Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy

Yuyu Chen[a,1], Avraham Ebenstein[b,1], Michael Greenstone[c,d,1,2], and Hongbin Li[e,1]

[a]Applied Economics Department, Guanghua School of Management, Peking University, Beijing 100871, China; [b]Department of Economics, Hebrew University of Jerusalem, Mount Scopus 91905, Israel; [c]Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02142; [d]National Bureau of Economic Research, Cambridge, MA 02138; and [e]China Data Center and Department of Economics, School of Economics and Management, Tsinghua University, Beijing 100084, China

Edited by William C. Clark, Harvard University, Cambridge, MA, and approved May 28, 2013 (received for review January 2, 2013)

This paper's findings suggest that an arbitrary Chinese policy that greatly increases total suspended particulates (TSPs) air pollution is causing the 500 million residents of Northern China to lose more than 2.5 billion life years of life expectancy. The quasi-experimental empirical approach is based on China's Huai River policy, which provided free winter heating via the provision of coal for boilers in cities north of the Huai River but denied heat to the south. Using a regression discontinuity design based on distance from the Huai River, we find that ambient concentrations of TSPs are about 184 μg/m^3 [95% confidence interval (CI): 61, 307] or 55% higher in the north. Further, the results indicate that life expectancies are about 5.5 y (95% CI: 0.8, 10.2) lower in the north owing to an increased incidence of cardiorespiratory mortality. More generally, the analysis suggests that long-term exposure to an additional 100 μg/m^3 of TSPs is associated with a reduction in life expectancy at birth of about 3.0 y (95% CI: 0.4, 5.6).

airborne particulate matter | unintended consequences of policy | premature mortality | health costs of coal combustion | Chinese environmental quality

Air quality in China is notoriously poor and recently has become an issue associated with increasing social unrest. Ambient concentrations of total suspended particulates (TSPs) between 1981–2001 were more than double China's National Annual Mean Ambient Air Quality Standard of 200 μg/m^3 (1) and five times the level that prevailed in the United States before the passage of the Clean Air Act in 1970. Furthermore, air quality is especially poor in Northern China, which is home to several of the world's most polluted cities (2). Following a career in the Southern China city of Shanghai, Premier Zhu Rongi reportedly quipped in 1999: "If I work in your Beijing [in Northern China], I would shorten my life at least five years" (3).

This paper examines the health consequences of these extraordinary pollution levels by exploiting a seemingly arbitrary Chinese policy that produced dramatic differences in air quality within China. During the 1950–1980 period of central planning, the Chinese government established free winter heating of homes and offices via the provision of free coal for fuel boilers as a basic right. The combustion of coal in boilers is associated with the release of air pollutants, and in particular emission of particulate matter that can be extremely harmful to human health (4, 5). Due to budgetary limitations, however, this right was only extended to areas located in North China, which is defined by the line formed by the Huai River and Qinling Mountain range (Fig. 1). Even today, the long-lived heating systems continue to make indoor heating much more common in the north.

This paper finds that the Huai River policy had dramatic impacts on pollution and human health. To the north of the Huai River, particulate concentrations are 184 μg/m^3 [95% confidence interval (CI): 61, 307], or 55% higher, and life expectancies are 5.5 y (95% CI: 0.8, 10.2) lower, almost entirely due to an increased incidence of cardiorespiratory mortality (and no effect on other causes). The estimates suggest that the 500 million residents of Northern China during the 1990s experienced a loss of more than 2.5 billion life years owing to the Huai River policy.

Furthermore, a research design based on this policy allows for a unique opportunity to estimate the effect of TSPs on human health, which can be applied to other countries, time periods, and settings. The resulting estimates suggest that long-term exposure to an additional 100 μg/m^3 of TSPs is associated with a reduction in life expectancy at birth of about 3.0 y (95% CI: 0.4, 5.6). This estimate is more than five times larger than the estimated impact of TSPs on life expectancy from fitting a conventional ordinary least-squares equation on the same data.

The study addresses several shortcomings in our understanding about the health effects of air pollution. First, the research design provides estimates of the impact of long-run exposure to TSPs on life expectancy. The policy caused long-run differences in TSP concentrations between cities north and south of the Huai River (6). Moreover, during the period in question, the *hukou* (household registrations) system restricted mobility, so in general individuals will be observed where they lived most of their lives. In contrast, studies that use data for the United States or other developed countries must assume no migration, which is undermined by the high rates of migration in the United States and the potential selection of location based on air pollution concentrations, or alternatively assume that short-run variation in air pollution is informative about life expectancy (7, 8). Second, the availability of a regression discontinuity design based on the Huai River policy provides an appealing quasi-experimental approach that can help to move the existing literature from documenting a robust association between particulates and health toward documenting a causal relationship (9). Third, China's air is extremely polluted and we are unaware of any previous direct evidence on the impact of air pollution on life expectancy at these concentrations, although important research has applied results from the United States to the Chinese setting (10). Fourth, the analysis is conducted with the most comprehensive data file ever compiled on mortality and air pollution in China or any other developing country.

Data Sources

The dataset for this analysis is based on several sources. We collected information on annual daily average concentrations of

Author contributions: Y.C., A.E., M.G., and H.L. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

See Commentary on page 12861.

[1]Y.C., A.E., M.G., and H.L. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: mgreenst@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300018110/-/DCSupplemental.

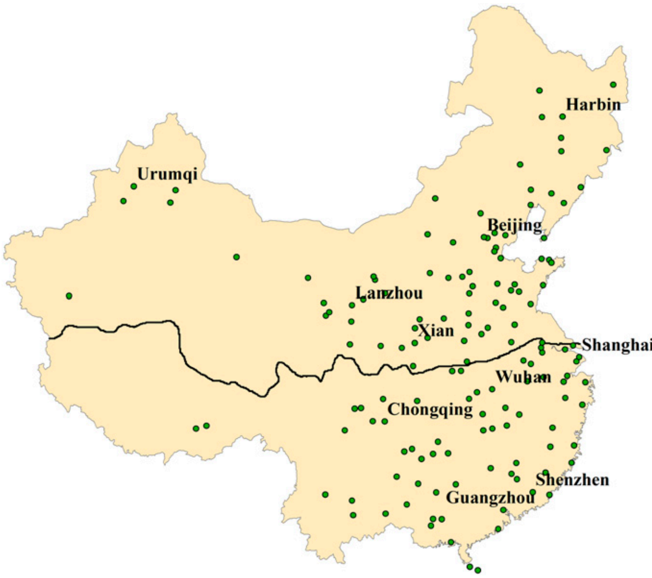


Fig. 1. The cities shown are the locations of the Disease Surveillance Points. Cities north of the solid line were covered by the home heating policy.

TSPs for 90 cities from 1981 to 2000. These data were compiled through a combination of hand entry from Chinese-language publications and access to electronic files (11). We obtained the data by combining the results of a World Bank project with information from hard copies of *China Environment Yearbook* to generate a single comprehensive file of air pollution across Chinese cities for our period.

Although there is recent evidence that air pollution readings in Chinese cities are manipulated by policymakers, with a tendency for officials to underreport pollution, we believe that manipulation was not a serious issue during the period we study (12). First, for the period of our study, government officials' evaluations were primarily based on economic growth rather than environmental indices (13). Second, the readings were also generally not widely available, which reduced the incentive to publish inaccurate information. Indeed, it has been reported extensively that Chinese officials monitored air pollution concentrations beginning in the late 1970s but this information was not publicly released until 1998 (14). Also, our analysis relies on testing for differences in air pollution readings near the Huai River, so unless the data were manipulated differently north and south of the river, mismeasurement would not bias the results in an obvious fashion. Moreover, even in the presence of manipulation of pollution concentrations, the estimates of the health impacts of living north of the Huai River will not be affected.

The mortality data are derived from China's Disease Surveillance Points (DSPs) system (15). The DSP is a set of 145 sites chosen to be nationally representative (benchmarked against the 1990 China census) so that it captures China's variation in wealth, urbanicity, and geographic dispersion. The DSP records all deaths and population counts at the sites and yields a nationally representative annual sample of deaths (16). The analysis will rely on the data taken from roughly 500,000 deaths recorded at sites between 1991 and 2000, and population counts by age and sex that are used to convert the recorded deaths into city-level mortality rates for ages 1, 5, and 10 y and 5-y increments through age 80. Additionally, these mortality rates are used to calculate an overall mortality rate based on China's age distribution in 2000 and life expectancy at birth, both measured annually for the 125 locations. In *SI Appendix*, we discuss why the results are unlikely to be driven by mismeasurement in a location's mortality rate from either migration or the transfer of sick rural residents to urban hospitals.

Importantly, the cause of death is also recorded after multiple validation checks. We classify causes of death as either cardiorespiratory or noncardiorespiratory. The cardiorespiratory causes of death that are those that have been linked to ambient air quality and include heart disease (17), stroke (18–20), lung cancer (21), and respiratory illnesses (22, 23). Causes of death presumably unrelated to air quality include other cancers, accidental or violent deaths, and various stomach ailments. Together, these two categories cover all causes of death.

We collected a range of determinants of mortality and life expectancy, besides TSPs, that are used as control variables in the subsequent statistical analysis. We obtained daily average temperature data for each location in the air pollution sample from the World Meteorological Organization that was used to calculate annual heating and cooling degree days (24). We also compiled a series of variables from China's 2000 census that are potentially related to health outcomes: average education of county residents, manufacturing's share of employment, the percentage of residents with urban registration, and the percentage of residents with access to tap water. The data file also includes an income variable taken from the DSP, which placed each site into one of four income categories.

