defecating, and has so far been developed and refined by many researchers (Noelker and Browdie, 2014). Researchers have proposed a variety of scales to measure a person's ability to perform ADLs such as the Katz scale and Barthel scale.The CHARLS investigated relative questions raised by these scales so that we could calculated the Katz index and the Barthel index. The lower these two indices, the lower the individual's ability to live independently, that is, the higher the degree of disability. The Barthel index is usually measured as a score out of 100. In order to make it comparable to the Katz index, we divide it by 100 to normalize it to a ratio between 0 and 1. The details of the questions, answers and evaluation methods proposed by CHARLS and these two scales are presented in Table B2, Table B3 and Table B4 [2].

B.7 Construction of YLD

Constructing Years Lived with Disability (YLD) involves several steps to quantify the burden of specific health conditions on the population. YLD is a measure used in the Global Burden of Disease (GBD) study to estimate the impact of non-fatal health outcomes. We construct YLD in a given city in four steps:

First, we identify the nine most common chronic conditions. These conditions include hypertension, dyslipidemia, diabetes, asthma, COPD, other chronic respiratory diseases, cardiovascular diseases, liver diseases, and kidney diseases. Cancer and stroke is not included in our analysis due to its extremely low self-reported incidence in our sample (we observe zero cancer and stroke case in 60 and 63 out of 121 cities in our baseline sample). Specifically, the prevalence of a given condition is calculated as the number of individuals affected by the condition divided by the total sample size of each city.

Second, disability weights are assigned to each health condition to capture the impact on individuals' functioning and quality of life. We obtain the disability weight for each chronic condition from the WHO global burden of disease study (Salomon et al., 2015). These weights range from 0 to 1, with 0 representing perfect health and 1 indicating a health state equivalent to death. Disability weights are generally derived through various methods such as population surveys, expert consensus, or disability-adjusted life year (DALY) valuation exercises. Since several types of these conditions encompass a wide variety of specific diseases, we calculate the average disability weights of specific diseases for these chronic conditions. We use the disability weight for hypertensive heart disease as hypertension's disability weight. And we set the disability weight for dyslipidemia to zero because it is treated as a type of risk factor instead of chronic disease. The disability weights for all chronic conditions are presented in Table B5.

2. Due to data limitations, we group whether the individual could dress and groom herself independently into one question

Third, we get YLD of each health condition by multiplying its prevalence with the corresponding disability weight. This accounts for the varying severity of different conditions and their impact on overall disability burden.

Finally, total burden of non-fatal health outcomes in the population is computed by aggregating the YLD of each health condition with comorbidity adjustment (WHO, 2020). This calculation yields the number of healthy years lost due to living with these chronic conditions.

It's important to note that constructing YLD involves several data limitations. The accuracy of the estimates depends on the quality and availability of data on prevalence for each chronic conditions as reported in CHARLS survey. We expect that our estimated YLD serves as a lower bound because of potential underreporting of disease incidences in CHARLS survey. For example, the self-reported incidence of cancer and stroke is extremely low in our sample.

Overall, YLD is a valuable metric for understanding the non-fatal burden of disease, complementing measures such as Years of Life Lost (YLL) that focus on premature mortality. It provides insights into the impact of health conditions on individuals' well-being, productivity, and overall quality of life. These estimates are crucial for informing public health policies, resource allocation, and healthcare planning to address the burden of disease and promote population health.

Appendix C Robustness Checks

C.1 Validating the RD design

We first assess the validity of applying the Qinling-Huaihe line as a geographic reference line in our RD analysis, as well as the RD assumption that all demographic and economic characteristics are smooth over distance to this reference line.

Figure 1, Panel (A) maps out the city-level implementation of winter heating policy and highlights the Qinling-Huaihe line as a policy boundary for the winter heating policy, with only a small number of cities did not exactly conform to the north-only rule. In the RD analysis, we follow the standard practice of Ito and Zhang (2020) and use the difference in latitude of cities relative to the Qinling-Huaihe line as the running variable. Specifically, due to the variation of latitude range covered by the Qinling-Huaihe line in different areas, we divide the river line into three segments, and in each segments, we measure a city's relative latitude to the middle point of the line's latitude range. Hence, we could control the regional fixed effect to avoid the endogeneity problem bias our estimates.

We adopt a falsification RD specification based on placebo reference line to assess the validity of this reference line. We verify that PM2.5 only changed discontinuously at the true reference line rather than any other placebo references. Results from these tests are presented in Figure A4.

Next, we test the smoothness of critical health-related factors over distance to this reference line. First, Figure A2 and Table B9 shows that nutrition intakes did not show any discontinuity change between the north and south. Then, Figure A5 and Table B10 show that city-level socioeconomic development, demographic characteristics and meteorological conditions are all similar on both sides of the reference line. In particular, population of cities on both sides of the line had similar demographic structure and faced similar economic circumstances and weather conditions. Finally, water pollution emissions from industrial activities did not exhibit any sharp difference between the north and south. Moreover, Ebenstein et al. (2017) found that the sulphur dioxide and nitrogen oxide concentration were also similar between north and south of the line. This helps rule out the possibility that the discontinuously higher level of air pollution in the north is driven by north-south differences in economic activities or other pollutant emissions.

Overall, we find that residents living in the north and south of the line had similar pattern of nutrition intakes, demographic structure, faced similar weather conditions and socioeconomic circumstances, but experienced a sharply higher level of PM2.5 concentration.

C.2 Endogenous Sorting Around the Cutoff

The identification of Sharp RD relies on the fact that the treatment is completely determined by the running variable and cannot be manipulated (Lee and Lemieux, 2010). The running variable in our setting is the latitude of the city and residents of cities north of the cutoff may choose to migrate to southern cities to avoid the health hazards of air pollution. On the one hand, We believe that people will only choose to move in cities with lower levels of air pollution. If this happens, the negative impacts of air pollution on chronic conditions we estimate are actually a lower bound on the true results, which does not affect our main conclusions. Meanwhile, Ding et al. (2021) points out that the mobility of the sampled population of CHARLS is low. On the other hand, we conduct McCrary test (McCrary, 2008) and density test (Cattaneo et al., 2020, 2021) in Figure A3, which indicates no jump in density of the running variable and demonstrate our argument above.

C.3 Sensitivity of Estimates to the bandwidth, Optimal Bandwidth Selection and Order of Local Polynomial Regression

We test wider bandwidths to check the sensitivity of our results to different bandwidths in Figure A7. In Table B11 and Table B12 we use another optimal bandwidth selection method, the CER-optimal bandwidth selection, for estimation. And the coefficient estimates obtained by local quardratic polynomial regression are shown in Table B13 and Table B14. Overall, it can be found that the results of the above tables are basically consistent with the results in the text, indicating that our conclusions are immune to bandwidth change, the optimal bandwidth selection method and the order of the polynomial.