To estimate the impact of long-run exposure to pollution, the location-level panel data are collapsed to a 125-observations, location-level, cross-sectional dataset, because the Huai River regression discontinuity design is fundamentally a cross-sectional design. This data file is obtained by averaging the annual location-specific measures of mortality rates, life expectancies, pollution concentrations, weather variables, and other covariates across the available years. Additionally, we used a geographic information system to identify the degrees latitude that each city centroid is north of the Huai River line and merged this information into the final dataset. *SI Appendix* provides more details on the procedure used to collapse the data file and the data sources.

Econometric Model

We use two approaches to estimate the relationship between TSPs and human health. The first approach is a "conventional" strategy that uses ordinary least squares to fit the following equation to the cross-sectional data file:

$$Y_j = \beta_0 + \beta_1 TSP_j + X_j\Gamma + \varepsilon_j, \qquad [1]$$

where TSP_j is the total suspended particulates concentration in city j, X_j is a vector of the observable characteristics of the city that might influence health outcomes other than air quality, and ε_j is a disturbance term. The dependent variable is Y_j, which is either a measure of mortality rates in city j or its residents' life expectancy, which is a simple function of age-specific mortality rates.

The coefficient β_1 measures the effect of TSP exposure on mortality, after controlling for the available covariates. Consistent estimation of β_1 requires that unobserved determinants of mortality do not covary with TSP_j after adjustment for X_j. Thus, the conventional approach rests on the assumption that linear adjustment for the limited set of variables available in the census removes all sources of confounding. With data from the United States, Chay et al. (9) have documented the sensitivity of the estimated TSP–adult mortality relationship to small changes in specification and sample, which is consistent with the possibility that omitted variables bias plagues the conventional approach.

The second approach leverages the regression discontinuity (RD) design implicit in the Huai River policy to measure its impact on TSP concentrations and life expectancy. The RD design was developed more than five decades ago and has been used successfully to test the causal nature of relationships in a wide range of fields including psychology, education, statistics, biostatistics, and economics (25, 26).

Table 1. Summary statistics

Variable	South (1)	North (2)	Difference in means (3)	Adjusted difference in means (4)	P value (5)
Panel 1: Air pollution exposure at China's Disease Surveillance Points					
TSPs, $\mu g/m^3$	354.7	551.6	196.8***	199.5***	<0.001/0.002
SO_2, $\mu g/m^3$	91.2	94.5	3.4	−3.1	0.812/0.903
NO_x, $\mu g/m^3$	37.9	50.2	12.3***	−4.3	<0.001/0.468
Panel 2: Climate at the Disease Surveillance Points					
Heating degree days	2,876	6,220	3,344***	482	<0.001/0.262
Cooling degree days	2,050	1,141	−910***	−183	<0.001/0.371
Panel 3: Demographic features of China's Disease Surveillance Points					
Years of education	7.23	7.57	0.34	−0.65	0.187/0.171
Share in manufacturing	0.14	0.11	−0.03	−0.15***	0.202/0.002
Share minority	0.11	0.05	−0.05	0.04	0.132/0.443
Share urban	0.42	0.42	0.00	−0.20*	0.999/0.088
Share tap water	0.50	0.51	0.02	−0.32**	0.821/0.035
Rural, poor	0.21	0.23	0.01	−0.33*	0.879/0.09
Rural, average income	0.34	0.33	0.00	0.24	0.979/0.308
Rural, high income	0.21	0.19	−0.02	0.27	0.772/0.141
Urban site	0.24	0.25	0.01	−0.19	0.859/0.241
Predicted life expectancy	74.0	75.5	1.54***	−0.24	<0.001/0.811
Actual life expectancy	74.0	75.5	1.55	−5.04**	0.158/0.044

The sample (n = 125) is restricted to DSP locations within 150 km of an air quality monitoring station. TSP ($\mu g/m^3$) in the years 1981–2000 before the DSP period is used to calculate city-specific averages. Degree days are the deviation of each day's average temperature from 65°F, averaged over the years 1981–2000 before the DSP period. The results in column (4) are adjusted for a cubic in degrees of latitude north of the Huai River boundary. Predicted life expectancy is calculated by OLS using all of the demographic and meteorological covariates shown. All results are weighted by the population at the DSP location. One DSP location is excluded due to invalid mortality data. *Significant at 10%, **significant at 5%, ***significant at 1%. Sources: China Disease Surveillance Points (1991–2000), *China Environment Yearbook* (1981–2000), and World Meteorological Association (1980–2000).

This paper's RD design exploits the discrete increase in the availability of free indoor heating as one crosses the Huai River line (with no availability to the south and, in principle, complete availability north of the line). Specifically, we separately test whether the Huai River policy caused a discontinuous change in TSPs at the river and a discontinuous change in life expectancy. The respective necessary assumptions are that any unobserved determinants of TSPs or mortality change smoothly as they cross the river. If the relevant assumption is valid, adjustment for a sufficiently flexible polynomial in distance from the river will remove all potential sources of bias and allow for causal inference.

In practice, we estimate the following equations to test for the impacts of the Huai River policy:

$$TSP_j = \alpha_0 + \alpha_1 N_j + \alpha_2 f(L_j) + X_j\kappa + \nu_j \quad \textbf{[2a]}$$

$$Y_j = \delta_0 + \delta_1 N_j + \delta_2 f(L_j) + X_j\phi + u_j, \quad \textbf{[2b]}$$

where j references a city or location in China. TSP_j is the average annual ambient concentration of TSPs in city j over the period 1980–2000 and Y_j is a measure of city j's mortality or life expectancy at birth. N_j is an indicator variable equal to 1 for locations that are north of the Huai River line, $f(L_j)$ is a polynomial in the degrees north of the Huai River, and X_j is a vector of the demographic and city characteristics, other than air quality, that are associated with mortality rates (*SI Appendix* gives details).

This design can also be used to develop estimates of the impact of TSP concentrations on life expectancy. Specifically, if the Huai River policy only influences mortality through its impact on TSPs, then it is valid to treat Eq. **2a** as the first stage in a two-stage least-squares (2SLS) system of equations. An important appeal of the 2SLS approach is that it produces estimates of the impact of units of TSPs on life expectancy, so the results are applicable in other settings. The second-stage equation is

$$Y_j = \beta_0 + \beta_1 T\hat{S}P_j + \beta_2 f(L_j) + X_j\Gamma + \varepsilon_j, \quad \textbf{[2c]}$$

where $T\hat{S}P_j$ represents the fitted values from estimating (Eq. **1**) and the other variables are as described above. The 2SLS approach offers the prospect of solving the confounding or omitted

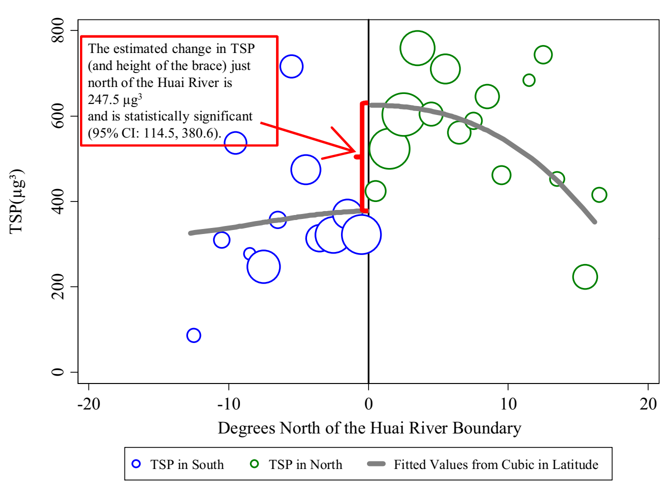


Fig. 2. Each observation (circle) is generated by averaging TSPs across the Disease Surveillance Point locations within a 1° latitude range, weighted by the population at each location. The size of the circle is in proportion to the total population at DSP locations within the 1° latitude range. The plotted line reports the fitted values from a regression of TSPs on a cubic polynomial in latitude using the sample of DSP locations, weighted by the population at each location.

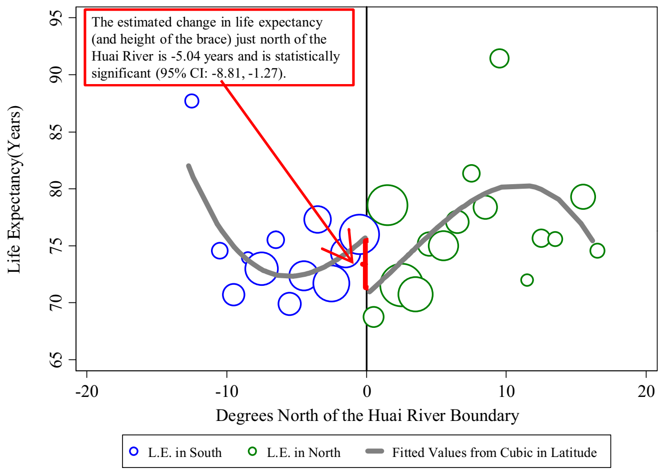


Fig. 3. The plotted line reports the fitted values from a regression of life expectancy on a cubic in latitude using the sample of DSP locations, weighted by the population at each location.

variables problem associated with the estimation of the air pollution–health effects relationship and is a solution to the attenuation bias associated with the mismeasurement of TSP.

Results

Summary Statistics. Table 1 reports the summary statistics for several of the key determinants of mortality rates and provides evidence on the validity of the RD design. Columns (1) and (2) report the means in cities south and north of the Huai River line. Column (3) reports the mean difference between the North and the South. Column (4) also reports the difference, but this time it is adjusted for a cubic polynomial in degrees north of the Huai River so that it is a test for a discontinuous change at the Huai River line. A direct test of the RD design's identifying assumption that unobservables change smoothly at the boundary is, of course, impossible, but it would nevertheless be reassuring if observable determinants change smoothly at the boundary. (This is analogous to the test in randomized control trials that observable determinants of the outcome are independent of treatment status.) Column (5) reports the P values associated with the tests that the differences in columns (3) and (4) are equal to zero.

Two key points emerge from this table. First, there are large differences in TSP exposure among Southern and Northern Chinese residents, but not for other forms of air pollution (e.g., sulfur dioxide and nitrous oxides) owing to the greater distances that they travel. Second, there are substantial differences in the determinants of mortality rates (e.g., temperature and predicted life expectancy) between the South and North, but adjustment for the cubic polynomial in latitude greatly reduces these differences and causes them to become statistically insignificant. This finding supports the validity of the RD design and the 2SLS approach to inferring the causal relationship between TSPs and life expectancy.

Graphical Analysis. The paper's primary findings are presented graphically in Figs. 2–4. Fig. 2 plots cities' TSP concentration against their degrees north of the Huai River boundary. The line is the fitted value from the estimation of the first-stage Eq. **1**, without adjustment for X_j. The circles represent the average TSP concentration across locations within 1° latitude distance bins from the Huai River; each circle's size is proportional to the number of DSP locations within the relevant 1° bin. The discontinuous change in ambient TSP concentrations to the north of the border indicates that the Huai River policy increased TSP concentrations in the North by about 200 μg/m^3. In contrast, *SI Appendix*, Figs. S1 and S2 graphically confirm that sulfur dioxide and nitrous oxide concentrations are approximately equal on the two sides of the Huai River.

Fig. 3 repeats the graphical exercise in Fig. 2, except here it plots life expectancy against degrees north of the Huai River boundary. The striking finding is that there is a discrete decline in life expectancy at the border of roughly 5 y, which mirrors the increase in TSPs. Together, Figs. 2 and 3 reveal a sharp increase in TSPs and a sharp decline in life expectancy at precisely the location where the Huai River policy went into effect. The results in Fig. 3 are also evident in *SI Appendix*, Fig. S3 in the sample of DSP sites within 5° latitude of the Huai River.

Fig. 4 graphically assesses the validity of this paper's approach by testing whether predicted life expectancy, calculated as the fitted value from an ordinary least-squares (OLS) regression of life expectancy on all covariates except TSPs (just as in Table 1), differs at the Huai River's border. Specifically, this equation includes all of the covariates listed in panels 2 and 3 of Table 1, and these variables collectively explain a substantial portion of the variation in life expectancy ($R^2 = 0.265$). It is evident that predicted life expectancy is essentially equal just to the north and south of the border (Table 1 reports a P value of 0.81 from a test of equal life expectancy). Also note that *SI Appendix*, Table S2 demonstrates that dietary and smoking patterns are similar in the North and South, suggesting that these determinants of life expectancy are unlikely to explain the sharp decline in life expectancy north of the border.

Regression Results. Table 2 reports the results from the application of the conventional OLS approach. The estimation of Eq. **1** with four different dependent variables is reported: the overall mortality rate, the cardiorespiratory mortality rate, the noncardiorespiratory-related mortality rate, and life expectancy. The set of controls are reported at the bottom of the table. The first three rows indicate that a 100 μg/m^3 increase in TSPs is associated with a statistically significant 3% increase in the mortality rate that is entirely due to an increase in cardiorespiratory causes of death. The column (2) specification in the final row indicates that a 100 μg/m^3 increase in TSPs is associated with a loss in life expectancy of about 0.52 y (95% CI: 0.07, 0.97).

Table 3 presents the RD results from the estimation of Eqs. **2a** and **2b** in panel 1 and the 2SLS results from Eq. **2c** in panel 2. In column (1), the specification includes a cubic in distance from the Huai River (measured in degrees of latitude). Column (2) adds the available covariates to the specification. Column (3) uses an alternative RD approach, which limits the sample to locations within

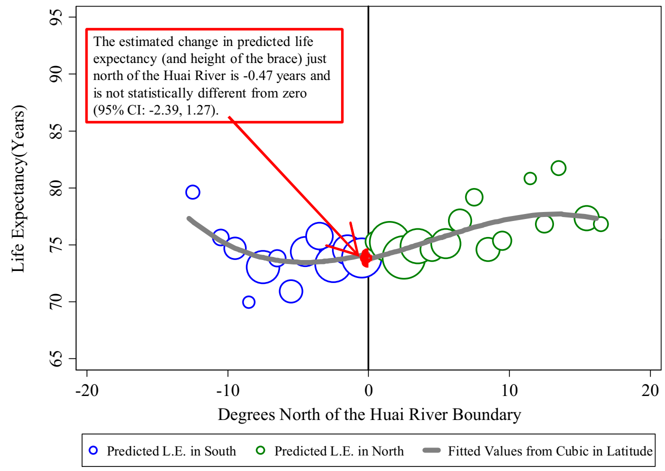


Fig. 4. The plotted line reports the fitted values from a regression of predicted life expectancy on a cubic in latitude using the sample of DSP locations, weighted by the population at each location. Predicted life expectancy is calculated by OLS using demographic and meteorological covariates (excluding TSPs).

SEE COMMENTARY

ENVIRONMENTAL SCIENCES

SUSTAINABILITY SCIENCE

Table 2. Impact of TSPs (100 μg/m^3) on health outcomes using conventional strategy (ordinary least squares)

Dependent variable	(1)	(2)
ln(All cause mortality rate)	0.03* (0.01)	0.03** (0.01)
ln(Cardiorespiratory mortality rate)	0.04** (0.02)	0.04** (0.02)
ln(Noncardiorespiratory mortality rate)	0.01 (0.02)	0.01 (0.02)
Life expectancy, y	−0.54** (0.26)	−0.52** (0.23)
Climate controls	No	Yes
Census and DSP controls	No	Yes

n = 125. Each cell in the table represents the coefficient from a separate regression, and heteroskedastic-consistent SEs are reported in parentheses. The cardiorespiratory illnesses are heart disease, stroke, lung cancer and other respiratory illnesses. The noncardiorespiratory-related illnesses are violence, cancers other than lung, and all other causes. Models in column (2) include demographic controls and climate controls reported in Table 1. Regressions are weighted by the population at the DSP location. *Significant at 10%, **significant at 5%, ***significant at 1%. Sources: China Disease Surveillance Points (1991–2000), *China Environment Yearbook* (1981–2000), and World Meteorological Association (1980–2000).

5° latitude of the Huai River and reduces the sample size from 125 to 69. With this smaller sample, latitude is modeled with a first-order polynomial. The specification details are noted at the bottom of the table.

The first row of Table 3, panel 1 confirms that the impact of the Huai River policy on TSPs is robust to adjustment for the covariates and the alternate RD approach (*SI Appendix*, Fig. S3). The remaining rows suggest that this policy increases mortality rate by 22–30% and that this is almost entirely due to higher rates of mortality among cardiorespiratory causes. The estimates in the final row indicate that there is a discontinuous decrease in life expectancy of ∼5 y to the north of the boundary.

Panel 2 in Table 3 reports on the 2SLS estimates. Column (2) estimates suggest that a 100 μg/m^3 increase in TSPs is associated with an increase in the overall mortality rate of 14% or 0.14 ln points (95% CI: 0.005, 0.275) and a decline in life expectancy of 3.0 y (95% CI: 0.39, 5.61). Again, the results seem to be driven entirely by higher mortality rates from cardiorespiratory causes of death.

The heterogeneity in the results across the population is explored in *SI Appendix*. *SI Appendix*, Table S3 reveals that the mortality and life expectancy findings hold for men and women. *SI Appendix*, Table S4 and Fig. S4 demonstrate that the impacts on cardiorespiratory mortality rates are generally evident over the entire course of the life cycle, not just for the young and old.

The basic results are also robust to a wide variety of specification checks. *SI Appendix*, Tables S5 and S6 document that the results are qualitatively unchanged by using alternative methods to assign TSP concentrations to DSP locations. *SI Appendix*, Table S7 shows that the panel 1 results are qualitatively unchanged by expanding the sample to the full set of 144 DSP locations (with valid mortality data) from the 125 locations with TSP data. Further, the results are unchanged by adjustment for distance from the coast (*SI Appendix*, Table S8).

We additionally explored the robustness of the results to alternative approaches to implementing the RD design. *SI Appendix*, Table S9 reports a set of goodness-of-fit tests that lead us to emphasize modeling distance from the Huai River with a cubic polynomial in latitude. *SI Appendix*, Table S10 documents that a first- or second-order polynomial in latitude best fits the data in the sample that is restricted to locations within 5° latitude of the river, and that the results are qualitatively similar using either approach. *SI Appendix*, Table S11 reports on specifications that allow the polynomial to differ to the north and south of the river. The goodness-of-fit statistics support separate quadratics north and south of the Huai River, and this specification suggests a somewhat smaller increase in TSPs at the border but a larger decline in life expectancy.