C.4 Bias-corrected RD Estimates with Robust Variance Estimator

Calonico et al. (2014) points out that when using non-parametric methods to estimate RD design, several existing bandwidth selection methods usually lead to too small confidence intervals for the coefficient estimates, thus overly rejecting the null hypothesis. Based on this, they propose a bias-corrected estimator with a novel robust standard error to correct the resulting estimation bias. We report the estimated coefficients and standard errors using this method in Table B15 and Table B16, which are also consistent with that in the text.

Appendix D Figures

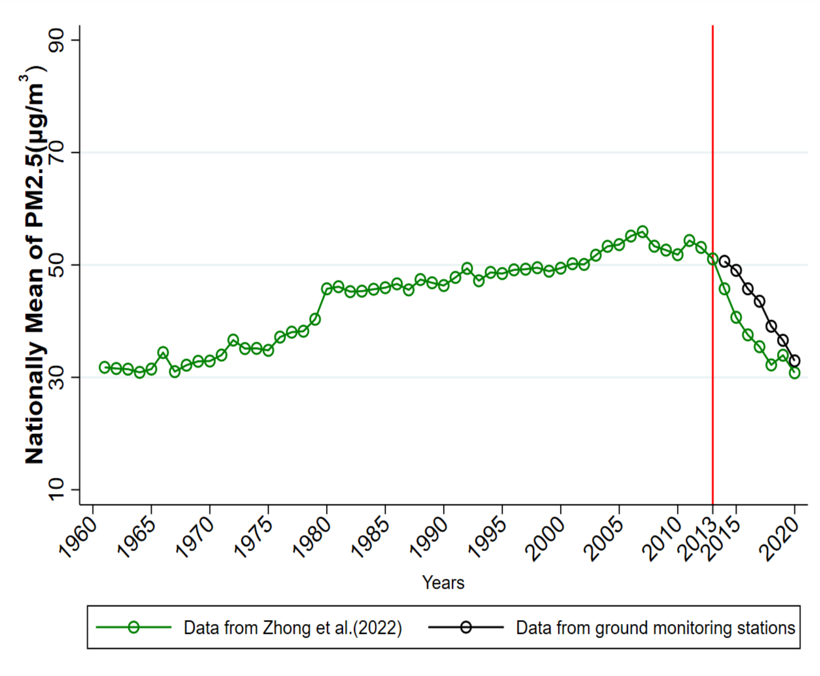
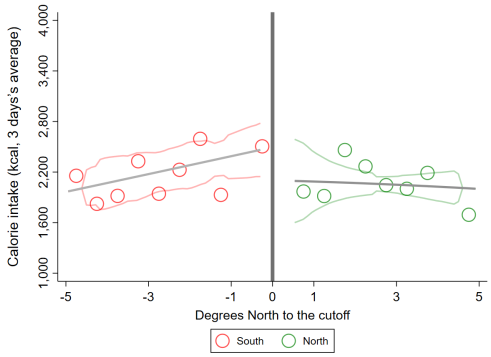
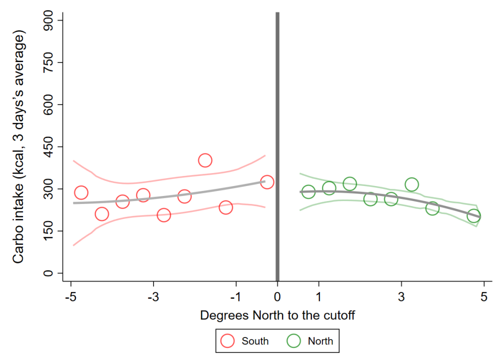


FIGURE A1
The annual trend of national PM2.5 concentration

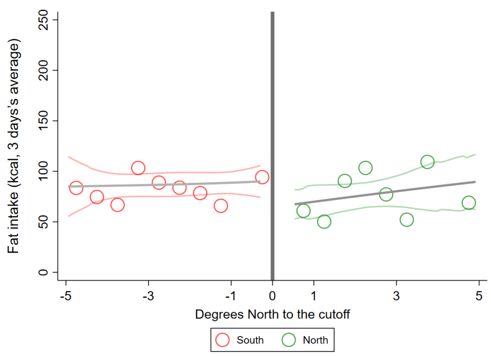
Notes: The green line represents the trend of annual PM2.5 concentration levels from 1960 to 2020 from (Zhong et al., 2022), while the black line represents the annual trend from 2014 to 2020 from ground-based monitoring stations. The red line indicates 2013.



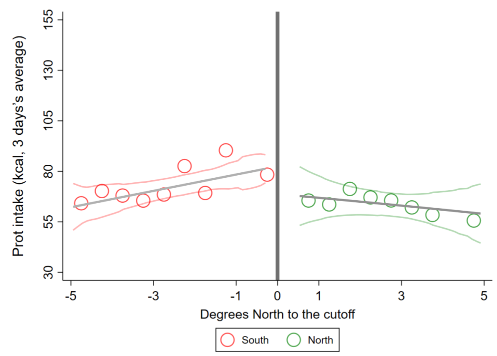
(A) Calorie intake (kcal)



(B) Carbohydrate intake (g)



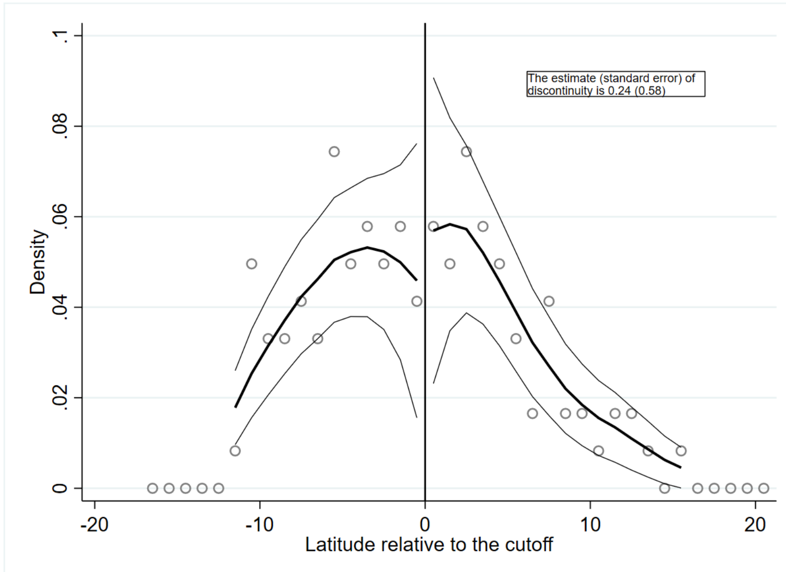
(C) Fat intake (g)