A natural concern related to the research design is that the government used the Huai River as the demarcation line for changes in other government policies related to public health, and this would confound the estimates of TSPs on health. This possibility is mitigated by the fact that the Huai River is not a border used for administrative purposes. The Huai River follows the January 0° average temperature line (Celsius), and this was in fact the basis of its choice as a method to divide the country for free heating. Further, local policies generally hew to administrative boundaries associated with cities and provinces; indeed, the Huai River cuts through several provinces. Nevertheless, we compiled some additional variables on city-level policies that are plausibly related to health from the *China City Statistical Yearbook*. We examine whether there is any discontinuity in these variables at the

Table 3. Using the Huai River policy to estimate the impact of TSPs (100 μg/m^3) on health outcomes

Dependent variable	(1)	(2)	(3)
Panel 1: Impact of "North" on the listed variable, ordinary least squares			
TSPs, 100 μg/m^3	2.48*** (0.65)	1.84*** (0.63)	2.17*** (0.66)
ln(All cause mortality rate)	0.22* (0.13)	0.26* (0.13)	0.30* (0.15)
ln(Cardiorespiratory mortality rate)	0.37** (0.16)	0.38** (0.16)	0.50*** (0.19)
ln(Noncardiorespiratory mortality rate)	0.00 (0.13)	0.08 (0.13)	0.00 (0.13)
Life expectancy, y	−5.04** (2.47)	−5.52** (2.39)	−5.30* (2.85)
Panel 2: Impact of TSPs on the listed variable, two-stage least squares			
ln(All cause mortality rate)	0.09* (0.05)	0.14** (0.07)	0.14* (0.08)
ln(Cardiorespiratory mortality rate)	0.15** (0.06)	0.21** (0.09)	0.23** (0.10)
ln(Noncardiorespiratory mortality rate)	0.00 (0.05)	0.04 (0.07)	0.00 (0.06)
Life expectancy, y	−2.04** (0.92)	−3.00** (1.33)	−2.44 (1.50)
Climate controls	No	Yes	Yes
Census and DSP controls	No	Yes	Yes
Polynomial in latitude	Cubic	Cubic	Linear
Only DSP locations within 5° latitude	No	No	Yes

The sample in columns (1) and (2) includes all DSP locations (n = 125) and in column (3) is restricted to DSP locations within 5° latitude of the Huai River boundary (n = 69). Each cell in the table represents the coefficient from a separate regression, and heteroskedastic-consistent SEs are reported in parentheses. Models in column (1) include a cubic in latitude. Models in column (2) additionally include demographic and climate controls reported in Table 1. Models in column (3) are estimated with a linear control for latitude. Regressions are weighted by the population at the DSP location. *Significant at 10%, **significant at 5%, ***significant at 1%. Sources: China Disease Surveillance Points (1991–2000), *China Environment Yearbook* (1981–2000), and World Meteorological Association (1980–2000).

Huai River in *SI Appendix*, Table S12. The results demonstrate that the null hypothesis that these variables are equal on both sides of the Huai River generally cannot be rejected. We also reestimate our Table 3 specification with the addition of these variables in *SI Appendix*, Table S13. We find that the addition of these variables to the Table 3 specifications leaves our results largely unchanged, and even suggests modestly larger losses of life expectancy than those reported in Table 3.

Interpretation. The 2SLS estimates suggest that a 100 $\mu g/m^3$ increase in TSPs leads to a decline in life expectancy of 3.0 y (95% CI: 0.4, 5.6). This estimate is more than five times larger than one obtained from the application of a conventional least-squares approach to the same data. The difference suggests that the OLS estimates understate the true effect owing to some combination of omitted variables bias (e.g., unobserved factors that improve health and are positively correlated with TSPs, such as income and hospital quality) and measurement error. Further, it is about half the magnitude of the estimated impact of an equal unit of particulate matter smaller than 2.5 μm ($PM_{2.5}$) from the pioneering study of Pope et al. (27). There are at least three important differences between the studies that complicate direct comparisons: (*i*) $PM_{2.5}$ is believed to be more lethal than the larger particles that also qualify as TSPs; (*ii*) the setting for the $PM_{2.5}$ results is the United States between 1980 and 2000, when particulate concentrations were just a fraction of current concentrations in China; and (*iii*) this study's results are based on a RD design, which can produce causal estimates in nonexperimental settings.

A related issue is whether the Huai River heating policy caused behavioral responses that amplify or mitigate the estimated mortality impacts of TSPs. For example, the heating policy very likely leads people in the North to spend more time indoors, where temperatures are presumably higher; this would be protective of human health. Alternatively, the greater time indoors may cause people to reduce their exercise and increase their exposure to indoor air pollution that would be harmful to health. Further, the free provision of coal is an in-kind transfer that increases households' disposable income, and this may cause northern households to alter their consumption patterns in ways that are protective (e.g., medical care) or harmful (e.g., tobacco or alcohol) for health. In the case where these behavioral responses affect mortality, the 2SLS estimates of the impact of TSPs on mortality would be invalid, although the estimated mortality effects of the Huai River heating policy (e.g., panel 1 of Table 3) would still be valid. Ultimately, the estimated impacts of TSPs on mortality should be interpreted with these caveats in mind because the necessary data to test for these behavioral responses are unavailable.

Conclusions

The analysis suggests that the Huai River policy, which had the laudable goal of providing indoor heat, had disastrous consequences for health, presumably due to the failure to require the installation of sufficient pollution abatement equipment. Specifically, it led to TSP concentrations that were 184 $\mu g/m^3$ higher (95% CI: 60, 308) or 55% higher in the North and reductions in life expectancies of 5.52 y (95% CI: 0.8, 10.2) in the North due to elevated rates of cardiorespiratory mortality.

The population in Northern China between 1990 and 2000 exceeded 500 million. Consequently, our results imply that the Huai River policy led to a staggering loss of over 2.5 billion life years. Furthermore, data from 2003 to 2008 indicate that PM_{10} (particulate matter smaller than 10 μm) concentrations are 22.9 $\mu g/m^3$ higher (95% CI: 13.5, 23.3) or 26% higher north of the Huai River, suggesting that residents of the North continue to have shortened lifespans. The TSP concentrations that prevailed during the study period greatly exceed the current concentrations in developed countries but are not atypical for many cities in developing countries today, such as India and China. These results may help explain why China's explosive economic growth has led to relatively anemic growth in life expectancy. More broadly, this paper's results may be useful in forming policy as developing countries search for the optimal balance between economic growth and environmental quality (28).

ACKNOWLEDGMENTS. We thank Douglas Almond for his insightful comments, criticisms, and support on all aspects of this paper and Janet Currie, Lucas Davis, Amy Finkelstein, Panle Jia, Damon Jones, Zhigang Li, Doug Miller, and Paulina Oliva for valuable comments. Ilya Faibushevitch, Joan Fang, Alison Flamm, Eyal Frank, Michael Freedman, Christine Pal, Susan Schwartz, and Yufei Wu provided excellent research assistance. Generous financial support was given by the Robert Wood Johnson Foundation, the Falk Institute, the Israel Foundation Trustees, and the National Natural Science Foundation of China: 71025004 and 71121001 (to H.L.), and 71073002 and 70903003 (to Y.C.).

1. Bi X, Feng Y, Wu J, Wang Y, Zhu T (2007) Source apportionment of PM10 in six cities of Northern China. *Atmos Environ* 41(5):903–912.
2. World Bank (2007) *Cost of Pollution in China: Economic Estimates of Physical Damages* (The World Bank, Washington, DC), Conference Ed.
3. Anonymous (August 21, 2004) Special report: A Great Wall of waste—China's environment." *Economist*, pp 55–57.
4. Dockery DW, et al. (1993) An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 329(24):1753–1759.
5. Pope CA, 3rd, et al. (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 287(9):1132–1141.
6. Almond D, Chen Y, Greenstone M, Li H (2009) Winter heating or clean air? Unintended impacts of China's Huai River policy. *Am Econ Rev* 99(2):184–190.
7. Ransom MR, Pope CA, 3rd (1992) Elementary school absences and PM10 pollution in Utah Valley. *Environ Res* 58(2):204–219.
8. Chay K, Greenstone M (2003) The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Q J Econ* 118(3):1121–1167.
9. Chay K, Dobkin C, Greenstone M (2003) The Clean Air Act of 1970 and adult mortality. *J Risk Uncertain* 27(3):279–300.
10. Cropper M 2010. What are the health effects of air pollution in China? *Is Economic Growth Sustainable?* (Palgrave-Macmillan, London).
11. Anonymous (1981–2001) *China Environment Yearbook* (China Environment Protection Agency and National Bureau of Statistics, Beijing).
12. Andrews SQ (2008) Inconsistencies in air quality metrics: Blue sky days and PM10 concentrations in Beijing. *Environ Res Lett* 3(3):034009.
13. Li H, Zhou L (2005) Political turnover and economic performance: The incentive role of personnel control in China. *J Public Econ* 89:1743–1762.
14. Andrews SQ 2008. Seeing through the smog: Understanding the limits of Chinese air pollution reporting. *Woodrow Wilson International Center for Scholars China Environment Forum 10*, pp 5–32.
15. Anonymous (1991–2001) *China's Disease Surveillance Points* (Chinese Center for Disease Control, Beijing).
16. Yang G, et al. (2005) Mortality registration and surveillance in China: History, current situation and challenges. *Popul Health Metr* 3(1):3.
17. Dominici F, et al. (2006) Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA* 295(10):1127–1134.
18. Hong YC, Lee JT, Kim H, Kwon HJ (2002) Air pollution: A new risk factor in ischemic stroke mortality. *Stroke* 33(9):2165–2169.
19. Hong YC, et al. (2002) Effects of air pollutants on acute stroke mortality. *Environ Health Perspect* 110(2):187–191.
20. Wellenius GA, Schwartz J, Mittleman MA (2005) Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries. *Stroke* 36(12): 2549–2553.
21. Kabir Z, Bennett K, Clancy L (2007) Lung cancer and urban air-pollution in Dublin: A temporal association? *Ir Med J* 100(2):367–369.
22. Pope CA, 3rd (1989) Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am J Public Health* 79(5):623–628.
23. Dab W, et al. (1996) Short term respiratory health effects of ambient air pollution: results of the APHEA project in Paris. *J Epidemiol Community Health* 50(1, Suppl 1): s42–s46.
24. World Meteorological Association. Data taken from the World Weather Watch Program according to WMO Resolution 40 (Cg-XII). Available at http://www.ncdc.noaa.gov/.
25. Cook TD, Campbell DT (1979) *Quasi-Experimentation: Design and Analysis for Field Settings* (Rand McNally, Chicago).
26. Lee D, Lemieux T (2010) Regression discontinuity designs in economics. *J Econ Lit* 48(2):281–355.
27. Pope CA, 3rd, Ezzati M, Dockery DW (2009) Fine-particulate air pollution and life expectancy in the United States. *N Engl J Med* 360(4):376–386.
28. Alberini A, et al. (1997) Valuing health effects of air pollution in developing countries: The case of Taiwan. *J Environ Econ Manage* 34:107–126.

SUPPORTING INFORMATION

Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy

PART 1: Description of Data

I. Mortality Rate and Life Expectancy Data

Our sample of mortality in China is taken from the Disease Surveillance Points system, which forms a nationally-representative sample of mortality for 1991-2000. These data are collected by the Chinese Center for Disease Control and were made available to the research team for this project.

Table S1 reports the death rates by cause at the DSP locations. The categories are selected from the International Classification of Disease Revision 9 (ICD-9). Cardiorespiratory causes of death are lung cancer (10), heart diseases (25, 26, 27, 28), vascular disease (29), and respiratory diseases (31, 32). Non-cardiorespiratory related illnesses are cancers other than lung (8, 9, 11, 12, 13, 14), accidents and violent deaths (47-56), and all other causes (1-7, 15-24, 30, 33-46, 99). It is worth noting that these death rates match the rates reported in Yang (1), an official tabulation of the death rates recorded by the Chinese Disease Surveillance Points system.

Since the paper's goal is to determine the relationship between mortality rates and long-term exposure to ambient air pollution, it is critical that the DSP accurately reports mortality among all permanent residents. There are two possible threats to the accurate calculation of such mortality rates. First, although China's *hukou* (household registration) system restricted migration, it did not prevent all migration, and indeed migration became more common in China by the end of our sample period. Second, sick rural residents are referred to urban hospitals with an unknown frequency; if their deaths are registered where they die instead of where they reside, our estimates could be biased.

Although we do not have a silver bullet for these problems, there are two primary reasons that we believe that mismeasurement of location-specific mortality rates are unlikely to explain the paper's results. First, the DSP's approach to assigning mortality rates to locations mitigates these potential problems because, like the *hukou* system, it assigns a death to an individuals' *hukou* or place of birth. At the DSP locations, the covered population is tracked from a registry designed at the outset of the program, and deaths are usually recorded by personnel at local hospitals. In the circumstance that a death occurs away from home, the family is instructed to report on the death wherever it occurs, and a site official is expected to follow up on and accurately record each death at the site of original registration. Further, the *hukou* system requires that all death certificates report the *hukou* or place of birth, and this information is checked by a local government official who provides an official stamp of verification to the death certificate. Second, the regression discontinuity design is likely to provide a solution to any residual confounding. Specifically, in the context of this design these two behaviors (migration and excess urban deaths) will not be a source of confounding unless their frequency changes discretely at the Huai River boundary. We are unaware of a reason that this would be the case.

Our analysis is conducted with cross-sectional data on mortality rates and their determinants. We converted the panel DSP data into a cross-section by taking averages of the cause-specific mortality rates, age-specific mortality rates, and life expectancy, from all annual observations for 1991-2000 for each of the 145 DSP locations.[1]

II. Air Pollution and Weather Data

The creation of the air pollution and weather exposure data sets involved several steps. We describe them here.

A. Creating a Panel Data Set of City-Level Pollution

The air pollution data used for the analysis were collected from China's Environmental Yearbooks[2], which contain nearly-comprehensive readings for 90 cities from 1981-2001 (2).

1. The data from 1981-1995 were taken from the World Bank online archive[3] of air quality readings, and supplemented by hand-entered readings from the Chinese language yearbooks. For this period, we have a total of 760 valid readings, 680 of which were taken from the World Bank online archive and 80 readings were entered by our team directly from print copies of the yearbooks.

[1] The DSP sample includes 1,374 DSP year-location observations, instead of 1,450 (145 x 10) as a result of difficulties encountered in the DSP collection process at certain sites in certain years.
[2] Issued by the State Statistical Bureau and accessed in print form by the authors at Harvard University's Fung Library.
[3] The World Bank online data is also taken from China's Environmental Yearbooks, but did not contain all available data for the cities included in our sample. Therefore, we supplemented these data with our hand-entered data from hard copies of the Yearbooks.

2. The data for 1995-2001 were entered by hand by two different researchers on our team, with nearly perfect data agreement. Out of a total of 508 readings for 1996-2001, there were 18 total cases of disagreement, and in these cases averages of the two readings were taken.

3. Steps 1 and 2 created a data set of 1,268 validated TSP measurements from 1981-2001 for 90 cities with monitoring stations. This was from a possible total of 1,890 readings, with most of the missing readings for the 1980s. The missing values were imputed (when possible) using linear interpolation from valid measurements in years prior to or following the missing value. The interpolation provided an estimate for an additional 138 readings, yielding a data set of 1,406 measurements for our analysis.

B. Creating a Panel Data Set of DSP location-Level Pollution Data from the City-Level Pollution Panel

The next step was to use the city-level panel data to create a DSP location-level panel of pollution data. This was done in three steps:

1. We first calculated the distance between each of these monitoring stations and our mortality sample taken from China's Disease Surveillance Points (DSP). The distance between each of these 90 stations and the 145 DSP sites or locations yielded a full matrix of 145 X 90 calculated distances.[4]

2. Our measure of air pollution for a DSP location in a year was calculated as the weighted average of air pollution at each station, with the weights determined by the inverse of the distance between the two points.[5] When a station had no valid TSP reading for a particular year, it was assigned a zero weight for that year and did not enter into the calculation.

3. At each DSP location, the air pollution exposure was measured as the average TSP reading in all previous years for which data are available. For example, the reading for a DSP location in 1991 is the average of TSP readings from 1981-1990 at the location. This creates a panel data set of year by DSP location mortality readings with validated air pollution data averaged across all previous years.

[4] The centroid of the county containing the DSP location and the centroid of the city containing the monitoring station were used to calculate an exact distance between the two.
[5] The results are robust to different choices for the functional form. This is discussed further in Part 4.

C. Creating a Cross-Sectional Data Set of DSP-Level Pollution Data from the Panel DSP Data File.

The panel data set was then averaged across all observations for each DSP location to create a single observation for each of the 145 DSP locations. The empirical analysis is restricted to locations within 150 kilometers of a monitoring station. Additionally, we dropped 1 DSP location for having invalid mortality data.[6] This represents the main sample for our analysis. It contains 125 DSP locations with valid mortality data within 150 kilometers of an air quality monitoring station.

A similar method was used to calculate weather exposure at each DSP location. Daily temperature readings for 1981-2001 were obtained from the World Meteorological Association (3). Our analysis is limited to the 339 weather stations with nearly-complete weather data for 1981-2001.[7] At each of these stations, we calculated the total heating and cooling degree days for each year. The total heating and cooling degree exposure for a DSP location in a given year was calculated as the weighted average of the degree days at the 339 stations, with the weights determined by the distance between the DSP location and the station. The weather exposure for a DSP location in a year is the average of heating and cooling degrees in all previous years during the sample period assigned in a manner similar to our method for assigning air pollution exposure.

III. Other Covariates

As a complement to the DSP, we also use the China 2000 county census data to control for confounders that might vary across locations in China. We assigned county-level information to each of the DSP locations by spatially merging the DSP locations with ArcGIS shape files of the 2000 census. For example, in the main regressions, we include controls for the average education, share in manufacturing, share of minority population, share with urban registration, and share with access to tap water. These data are taken from the Harvard Geospatial Library collection of China census data, linked to the DSP locations. Information on the average

[6] The DSP location in the Meilie District of Sanming City, Fujian Province, is excluded with a calculated life expectancy of 182 years. This was determined to be an outlier, with the unrealistic life expectancy due to either unregistered deaths or over-registered population counts, at the DSP location.

[7] Only stations with at least 360 days of data for every year in the period were included in the sample of weather stations. If a station had fewer than 360 readings in a particular year, the station was assigned a zero weight when estimating the weather for the DSP location for that year.

education, the share of minority population and share with urban registration are based on tabulations of the Short Form (100% sample) of the survey, and information on manufacturing employment and access to tap water are taken from the Long Form (9.5% sample) of the survey.

PART 2: Information on Dietary Patterns and Smoking Rates Regionally

We examined diet and smoking patterns in North and South China as they are potential confounders for determining the causal relationship between TSP and mortality rates. These variables come from the China Health and Nutrition Survey (1989-2006) and the China Household Income Survey (1995), respectively. Neither survey contains geographic identifiers at a level that would allow for reliable assignment to DSP locations, so these variables cannot be included in the regressions. Instead, this subsection assesses whether there are differences in these variables in the North and South which may be informative about differences at the river's border, since our identification strategy requires that no variables relevant to human health (other than air pollution) change discontinuously at the river.

Table S2 demonstrates that dietary patterns in terms of broad caloric intake are similar. Note also that the main dietary difference between Southern and Northern China is that the southern provinces have a diet with greater intake of rice and Northern China has greater intake of wheat. This would presumably not affect non-cardiorespiratory related illnesses differently than it affects non-cardiorespiratory related illnesses. In fact, the Yangtze River is generally thought the cultural divide between Northern and Southern China that is relevant in terms of diet, rather than the Huai River. Smoking rates are also quite similar between the South and North for men (69.2 versus 71.8), whereas the rates are somewhat higher in the North than South for women (2.7 versus 6.9). These differences seem quite small relative to the regional differences in cardiorespiratory illness rates. Also note that our empirical specifications control for a polynomial in the latitude of a DSP location, so the Two-stage least squares estimates exploit the discrete change in TSP near the Huai River. We suspect that it is unlikely that dietary patterns or smoking rates change nearly as dramatically as TSP near this boundary.