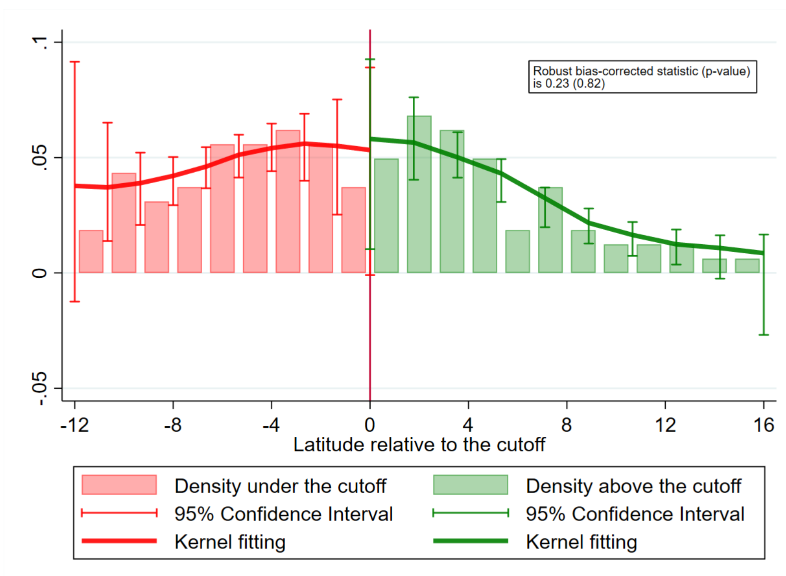
(D) Protein intake (g)

FIGURE A2
RD Estimates of North-South Gaps in Nutrition Intakes

Notes: This table reports the estimated north-south gap in nutrition intakes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on socioeconomic characteristics and industrial emissions are obtained from China Urban Statistical Yearbook. Data on nutrition intakes are from CHNS 2011 survey. Each dot represents the coefficient from a separate RD regression based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth and a set of city-level control variables. Each vertical line refers to confidence interval (CI) of the RD estimates.



(A) **[McCrary (2008)]'s Density Test**



(B) **Cattaneo et al. (2020, 2021)'s Density Test**

FIGURE A3

Notes: This figure plots the two density tests of latitudes of cities in CHARLS on both sides of the reference line. Panel (A) depicts the McCrary (2008)'s density test of discontinuity and Panel (B) shows (Cattaneo et al., 2020, 2021)'s density test. Both tests show there is no discontinuous change in density of the sample observations across the cutoff.

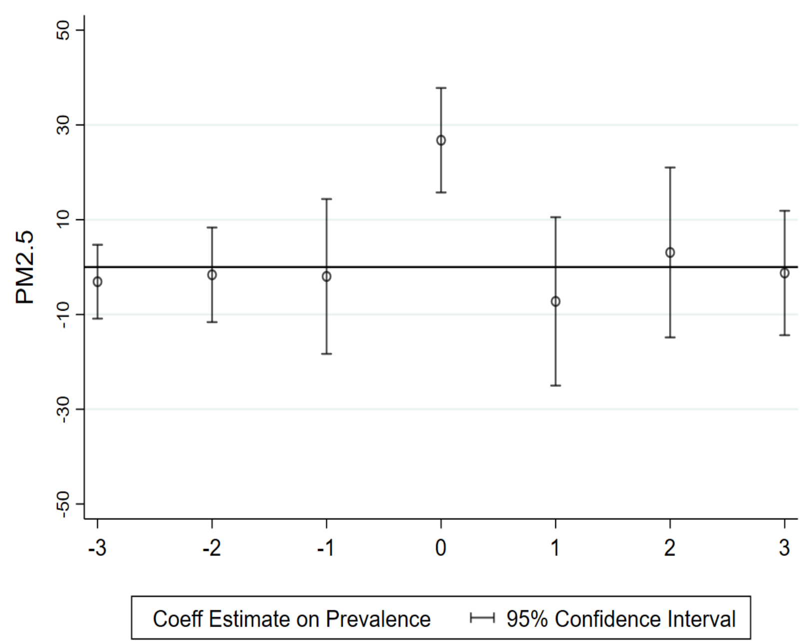
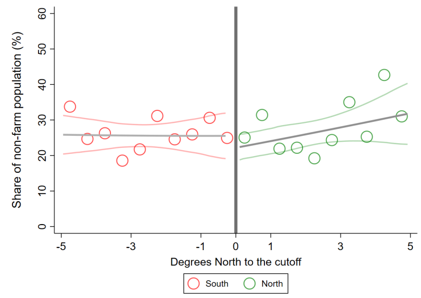
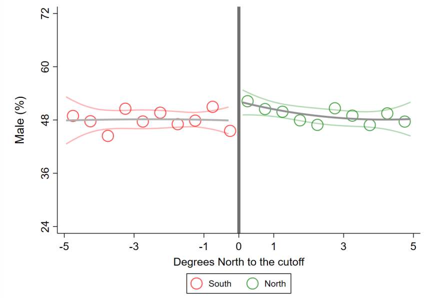


FIGURE A4
RD estimates of the change in PM2.5 concentration at different placebo cutoffs

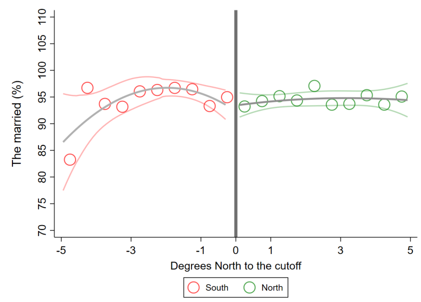
Notes: This figure plots the estimates of north-south gap in long-term PM2.5 concentrations across placebo reference lines. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022). Each dot represents the coefficient from a separate RD regression based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Each vertical line refers to confidence interval (CI) of the RD estimates.