PART 3: Heterogeneity by Gender and Age

I. Estimates by Gender

As shown in Table S3, the results when estimated separately for men and women fail to reveal important gender differences in the effect of TSP on mortality. For example, the saturated model with a full set of controls in the overall sample indicates that an additional 100 μg/m³ of TSP imposes a 2.5/3.6 year cost in life expectancy at birth for men and women respectively. The results are less precisely estimated than what we find for the overall sample, but the estimates are statistically significant at the 10 and 5 percent levels respectively. One interpretation of a larger effect among women than men is that Chinese men have much higher smoking rates than Chinese women, and so ambient air pollution is a larger factor in determining female respiratory health than it is for men.

II. Estimates by Age

In Table S4, we examine the relationship between TSP exposure and mortality rates separately for different age groups. The results indicate that air pollution has a deleterious effect on health throughout the life cycle. Among infants, we observe a large impact on mortality for both cardiorespiratory related and non-cardiorespiratory related mortality. The non-cardiorespiratory mortality impact for infants appears to be due to the crude classification of the causes of mortality among infants in the DSP data.[8] Indeed, the difficulty in assigning causes to infant deaths leads us to conclude that the all cause mortality results are the most meaningful for infants.

At older ages, the results indicate that there is a stronger relationship between age-specific mortality for cardiorespiratory related causes than for non-cardiorespiratory related causes. For example, the results implicating air pollution in higher mortality rates for cardiorespiratory related illnesses are most pronounced for those ages 40 and older, where we estimate coefficients that range from .08 to .28, with the largest effects observed among those

[8] Nearly half of all infants deaths are assigned to causes of death that are almost never used for deaths at older ages. Specifically, 37.9% of infant deaths are classified as “certain conditions originating in the perinatal period” and 9.5% are assigned to “congenital anomalies.” These causes of death fall outside the range of ICD-9 codes for our cardiorespiratory category. They are very broad and non-specific causes of death and we suspect that many of the deaths are likely to be related to cardiac or respiratory conditions. However rather than change the definition of cardiorespiratory diseases for infants only, the analysis keeps the definition of cardiorespiratory mortality constant across all ages. The bottom line is that given the crude classification system for infant causes of death, we conclude that the all cause mortality results are the most meaningful for infants.

ages 60-80. The results for non-cardiorespiratory related causes are smaller and statistically insignificant. These results suggest that air pollution has a noticeable impact on human health at all ages, with slightly larger estimated effects among the elderly, possibly due to their longer-term exposure and/or their weaker condition. Figure S4 plots the TSP coefficients from the cardiorespiratory mortality rate regressions.

PART 4: Robustness of Results to Alternative Specifications and Samples

In this section, we examine the robustness of our results according to several different tests to confirm whether our results were affected qualitatively by the decisions made in our paper along several dimensions, such as data assignment, sample selection, and functional forms of our models.

I. Assignment of TSP Concentrations and Weather Variables to DSP Locations

Table S5 examines the sensitivity of the results to alternative approaches to calculating distance weighted averages of TSP and weather for each DSP location, and to our choice of whether to use the closest monitoring station versus a distance-weighted set of monitoring stations. One potential concern is that using distance-weighted averages between a DSP location and TSP readings from monitors will cause the calculated TSP change to vary smoothly with latitude, and this may incorrectly attenuate the estimated effect of the Huai River policy on TSP concentrations. We chose the distance-weighted method as it represents a conservative estimate of the impact of the policy on TSP, though we acknowledge that other choices could have been made. In Table S5, each cell represents a separate regression, with a dummy for "North" the reported independent variable. In the first panel, TSP is the dependent variable. In the second panel, heating and cooling degree days are the dependent variables in the first and second rows respectively. All models include a cubic polynomial in latitude.

In columns (1)-(4) of Table S5, the TSP and weather variables for a given DSP location are calculated as the weighted average of the nearby monitoring stations, with the weight given by the inverse of the distance, the square of the distance, the cube of the distance, and the quartic of the distance respectively. The results are not very sensitive to using alternative methods for assigning TSP and weather variables to DSP locations. For example, the estimated first stage relationship varies between 244.4 μg/m³ and 247.5 μg/m³, and all estimates are statistically significant at the 5% level. The relationship between “North” and the weather measures is qualitatively similar across the specifications as well, suggesting that our results are robust to different treatment of TSP and weather assignment to the DSP locations.

In columns (5)-(11) of Table S5, we examine the sensitivity of the results to alterations in the threshold for the cases where we only use the nearest monitor (rather than a weighted average). The analysis in the paper assigns TSP and weather based only on the nearest station when the nearest station is within 25 km of the DSP location and uses weighted averages across stations in cases where the nearest station is further than 25 km but less than 150 km away. The estimated first-stage relationship varies between 199.5 $\mu g/m^3$ and 252.2 $\mu g/m^3$, with larger estimates found for specifications with smaller values of the threshold which is consistent with some modest attenuation due to measurement error or due to our distance-weighting method that may attenuate the estimated impact of the policy. The weather regressions are generally unaffected by these alterations in the assignment rule. Overall, there is little evidence that the results are affected in any meaningful way by the rule to assign TSP and weather to DSP locations.

II. Maximum Distance Between Monitoring Stations and DSP Locations

The main analysis is restricted to all DSP locations within 150 kilometers of TSP monitoring station. Table S6 examines the robustness of the results to using stricter criteria and only including DSP locations within 150, 125, 100, 75, 50, or 25 kilometers of a monitoring station. To the extent that there is a pattern, it appears that the impact of the Huai River policy on TSP and life expectancy is larger in the more restricted samples where we focus on a smaller number of DSP locations with closer monitoring stations. However, on the whole, these estimates are qualitatively similar to those in Table 3 in that they continue to suggest sharp declines in TSP and life expectancy (due to increases in the cardiorespiratory related mortality rate) north of the Huai River.

III. Expanding the Sample to Include All DSP Locations

Table S7 examines the relationship between health outcomes and living north of the Huai River among *all* DSP locations, even those that are not within 150 kilometers of an air quality monitoring station. For this larger sample of locations without complete TSP data, we simply report the reduced form relationship between health outcomes and an indicator variable for "North". The sample includes all 144 DSP locations, relative to the sample of 125 locations in the main analysis. Among DSP locations within 5 degrees latitude of the Huai River, the expanded sample includes 77 locations, relative to the 69 locations with valid TSP data in the primary sample.

In this larger sample, the estimated effect of living north of the Huai River is qualitatively similar to the results reported in the main analysis. Specifically, all-cause mortality rates are estimated to be 17-25% higher and this is entirely driven by higher cardiorespiratory related mortality rates that are elevated by 31-43%. Overall, the table indicates that living north of the Huai River causes a 3.7-4.7 year decline in life expectancy.

IV. Adjustment for Distance from the Coast

The specifications that produce the estimates in Table S8 are identical to those in Table 3, except that in columns (2) and (3) there is an additional control variable for the distance in meters from China's coast. The adjustment for distance from the Coast aims to account for factors related to China's rapid integration into the global economy and manufacturing boom in coastal cities. As reflected, the results controlling for a DSP location's distance from the coast are very similar to the results we report in Table 3, and in fact are slightly more precisely estimated.

V. Alternative Approaches to Implementation of the Regression Discontinuity Design

We additionally explored the robustness of the results to alternative approaches to implementing the RD design, noting that the design is demanding of the data. Table S9 replicates the column (2) specification from Table 3 but with alternative choices for the polynomial in degrees latitude from the Huai River, starting with a linear polynomial in column (1) and then progressively adding a higher order term to each subsequent specification as one moves from left to right. For example, column (3) includes latitude, its square, and its cube. Below the point estimate and its standard errors, we report two measures of goodness of fit; the p-value associated with a chi-squared test statistic from the addition of the nth term in latitude is in square brackets and the Akaike Information Criterion (AIC) statistic is in braces (4).

The results support our emphasis on a cubic in degrees latitude from the Huai River. Across the outcome variables, the chi-squared test statistics favor the cubic polynomial. For TSPs, the AIC statistic is minimized with a quartic polynomial, although the cubic is the next favored model. In the case of all cause and cardiorespiratory mortality rates and life expectancy, the AIC criteria indicate that the cubic is the optimal order of the polynomial. Both statistics suggest that the linear and quadratic polynomials fail to fit the data as well as the higher order models; these poorer fitting models fail to find a significant relationship between living north of the river and mortality rates, which may be because the lower-order polynomials fail to absorb the variation in mortality patterns in cities that are farthest north and south. Notably, the estimates from the better fitting polynomials (i.e., the cubic, quartic, and quantic) are all similar.

Tables S10 and S11 explore the sensitivity of the results to limiting the sample to DSP locations near the Huai River, and to allowing the polynomial to vary to the north and south of the river. In the former case, the goodness of fit statistics favor the linear and quadratic polynomials; the results from these models are similar to those in Table 3. In the latter case, the goodness of fit statistics strongly support the use of a quadratic polynomial that is allowed to vary to the north and south of the river. Compared to the Table 3 estimates, this specification suggests a smaller increase in TSP at the border but a larger decline in life expectancy.

VI. Do Other Government Policies Change at the Huai River?

A natural concern related to the research design is that the government used the Huai River as the demarcation line for changes in other government policies related to public health, and this would confound the estimates of TSP on health. This possibility is mitigated by the fact that the Huai River is not a border used for administrative purposes. The Huai River follows the January zero degree average temperature line (Celsius), and this was in fact the basis of its choice as a method to divide the country for free heating. Further, local policies generally hew to administrative boundaries associated with cities and provinces; indeed, the Huai River cuts through several provinces. Nevertheless, we identified and compiled variables on policies that are plausibly related to health across 261 cities from the China City Statistical Yearbooks. Table S12 reports on the results of fitting regressions for these policy variables, where the parameter of interest is associated with an indicator for North. The outcome variables are: counts of hospitals, beds in hospitals, doctors, and social service workers; total health care, sports and social welfare workers; and, budgetary expenditures of local governments. These variables are calculated as averages for the years 1996 through 2000 (data are unavailable for 1991-1995). Column (1) reports on the parameter associated with the indicator for North without adjustment for any covariates. Column (2) adjusts the estimate for a cubic in latitude, while column (3) restricts the sample to cities that are roughly 5° latitude north or south of the Huai River (i.e., cities between 27.5° and 37.5° latitude) and adjusts for a linear function of latitude. The restricted sample of cities in column (3) includes 138 cities.

The results fail to suggest any substantial differences in policies north and south of the Huai river. The column (2) specification indicates that the null hypothesis of equality can be rejected at the 10% level for two variables: the number of social service workers and the number of health care, sports and social welfare workers. The column (3) specification indicates that the null hypothesis of equality can only be rejected at the 10% level for the number of hospitals. However, there is little evidence of differences in these variables from the other specifications, or for the other variables in any of the specifications.

Table S13 further explores the possibility that the paper's results are due to differences in other policies north and south of the Huai River. It reports on specifications that are identical to those in Table 3, except that in columns (2) and (3) the full set of policy variables in Table S12 are included as covariates. The results are qualitatively identical to those in Table 3 but estimated with slightly more precision. If anything, they suggest that the Huai River policy and TSP exposure lead to larger losses of life expectancy than indicated in Table 3.

References

1. Yang Gonghuan, Jianping Hu, Ke Quin Rao, Jemin Ma, Chalapati Rao, Alan D. Lopez. 2005. "Mortality registration and surveillance in China: History, current situation and challenges." *Population Health Metrics*, 3(3).
2. *China's Environmental Yearbook*. 1981-2001. China Environmental Protection Agency and National Bureau of Statistics.
3. *World Meteorological Association.* Data taken from the World Weather Watch Program according to WMO Resolution 40 (Cg-XII). http://www.ncdc.noaa.gov/
4. Akaike, Hirotugu (1974). "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* 19 (6): 716–723.

Table S1

Age-adjusted death rates (per 100,000) by cause in China, 1991-2000

	South	North	Difference In Means	p-value
	(1)	(2)	(3)	(4)
All Cause Mortality Rate	624.9	644.9	20.0	0.589
	(155.1)	(221.3)	(36.9)	
<u>Cardiorespiratory Mortality Rate</u>				
All Cardiorespiratory Mortality	353.7	400.3	46.6*	0.100
	(114.7)	(166.2)	(28.1)	
Heart	108.7	147.7	39.0***	0.005
	(55.2)	(76.6)	(13.6)	
Stroke	102.7	133.7	31.0***	0.005
	(46.2)	(64.2)	(10.8)	
Lung Cancers	21.1	25.6	4.5**	0.050
	(12.7)	(12.5)	(2.3)	
Respiratory Illnesses	121.2	93.3	-27.9*	0.075
	(74.8)	(84.2)	(15.5)	
<u>Non-Cardiorespiratory Mortality Rate</u>				
All Non-Cardiorespiratory Mortality	271.2	244.5	-26.6*	0.065
	(74.4)	(78.5)	(14.3)	
Cancers Other than Lung	82.5	91.1	8.5	0.307
	(41.8)	(39.9)	(8.3)	
Accidents / Violence	72.4	59.9	-12.5**	0.016
	(30.0)	(26.9)	(5.1)	
Other	116.3	93.6	-22.7***	0.004
	(53.9)	(36.6)	(7.7)	

Source : Chinese Disease Surveillance Points Mortality Registration System (DSP)

Note : N=145. The cardiorespiratory mortality causes are heart disease, stroke, lung cancer and other respiratory illnesses (e.g. pneumonia, bronchitis). The non-cardiorespiratory mortality causes are violence, cancers other than lung, and all other causes (e.g. diseases of the digestive system). Age adjustment is performed by calculating age-specific death rates and creating weighted averages using the population structure in China's 2000 census. The classification of North is determined by whether the DSP location is covered by the home heating policy, which is all cities north of the line formed by the Huai River and Qinling Mountain range (approximately 33^{o} latitude). The DSP has 75 sites located north of the Huai River boundary and 70 sites south of the Huai River boundary. Heteroskedastic-consistent standard errors are reported in parentheses.

Table S2

Dietary and smoking habits, South and North of the Huai River.

	South	North
	(1)	(2)
Dietary Patterns		
Total Caloric Intake	2,313	2,263
Carbohydrates (%)	15.04	15.32
Fat (%)	2.95	2.77
Protein (%)	2.87	2.98
Other (%)	79.14	78.93
Smoking Rates by Gender		
Men	0.692	0.718
Women	0.027	0.069

Source : Smoking rates are taken from the China Household Income Survey (CHIS, 1995). Dietary information is taken from the China Health and Nutrition Survey (CHNS, 1989-2006).

Note : The smoking rates are shown for the DSP locations which were in the 19 provinces included in the CHIS (1995), which includes 102 of the 145 DSP locations. Information on diet is shown for DSP locations located in the 9 provinces included in the CHNS, which includes 55 of the 145 locations.

Table S3

Using the Huai River Policy to Estimate the Impact of Total Suspended Particulates (100 μg/m^3) on Health Outcomes, Males and Females

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares						
TSP (100 μg/m^3)	2.48***	1.84***	2.17***	2.48***	1.84***	2.17***
	(0.65)	(0.63)	(0.66)	(0.65)	(0.63)	(0.66)
ln(All Cause Mortality Rate)	0.20*	0.25**	0.31**	0.2400	0.27*	0.26
	(0.12)	(0.12)	(0.14)	(0.15)	(0.15)	(0.17)
ln(Cardiorespiratory Mortality Rate)	0.33**	0.36**	0.53***	0.43**	0.41**	0.46**
	(0.14)	(0.15)	(0.18)	(0.18)	(0.19)	(0.21)
ln(Non-Cardiorespiratory Mortality Rate)	0.03	0.10	0.02	-0.05	0.03	-0.04
	(0.12)	(0.12)	(0.13)	(0.14)	(0.14)	(0.14)
Life Expectancy (years)	-3.99*	-4.52*	-5.78*	-6.01**	-6.55**	-4.35
	(2.39)	(2.30)	(2.94)	(2.87)	(2.83)	(3.00)
Panel 2: Impact of TSP (100 μg/m^3) on the Listed Variable, Two-stage Least Squares						
ln(Total Death Rate)	0.08*	0.14**	0.15*	0.10*	0.14*	0.12
	(0.05)	(0.07)	(0.08)	(0.06)	(0.08)	(0.08)
ln(All Cause Mortality Rate)	0.13**	0.19**	0.25**	0.17**	0.22**	0.21**
	(0.05)	(0.09)	(0.10)	(0.07)	(0.10)	(0.10)
ln(Cardiorespiratory Mortality Rate)	0.01	0.06	0.01	-0.02	0.02	-0.02
	(0.05)	(0.06)	(0.06)	(0.06)	(0.08)	(0.07)
ln(Non-Cardiorespiratory Mortality Rate)	-1.61*	-2.46*	-2.66*	-2.43**	-3.56**	-2.00
	(0.92)	(1.27)	(1.57)	(1.05)	(1.55)	(1.50)
Climate Controls	N	Y	Y	N	Y	Y
Census and DSP Controls	N	Y	Y	N	Y	Y
Polynomial in Latitude	Cubic	Cubic	Linear	Cubic	Cubic	Linear
Only DSP sites within 5^0 latitude	N	N	Y	N	N	Y

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : This table reports results estimated in an identical manner to those presented in Table 3 but separately for men and women.

Table S4

Using the Huai River Policy to Estimate the Impact of Total Suspended Particulates (100 μg/m^3) on Health Outcomes

Age	All Cause Mortality	Cardiorespiratory Mortality	Non-Cardiorespiratory Mortality
		(1)	(2)
0	0.55***	0.36*	0.63***
	(0.20)	(0.22)	(0.22)
1	0.01	-0.31	-0.27
	(0.18)	(0.26)	(0.17)
5	-0.04	-0.13	-0.03
	(0.12)	(0.19)	(0.12)
10	-0.11	0.21	-0.09
	(0.13)	(0.26)	(0.14)
15	0.01	0.17	0.00
	(0.10)	(0.17)	(0.13)
20	0.10	0.08	0.12
	(0.09)	(0.21)	(0.09)
25	0.13*	0.21	0.12
	(0.08)	(0.15)	(0.08)
30	0.01	0.20	-0.03
	(0.07)	(0.16)	(0.08)
35	0.05	0.04	0.06
	(0.06)	(0.1)	(0.07)
40	0.12	0.18*	0.09
	(0.07)	(0.11)	(0.07)
45	0.08	0.21*	0.01
	(0.07)	(0.12)	(0.07)
50	0.07	0.17*	0.01
	(0.06)	(0.09)	(0.08)
55	0.00	0.08	-0.07
	(0.07)	(0.1)	(0.1)
60	0.17**	0.21**	0.14
	(0.08)	(0.09)	(0.1)
65	0.14**	0.20**	0.08
	(0.07)	(0.09)	(0.09)
70	0.14*	0.19**	0.04
	(0.07)	(0.09)	(0.09)
75	0.19**	0.24**	0.10
	(0.09)	(0.11)	(0.11)
80	0.13	0.19	0.00
	(0.11)	(0.13)	(0.14)

* significant at 10% ** significant at 5%. *** significant at 1%.