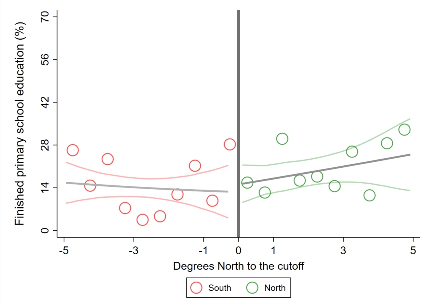
(A) Share of urban pop



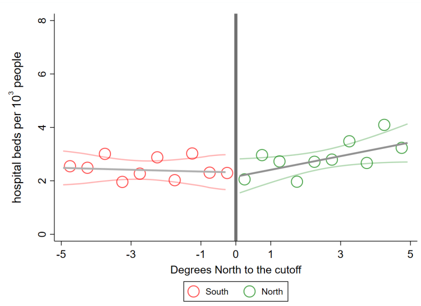
(B) Share of male pop



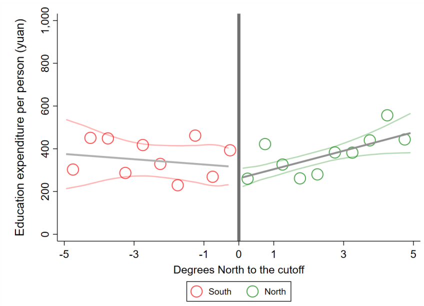
(C) Share of married



(D) Finished primary school



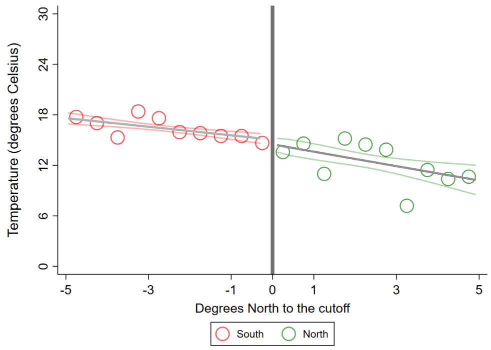
(E) Hospital beds per 10^3 people



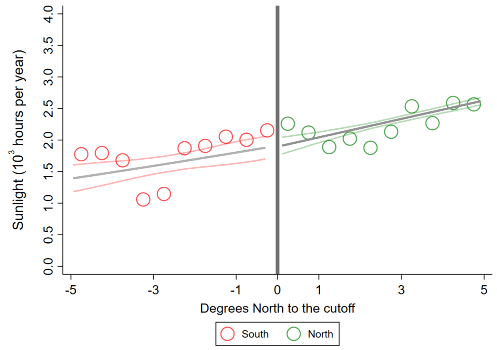
(F) Education expenditure per person (yuan)

FIGURE A5
RD Estimates of North-South Gaps in Socioeconomic Factors

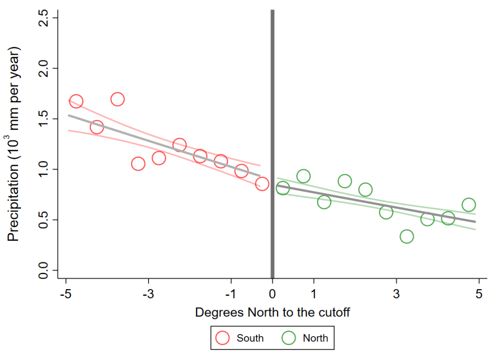
Notes: This table reports the estimated north-south gap in Socioeconomic factors. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on socioeconomic characteristics and industrial emissions are obtained from China Urban Statistical Yearbook. Data on demographic characteristics are from CHARLS 2011 survey. Data on meteorological characteristics are from the Chinese daily surface meteorological dataset (V3.0) provided by the National Meteorological Science Data Center. Both of socioeconomic characteristics and industrial emissions are annual average from 1990 to 2011, while the meteorological variables are annual average from 1980 to 2010. Each dot represents the coefficient from a separate RD regression based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth and a set of city-level control variables. Each vertical line refers to confidence interval (CI) of the RD estimates.



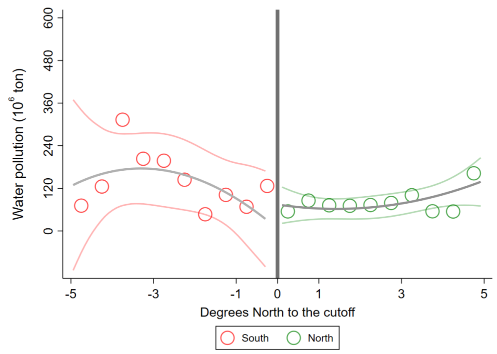
(A) Temperature (degrees Celsius)



(B) Sunlight (10^3 hours)



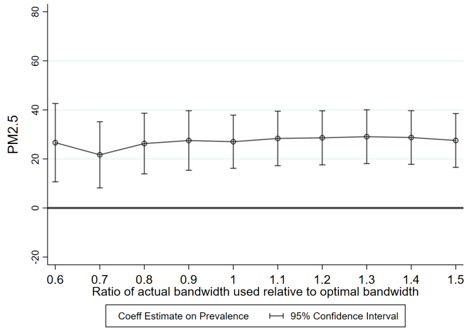
(C) Precipitation (10^3 mm)



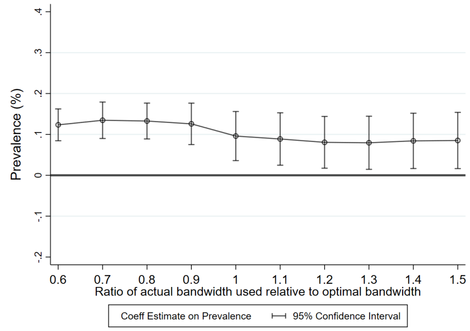
(D) Water pollution (10^6 ton)

FIGURE A6
RD Estimates of North-South Gaps in Meteorological Factors and Water Pollution

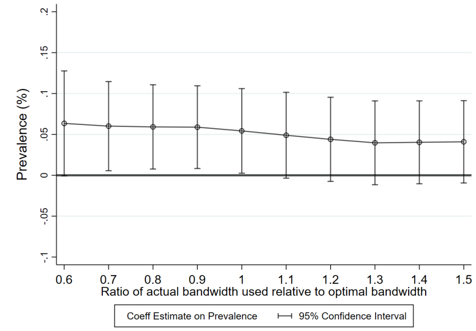
Notes: This table reports the estimated north-south gap in meteorological factors and water pollution. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on nutrition intakes are from CHNS 2011 survey. Data on industrial emissions are obtained from China Urban Statistical Yearbook. Data on meteorological characteristics are from the Chinese daily surface meteorological dataset (V3.0) provided by the National Meteorological Science Data Center. Data on water pollution are annual average from 1990 to 2011, while the meteorological variables are annual average from 1980 to 2010. Each dot represents the coefficient from a separate RD regression based on local linear distance function, epanechnikov kernel, optimal minimal standard error (MSE) bandwidth and a set of city-level control variables. Each vertical line refers to confidence interval (CI) of the RD estimates.



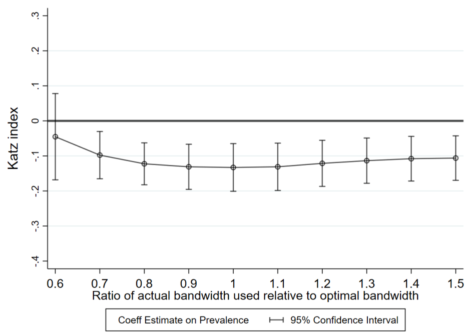
(A) **PM25**



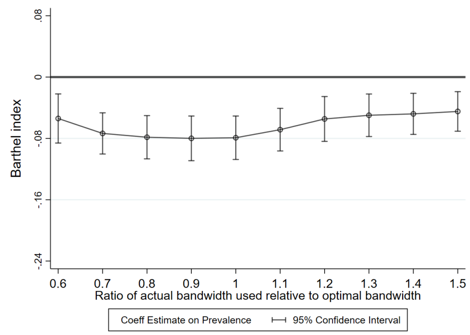
(B) **Multimorbidity ($\geq$ 2 conditions)**



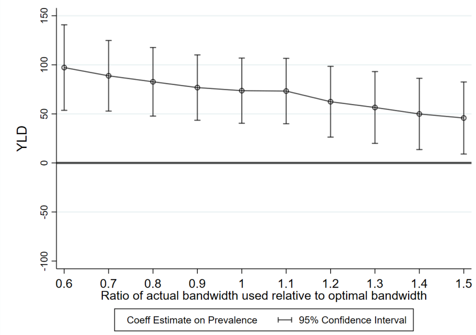
(C) **Multimorbidity ($\geq$ 3 conditions)**



(D) **Katz index**



(E) **Barthel index**



(F) **Years lived with disability (YLD)**

FIGURE A7
Robustness Tests on Alternative Bandwidths

Notes: This table reports the estimated north-south gap in long-term PM2.5 concentrations and population health outcomes using different multiples of the MSE-optimal bandwidth. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. Each dot represents the coefficient from a separate RD regression based on local linear distance function, epanechnikov kernel and and a set of city-level control variables. Each vertical line refers to confidence interval (CI) of the RD estimates.

Appendix E Tables

TABLE B1 The profile of chronic conditions in CHARLS

Chronic condition	Subcategory of specific diseases
Hypertension	Hypertension
Dyslipidemia	Elevation of low density lipoprotein, triglycerides (TGs),and total cholesterol, or a low high density lipoprotein level
Diabetes	Diabetes or high blood sugar
Asthma	Asthma
Lung diseases	Chronic lung diseases, such as chronic bronchitis (excluding asthma, tumor or cancer)
Liver diseases	Hepatic diseases (except fatty liver, tumors, and cancer)
Heart diseases	Heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems
Kidney diseases	Chronic nephropathy (except for tumor or cancer)
Stomach diseases	Gastropathy (except for tumor or cancer)

TABLE B2 Questions on Activities of Daily Living (ADL) in CHARLS

Number	Questions
1	Do you have any difficulty with running or jogging about 1 kilometer?
2	Do you have any difficulty with walking about 1 kilometer?
3	Do you have any difficulty with walking about 100 meters?
4	Do you have difficulty with climbing several flights of stairs without resting?
5	Because of health and memory problems, do you have any difficulty with bathing or showering?
6	Because of health and memory problems, do you have any difficulty with dressing? Dressing includes taking clothes out from a closet, putting them on, buttoning up, and fastening a belt.
7	Because of health and memory problems, do you have any difficulties with using the toilet, including getting up and down?
8	Do you have any difficulty with getting into or out of bed?
9	Because of health and memory problems, do you have any difficulties with controlling urination and defecation? If you use a catheter (conduit) or a pouch by yourself, then you are not considered to have difficulties.
10	Because of health and memory problems, do you have any difficulty with eating, such as cutting up your food?
Answers	
A	No, I do not have any difficulty.
B	I have difficulty but can still do it.
C	Yes, I have difficulty and need help.
D	I can not do it.

TABLE B3 Scale of Katz index

Katz index activity	Independent	Partially dependent	Fully dependent
Bathing (sponge, shower, or bathtub)	Receives no assistance	Receives assistance in bathing only one part of the body (such as back or a leg)	Receives assistance in bathing more than one part of the body (or no bathed)
Dressing (gets clothes from closets and drawers and gets dressed)	Receives no assistance	Receives assistance only in tying shoes	Receives assistance in getting clothes or in getting dressed, or stays partly or completely undressed
Going to toilet (for bowel and urine elimination; cleaning self after elimination, and arranging clothes)	Receives no assistance	Receives assistance in going to ”toilet room” or in cleaning self or in arranging clothes	Does not go to the room termed “toilet” for the elimination process
Transfer (moves in and out of bed or chair)	Receives no assistance	Moves in and out of bed or chair with assistance	Does not get out of bed
Continence (physiological process of elimination from bladder and bowel)	Controls urination and bowel movement completely by self	Has occasional ”accidents”	Supervision helps keep urine or bowel control; catheter is used or is incontinence
Feeding (process of getting food from a plate or equivalent into the mouth)	Receives no assistance	Feeds self except for getting assistance in cutting meat or buttering bread	Receives assistance in feeding or is fed partly or completely by using tubes or intravenous fluids

Scoring and definitions

A. Independent in feeding, continence, transferring, going to toilet, dressing, and bathing.
B. Independent in all but one of these functions.
C. Independent in all but bathing, and one additional function.
D. Independent in all but bathing, dressing, and one additional function.
E. Independent in all but bathing, dressing, going to toilet, and one additional function.
F. Independent in all but bathing, dressing, going to toilet, transferring, and one additionalfunction.
G. Dependent in all six functions
Other: dependent in At least two functions, but not classifiable as C, D, E, or F.

TABLE B4 Scale of Barthel index

Barthel Index Activity	Score
Feeding	0 = unable 5 = needs help cutting, spreading butter, etc., or requires modified diet 10 = independent
Bathing	0 = dependent 5 = independent (or in shower)
Grooming	0 = needs to help with personal care 5 = in dependent face/hair/teeth/shaving (implements provided)
Dressing	0 = dependent 5 = needs help but can do about half unaided 10 = in dependent (in eluding buttons, zips, laces, etc.)
Bowels	0 = incontinent (or needs to be given enemas) 5 = occasional accident 10 = continent
Bladder	0 = incontinent, or catheterized and unable to manage alone 5 = occasional accident 10 = continent
Toilet Use	0 = dependent 5 = needs some help, but can do something alone 10 = independent (on and off, dressing, wiping)
Transfers (Bed-Chair)	0 = unable, no sitting balance 5 = major help (one or two people, physical), can sit 10 = minor help (verbal or physical) 15 = independent
Mobility (on Level Surfaces)	0 = immobile or <50 yards 5 = wheelchair independent, including corners, >50 yards 10 = walks with help of one person (verbal or physical) >50 yards 15 = independent (but may use any aid; for example, stick) >50 yards
Stairs	0 = unable 5 = needs help (verbal, physical, carrying aid) 10 = independent

TABLE B5 The disability weights of chronic conditions in CHARLS

Name of chronic conditions	Dsiability weight
Hypertension	0.072
Dyslipidemia	0
Diabetes	0.148
Asthma	0.061
Lung diseases	0.344
Liver diseases	0.220
Heart diseases	0.112
Kidney diseases	0.347
Stomach diseases	0.172

Notes: This table reports the disability weights for all chronic conditions reported by the CHARLS survey. The original disability weight for each health condition is obtained from the WHO global burden of disease study (Salomon et al., 2015).

TABLE B6 Summary Statistics of Covariates

Dependent variables	(1) North	(2) South	(3) Raw Diff (1)-(2)
Economic Conditions			
GDP per capita (10^3 yuan, 1990 RMB)	7.192 (4.309)	7.141 (6.870)	0.051 (1.058)
Share of value added of primary industry	0.203 (0.097)	0.216 (0.092)	-0.013 (0.017)
Share of value added of secondary industry	0.447 (0.097)	0.433 (0.077)	0.014 (0.016)
Demographics			
Share of non-farm population	0.323 (0.144)	0.262 (0.134)	0.061 (0.025)
Average years of education	9.094 (0.849)	8.659 (0.905)	0.435 (0.160)
Weather Conditions			
Perticipation (10^3 mm)	0.597 (0.181)	1.374 (0.313)	-0.777 (0.047)
Sunlight (10^3 hour)	2.402 (0.303)	1.677 (0.336)	0.725 (0.058)
Temperature (degrees Celsius)	10.329 (3.823)	17.625 (2.351)	-7.297 (0.570)
Other Polluting Activities			
Water pollution (10^6 ton)	71.958 (64.716)	90.593 (157.929)	-18.635 (22.435)
Observations	57	64	121

Notes: This table presents the average values of relevant control variables in the north and south of the reference line. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line, of which 57 are in the north and 64 are in the south. Data on socioeconomic characteristics are obtained from China Urban Statistical Yearbook. Data on meteorological characteristics are from the Chinese daily surface meteorological dataset (V3.0) provided by the National Meteorological Science Data Center. The socioeconomic variables are all annual average from 1990 to 2011, while the meteorological variables are annual average from 1980 to 2010. Standard deviations are reported in parentheses.

TABLE B7 RD Estimated North-south Gap in each chronic condition

Chronic Disease	(1)	(2)	Chronic Disease	(3)	(4)
Hypertension	0.022	0.017	**Asthma**	0.034**	0.027*
	(0.028)	(0.026)		(0.017)	(0.016)
Bandwidth (North)	3.811	4.669	Bandwidth (North)	4.034	4.341
Bandwidth (South)	3.005	2.917	Bandwidth (South)	2.459	3.336
Diabetes	0.024	0.039**	**Stomach diseases**	0.205***	0.177***
	(0.017)	(0.017)		(0.041)	(0.040)
Bandwidth (North)	2.622	3.705	Bandwidth (North)	2.269	2.679
Bandwidth (South)	1.756	1.796	Bandwidth (South)	2.162	2.300
Dyslipidemia	0.047	0.049*	**Liver diseases**	0.026**	0.020**
	(0.030)	(0.028)		(0.012)	(0.010)
Bandwidth (North)	3.100	3.762	Bandwidth (North)	2.306	2.682
Bandwidth (South)	2.822	2.914	Bandwidth (South)	3.012	2.760
Heart diseases	0.016	0.013	**Kidney diseases**	0.066***	0.068***
	(0.016)	(0.014)		(0.023)	(0.023)
Bandwidth (North)	1.984	2.617	Bandwidth (North)	2.035	2.227
Bandwidth (South)	2.973	2.960	Bandwidth (South)	2.318	2.391
Lung diseases	0.053***	0.048***			
	(0.013)	(0.011)			
Bandwidth (North)	1.940	2.408			
Bandwidth (South)	2.461	2.575			
Observations	121	121		121	121
Longitude-region FEs	Yes	Yes		Yes	Yes
Control Variables	Yes	Yes		Yes	Yes
Bandwidth selection	MSE	MSE		MSE	MSE
Kernel Function	Epa	Tri		Epa	Tri

Notes: This table reports the estimated north-south gap in each chronic condition. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on chronic conditions are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B8 RD Estimated North-south Gap in Life-style Adjustments

Outcomes	(1)	(2)
Frequency of exercise (hrs/week)	9.740**	8.605**
	(4.110)	(3.719)
Bandwidth (North)	3.412	4.249
Bandwidth (South)	2.420	2.687
Years of alcohol drinking	-4.548***	-4.709***
	(1.674)	(1.632)
Bandwidth (North)	2.393	4.170
Bandwidth (South)	2.534	2.678
Years of smoking	-5.009***	-5.086***
	(0.874)	(0.845)
Bandwidth (North)	3.553	4.132
Bandwidth (South)	2.113	2.099
Observations	121	121
Longitude-region FEs	Yes	Yes
Control Variables	Yes	Yes
Bandwidth Selection	MSE	MSE
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in Life-style Adjustments. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on Life-style Adjustments are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B9 RD Estimated North-south Gap in nutrition intakes

Outcomes	(1)	(2)
Calorie (kcal, 3 days's average)	-357.908	-402.722
	(258.029)	(275.525)
Carbohydrate (g, 3 days's average)	-51.144	-59.599
	(49.885)	(53.534)
Fat (g, 3 days's average)	-9.708	-10.222
	(7.289)	(7.313)
Protein (g, 3 days's average)	-10.589	-10.877
	(7.855)	(8.262)
Bandwidth (North)	7.000	7.000
Bandwidth (South)	7.000	7.000
Observations	45	45
Longitude-region FEs	Yes	Yes
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in nutrition intakes. The estimation sample includes 45 cities located within 13 latitudes along the Qinling-Huaihe line. Data on nutrition intakes are from CHNS 2011 survey. The table reports RD estimated coefficients from Equations 2 based on local linear distance function and epanechnikov kernel (Column 1) or triangular kernel (Column 2).The bandwidths used in each regression are set to 7 to ensure there are enough observations to perform calculations. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B10 RD Estimated North-south Gap in Socioeconomic and Meteorological Factors

Outcomes	(1)	(2)
Panel A: Socioeconomic characteristics		
Hospital beds per 10^3 people	0.104	0.144
	(0.588)	(0.560)
Bandwidth (North)	3.619	4.178
Bandwidth (South)	2.789	2.942
Education expenditure per person (yuan)	-47.635	-48.194
	(61.926)	(60.998)
Bandwidth (North)	3.264	2.786
Bandwidth (South)	2.458	2.567
Share of non-farm population	0.024	0.018
	(0.040)	(0.039)
Bandwidth (North)	2.338	3.021
Bandwidth (South)	2.947	3.017
Share of male population	0.032	0.036
	(0.023)	(0.024)
Bandwidth (North)	3.100	3.862
Bandwidth (South)	2.764	2.888
Share finished primary school	-0.220	-0.222
	(0.177)	(0.175)
Bandwidth (North)	2.431	2.652
Bandwidth (South)	1.922	2.081
Share of married population	-0.009	-0.011
	(0.018)	(0.016)
Bandwidth (North)	2.194	2.845
Bandwidth (South)	2.204	2.303
Panel B: Meteorological conditions and water pollution		
Temperature (degrees Celsius)	-0.133	-0.061
	(0.486)	(0.458)
Bandwidth (North)	2.506	2.988
Bandwidth (South)	2.581	2.704
Sunlight (10^3 hours)	-0.115	-0.133*
	(0.079)	(0.074)
Bandwidth (North)	2.978	3.745
Bandwidth (South)	2.330	2.520
Precipitation (10^3 mm)	0.034	0.030
	(0.114)	(0.113)
Bandwidth (North)	3.481	3.813
Bandwidth (South)	2.506	2.635
Water pollution (10^6 ton)	-10.534	-13.263
	(35.068)	(33.845)
Bandwidth (North)	2.392	2.845
Bandwidth (South)	2.454	2.699
Observations	121	121

Notes: This table reports the estimated north-south gap in critical health-related factors. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on socioeconomic characteristics and industrial emissions are obtained from China Urban Statistical Yearbook. Data on demographic characteristics are from CHARLS 2011 survey. Data on meteorological characteristics are from the Chinese daily surface meteorological dataset (V3.0) provided by the National Meteorological Science Data Center. Both of socioeconomic characteristics and industrial emissions are annual average from 1990 to 2011, while the meteorological variables are annual average from 1980 to 2010. The table reports RD estimated coefficients from Equations 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2) and optimal minimal standard error (MSE) bandwidth. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B11 RD Estimated North-south Gap in PM2.5 Levels, Multimorbidity, Disability, and YLD using the CER-optimal bandwidth selection

Outcomes	(1)	(2)
Panel A: Long-term PM2.5 concentration (1980-2010)		
PM25 ($\mu g/m^3$)	25.336***	25.956***
	(6.341)	(5.939)
Bandwidth (North)	1.684	1.894
Bandwidth (South)	1.988	1.973
Panel B: Multimorbidity		
Multimorbidity ($\geq$ 2 conditions)	0.134***	0.135***
	(0.023)	(0.021)
Bandwidth (North)	2.015	2.618
Bandwidth (South)	2.583	2.529
Multimorbidity ($\geq$ 3 conditions)	0.059**	0.057**
	(0.026)	(0.023)
Bandwidth (North)	2.082	3.256
Bandwidth (South)	2.478	2.342
Panel C: Disability		
Katz index	-0.121***	-0.125***
	(0.031)	(0.029)
Bandwidth (North)	3.248	3.868
Bandwidth (South)	1.815	1.770
Barthel index	-0.078***	-0.076***
	(0.014)	(0.014)
Bandwidth (North)	3.287	3.532
Bandwidth (South)	1.893	1.822
Panel D: Quality of life		
YLD (per 10^3 people per year)	84.866***	81.361***
	(18.308)	(15.667)
Bandwidth (North)	1.638	1.783
Bandwidth (South)	2.191	2.181
Observations	121	121
Longitude-region FEs	Yes	Yes
Control Variables	Yes	Yes
Bandwidth Selector	CERtwo	CERtwo
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in long-term PM2.5 concentrations and population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 1 and 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal coverage error rate (CER) bandwidth, and a set of city-level control variables. The optimal CER bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B12 IV Estimated Marginal Effect of PM2.5 on Multimorbidity, Disability, and YLD using the CER-optimal bandwidth selection

	(1)	(2)	(3)	(4)
Panel A: Marginal effects on multimorbidity				
Outcomes:	Multimorbidity ($\geq$ 2 conditions)		Multimorbidity ($\geq$ 3 conditions)	
PM2.5 ($\mu g/m^3$)	0.051**	0.051***	0.019**	0.022**
	(0.021)	(0.019)	(0.010)	(0.009)
Bandwidth (North)	1.732	1.964	1.759	2.011
Bandwidth (South)	1.896	1.878	3.186	2.970
Panel B: Marginal effects on disability				
Outcomes:	Katz Index		Barthel Index	
PM2.5 ($\mu g/m^3$)	-0.047*	-0.047**	-0.021*	-0.019**
	(0.025)	(0.019)	(0.011)	(0.009)
Bandwidth (North)	1.678	2.049	1.492	1.688
Bandwidth (South)	2.671	2.581	2.316	2.435
Panel C: Marginal effects on quality of life				
Outcomes:	YLD			
PM2.5 ($\mu g/m^3$)	31.526***	30.449***		
	(10.300)	(8.618)		
Bandwidth (North)	1.567	1.770		
Bandwidth (South)	2.779	2.303		
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth selection	CERtwo	CERtwo	CERtwo	CERtwo
Kernel Function	Epa	Tri	Epa	Tri

Notes: This table reports the estimated marginal effect of ten-unit increase ($\mu g/m^3$) in long-term PM2.5 concentrations on population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and that on health outcomes are from CHARLS 2011 survey. We estimate the impact of PM2.5 on the listed health outcomes using local linear regression, treating distance from the Qinling-Huaihe line as the forcing variable and PM2.5 as the treatment variable, with the Qinling-Huaihe line representing a "fuzzy" discontinuity in the level of long-term PM2.5 exposure. We also adopt local linear distance function, epanechnikov kernel (Columns 1 and 3) or triangular kernel (Columns 2 and 4), optimal coverage error rate (CER) bandwidth, and a set of city-level control variables. The optimal CER bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B13 RD Estimated North-south Gap in PM2.5 Levels, Multimorbidity, Disability, and YLD using the local quardratic polynomial regression

Outcomes	(1)	(2)
Panel A: Long-term PM2.5 concentration (1980-2010)		
PM2.5 ($\mu g/m^3$)	26.673***	24.283***
	(8.832)	(8.601)
Bandwidth (North)	3.219	3.876
Bandwidth (South)	3.156	3.236
Panel B: Multimorbidity		
Multimorbidity ($\geq$ 2 conditions)	0.124***	0.131***
	(0.047)	(0.043)
Bandwidth (North)	2.822	3.359
Bandwidth (South)	3.003	3.047
Multimorbidity ($\geq$ 3 conditions)	0.074**	0.073**
	(0.038)	(0.034)
Bandwidth (North)	3.181	3.785
Bandwidth (South)	3.429	3.498
Panel C: Disability		
Katz index	-0.131***	-0.127***
	(0.041)	(0.037)
Bandwidth (North)	4.723	6.227
Bandwidth (South)	2.656	2.885
Barthel index	-0.068***	-0.073***
	(0.021)	(0.023)
Bandwidth (North)	3.093	3.841
Bandwidth (South)	2.531	2.807
Panel D: Quality of life		
YLD (per 10^3 person per year)	74.979**	74.566***
	(30.855)	(28.625)
Bandwidth (North)	2.903	3.600
Bandwidth (South)	3.553	3.666
Observations	121	121
Longitude-region FEs	Yes	Yes
Control Variables	Yes	Yes
Bandwidth Selector	MSE	MSE
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in long-term PM2.5 concentrations and population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 1 and 2 based on local quardratic distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B14 IV Estimated Marginal Effect of PM2.5 on Multimorbidity, Disability, and YLD using the local quardratic polynomial regression

	(1)	(2)	(3)	(4)
Panel A: Marginal effects on multimorbidity				
Outcomes:	Multimorbidity ($\geq$ 2 conditions)		Multimorbidity ($\geq$ 3 conditions)	
PM2.5 ($\mu g/m^3$)	0.039*	0.054*	0.016	0.026
	(0.024)	(0.030)	(0.012)	(0.017)
Bandwidth (North)	2.901	3.620	3.644	4.022
Bandwidth (South)	3.982	3.509	4.853	4.040
Panel B: Marginal effects on disability				
Outcomes:	Katz Index		Barthel Index	
PM2.5 ($\mu g/m^3$)	-0.034	-0.048*	-0.024	-0.018
	(0.024)	(0.028)	(0.016)	(0.011)
Bandwidth (North)	3.438	4.046	2.987	3.690
Bandwidth (South)	2.913	3.142	3.310	4.232
Panel C: Marginal effects on quality of life				
Outcomes:	YLD			
PM2.5 ($\mu g/m^3$)	28.058	28.348		
	(17.794)	(17.440)		
Bandwidth (North)	3.671	3.932		
Bandwidth (South)	3.626	3.758		
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth selection	MSE	MSE	MSE	MSE
Kernel Function	Epa	Tri	Epa	Tri

Notes: This table reports the estimated marginal effect of ten-unit increase ($\mu g/m^3$) in long-term PM2.5 concentrations on population health outcomes. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and that on health outcomes are from CHARLS 2011 survey. We estimate the impact of PM2.5 on the listed health outcomes using local linear regression, treating distance from the Qinling-Huaihe line as the forcing variable and PM2.5 as the treatment variable, with the Qinling-Huaihe line representing a "fuzzy" discontinuity in the level of long-term PM2.5 exposure. We also adopt local quardratic distance function, epanechnikov kernel (Columns 1 and 3) or triangular kernel (Columns 2 and 4), optimal minimal standard error (MSE) bandwidth, and control for a set of city-level variables in regression analysis. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B15 RD Estimated North-south Gap in PM2.5 Levels, Multimorbidity, Disability, and YLD based on bias correction and robust variance adjustment

Outcomes	(1)	(2)
Panel A: Long-term PM2.5 concentration (1980-2010)		
PM2.5 ($\mu g/m^3$)	25.596***	26.368***
	(6.907)	(7.589)
Bandwidth (North)	3.985	3.676
Bandwidth (South)	4.021	4.135
Panel B: Multimorbidity		
Multimorbidity ($\geq$ 2 conditions)	0.127***	0.129***
	(0.033)	(0.039)
Bandwidth (North)	6.231	4.382
Bandwidth (South)	5.505	5.975
Multimorbidity ($\geq$ 3 conditions)	0.046*	0.061*
	(0.026)	(0.031)
Bandwidth (North)	7.162	4.499
Bandwidth (South)	4.787	5.008
Panel C: Disability		
Katz index	-0.162***	-0.170***
	(0.038)	(0.043)
Bandwidth (North)	7.540	6.959
Bandwidth (South)	3.377	3.322
Barthel index	-0.095***	-0.096***
	(0.017)	(0.018)
Bandwidth (North)	6.108	6.518
Bandwidth (South)	3.854	3.974
Panel D: Quality of life		
YLD (per 10^3 people per year)	95.287***	98.264***
	(21.283)	(22.675)
Bandwidth (North)	3.674	3.459
Bandwidth (South)	4.813	5.019
Observations	121	121
Longitude-region FEs	Yes	Yes
Control Variables	Yes	Yes
Bandwidth selection	MSE	MSE
Kernel Function	Epa	Tri

Notes: This table reports the estimated north-south gap in long-term PM2.5 concentrations and population health outcomes based on bias correction and robust variance adjustment. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and data on health outcomes are from CHARLS 2011 survey. The table reports RD estimated coefficients from Equations 1 and 2 based on local linear distance function, epanechnikov kernel (Column 1) or triangular kernel (Column 2), optimal minimal standard error (MSE) bandwidth, and a set of city-level control variables. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.

TABLE B16 IV Estimated Marginal Effect of PM2.5 on Multimorbidity, Disability, and YLD based on bias correction and robust variance adjustment

	(1)	(2)	(3)	(4)
Panel A: Marginal effects on multimorbidity				
Outcomes:	Multimorbidity ($\geq$ 2 conditions)		Multimorbidity ($\geq$ 3 conditions)	
PM2.5 ($\mu g/m^3$)	0.060***	0.060***	0.012	0.017
	(0.022)	(0.020)	(0.013)	(0.012)
Bandwidth (North)	3.418	3.612	3.310	3.682
Bandwidth (South)	4.007	3.982	6.434	5.879
Panel B: Marginal effects on disability				
Outcomes:	Katz Index		Barthel Index	
PM2.5 ($\mu g/m^3$)	-0.034	-0.038**	-0.008	-0.011
	(0.021)	(0.019)	(0.011)	(0.011)
Bandwidth (North)	3.514	4.113	3.256	3.647
Bandwidth (South)	5.565	5.407	4.988	5.169
Panel C: Marginal effects on quality of life				
Outcomes:	YLD			
PM2.5 ($\mu g/m^3$)	20.762*	31.520***		
	(11.065)	(11.196)		
Bandwidth (North)	3.356	3.786		
Bandwidth (South)	5.652	4.839		
Observations	121	121	121	121
Longitude-region FEs	Yes	Yes	Yes	Yes
Control Variables	Yes	Yes	Yes	Yes
Bandwidth selection	MSE	MSE	MSE	MSE
Kernel Function	Epa	Tri	Epa	Tri

Notes: This table reports the estimated marginal effect of ten-unit increase ($\mu g/m^3$) in long-term PM2.5 concentrations on population health outcomes based on bias correction and robust variance adjustment. The estimation sample includes 121 cities located within 16 latitudes along the Qinling-Huaihe line. Data on long-term PM2.5 concentrations are obtained from Zhong et al. (2022) and that on health outcomes are from CHARLS 2011 survey. We estimate the impact of PM2.5 on the listed health outcomes using local linear regression, treating distance from the Qinling-Huaihe line as the forcing variable and PM2.5 as the treatment variable, with the Qinling-Huaihe line representing a "fuzzy" discontinuity in the level of long-term PM2.5 exposure. We also adopt local linear distance function epanechnikov kernel (Columns 1 and 3) or triangular kernel (Columns 2 and 4), optimal minimal standard error (MSE) bandwidth, and control for a set of city-level variables in regression analysis. The optimal MSE bandwidths are chosen separately for each side of the RD cutoff and in each regression. Robust standard errors clustered at the city level are reported in parentheses. *10% significance level; **5% significance level; ***1% significance level.