Source: China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : N=125. The sample is restricted to DSP locations within 150 kilometers of an air quality monitoring station. Each cell in the table represents the coefficient from a separate regression, and heteroskedastic-consistent standard errors are reported in parentheses. The models are estimated in an identical manner to those in Table 3 but separately by age group.

Table S5

Robustness checks using weighted averages of TSP across the air quality and weather monitoring stations

	Weighted by:				Only use Nearest Station for Sites within (x) of a Monitor:						
	Distance	Distance^2	Distance^3	Distance^4	10 km	25 km	50 km	75 km	100 km	125 km	150 km
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel 1: Air Pollution Exposure at the Disease Surveillance Points											
Total Suspended	247.5***	246.0***	245.0***	244.4***	252.2***	247.5***	253.1***	234.5***	202.5***	200.1***	199.5***
Particulates (μg/m^3)	(65.2)	(65.7)	(66.1)	(66.4)	(65.6)	(65.2)	(67.4)	(59.6)	(63.5)	(62.9)	(62.9)
Panel 2: Weather Patterns at the Disease Surveillance Points											
Total Heating	465.2	485.7	492.4	493.8	455.8	465.2	673.1*	544.7	400.5	484.9	482.4
Degree Days (000s)	(353.6)	(348.2)	(351.8)	(357.5)	(354.2)	(353.6)	(367.9)	(384.8)	(381.7)	(430.8)	(428.3)
Total Cooling	-240.9	-242.0	-242.1	-240.7	-238.4	-240.9	-329.8*	-263.7	-166.1	-184.5	-182.6
Degree Days (000s)	(177.0)	(177.9)	(180.3)	(182.5)	(177.5)	(177.0)	(190.0)	(197.0)	(195.9)	(205.2)	(203.3)

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : N=125. Each cell in the table represents the coefficient from a separate regression, with a dummy for "North" the reported independent variable. Heteroskedastic-consistent standard errors are reported in parentheses. The first panel has TSP as the dependent variable, and the second panel has heating and cooling degree days as the dependent variable. All models include a cubic in latitude. The sample is restricted to DSP points within 150 kilometers of an air quality monitoring station, and calculation of TSP is only made using stations within 150km. In the first four columns, TSP for a given DSP site is calculated as the weighted average of the nearby monitoring stations, with the weight given by the inverse of the distance, the square of the distance, the cube of the distance, and the quartic respectively. In columns 5-11, we assign to DSP sites sufficiently close to a monitoring station the value for the station, and assign all others using weighted averages where the weight is given by the inverse of the distance. For example, in column 5, any DSP location within 10 kilometers of a station is assigned the value at the closest station instead of a weighted average of the value at multiple stations.

Table S6

Robustness checks of choice of acceptable distance from monitoring station

	<150KM	<125KM	<100KM	<75KM	<50KM	<25KM
	(1)	(2)	(3)	(4)	(5)	(6)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares						
TSP (100 μg/m^3)	1.84***	1.88***	1.49**	2.17**	4.81***	2.91**
	(0.63)	(0.65)	(0.64)	(1.06)	(1.50)	(1.08)
ln(All Cause Mortality Rate)	0.26*	0.22*	0.30**	0.52***	0.30	0.45*
	(0.13)	(0.13)	(0.13)	(0.15)	(0.18)	(0.23)
ln(Cardiorespiratory Mortality Rate)	0.38**	0.34**	0.42**	0.70***	0.55***	0.66**
	(0.16)	(0.17)	(0.17)	(0.19)	(0.20)	(0.27)
ln(Non-Cardiorespiratory Mortality Rate)	0.08	0.06	0.12	0.25*	0.00	0.18
	(0.13)	(0.12)	(0.11)	(0.14)	(0.21)	(0.24)
Life Expectancy (years)	-5.52**	-4.89*	-5.38**	-9.49***	-9.67***	-8.20*
	(2.39)	(2.47)	(2.52)	(2.85)	(3.54)	(4.12)
Panel 2: Impact of TSP (100 μg/m^3) on the Listed Variable, Two-stage Least Squares						
ln(All Cause Mortality Rate)	0.14**	0.12*	0.21**	0.24**	0.06	0.15
	(0.07)	(0.06)	(0.10)	(0.12)	(0.05)	(0.11)
ln(Cardiorespiratory Mortality Rate)	0.21**	0.18**	0.29**	0.32**	0.11*	0.23
	(0.09)	(0.09)	(0.13)	(0.15)	(0.06)	(0.15)
ln(Non-Cardiorespiratory Mortality Rate)	0.04	0.03	0.08	0.12	0.00	0.06
	(0.07)	(0.06)	(0.08)	(0.09)	(0.04)	(0.09)
Life Expectancy (years)	-3.00**	-2.61**	-3.62*	-4.38**	-2.01**	-2.82
	(1.33)	(1.28)	(1.95)	(2.18)	(0.84)	(2.02)
Observations	125	112	94	69	50	28

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : Each cell in the table represents the coefficient from a separate regression. Heteroskedastic-consistent standard errors are reported in parentheses. All models have a cubic in latitude, demographic controls from China's 2000 census and wealth controls from the DSP, and weather controls. Each column reports these results for the restricted sample of DSP locations with a monitoring station within the listed distance.

Table S7

Relationship between Living North of the Huai River and Health Outcomes using All Disease Surveillance Points

	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
ln(All Cause Mortality Rate)	0.17	0.22*	0.25*
	(0.12)	(0.12)	(0.14)
ln(Cardiorespiratory Mortality Rate)	0.31**	0.32**	0.43**
	(0.14)	(0.15)	(0.18)
ln(Non-Cardiorespiratory Mortality Rate)	-0.03	0.07	-0.01
	(0.11)	(0.12)	(0.11)
Life Expectancy (years)	-3.74*	-4.54**	-4.65*
	(2.16)	(2.26)	(2.69)
Climate Controls	N	Y	Y
Census and DSP Controls	N	Y	Y
Polynomial in Latitude	Cubic	Cubic	Linear
Only DSP locations within 5^0 latitude	N	N	Y

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : The sample in columns (1) and (2) includes all DSP locations with valid mortality data (N=144) and in column (3) is restricted to DSP locations within 5^0 latitude of the Huai River boundary (N=77). In this table, we include all DSP locations, including those that are not within 150 kilometers of an air quality monitoring station. One DSP location is excluded due to invalid mortality data. Each cell in the table represents the coefficient from a separate regression, and heteroskedastic-consistent standard errors are reported in parentheses. The cardiorespiratory mortality causes are heart disease, stroke, lung cancer and other respiratory illnesses. The non-cardiorespiratory mortality causes are violence, cancers other than lung, and all other causes. Models in column (1) include a cubic in latitude. Models in column (2) additionally include controls for average education, share employed in manufacturing, share minority, share with urban registration, share with tap water, and controls for the heating and cooling degrees days between 1981 and 2000 prior to the year being analyzed. Degree days are the sum of the difference between the temperature and 65^0F. Models in column (3) are estimated with a linear control for latitude. Regressions are weighted by the population at the DSP location.

Table S8

Using the Huai River Policy to Estimate the Impact of Total Suspended Particulates (100 µg/m^3) on Health Outcomes Controlling for Distance from the Coast

	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
TSP (100 µg/m^3)	2.48***	1.86***	2.19***
	(0.65)	(0.63)	(0.66)
ln(All Cause Mortality Rate)	0.22*	0.26**	0.32**
	(0.13)	(0.13)	(0.15)
ln(Cardiorespiratory Mortality Rate)	0.37**	0.39**	0.54***
	(0.16)	(0.16)	(0.18)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.08	0.01
	(0.13)	(0.13)	(0.13)
Life Expectancy (years)	-5.04**	-5.59**	-5.55**
	(2.47)	(2.35)	(2.78)
Panel 2: Impact of TSP (100 µg/m^3) on the Listed Variable, Two-stage Least Squares			
ln(All Cause Mortality Rate)	0.09*	0.14**	0.15*
	(0.05)	(0.07)	(0.08)
ln(Cardiorespiratory Mortality Rate)	0.15**	0.21**	0.25**
	(0.06)	(0.09)	(0.10)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.04	0.01
	(0.05)	(0.07)	(0.06)
Life Expectancy (years)	-2.04**	-3.00**	-2.54*
	(0.92)	(1.31)	(1.50)
Climate Controls	N	Y	Y
Census and DSP Controls	N	Y	Y
Distance from Coast	N	Y	Y
Polynomial in Latitude	Cubic	Cubic	Linear
Only DSP locations within 5^0 latitude	N	N	Y

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : These results are estimated in an identical manner to those presented in Table 3 but include an additional control for the DSP location's distance from the coast (in meters).

Table S9

Robustness checks of choice of functional form for latitude

	Linear & Controls	Quadratic & Controls	Cubic & Controls	Quartic & Controls	Quintic & Controls
	(1)	(2)	(3)	(4)	(5)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares					
TSP (100 μg/m^3)	2.89***	2.63***	1.84***	1.95***	1.52**
	(0.56)	(0.49)	(0.63)	(0.59)	(0.72)
	[0.988]	[0.068]	[0.148]	[0.229]	[0.671]
	{492.4}	{489}	{487.2}	{486.3}	{487.5}
ln(All Cause Mortality Rate)	0.12	0.09	0.26*	0.26**	0.37**
	(0.10)	(0.10)	(0.13)	(0.13)	(0.16)
	[0.276]	[0.215]	[0.035]	[0.908]	[0.409]
	{39.88}	{38.8}	{34.11}	{35.92}	{36.13}
ln(Cardiorespiratory Mortality Rate)	0.13	0.09	0.38**	0.39**	0.47**
	(0.13)	(0.13)	(0.16)	(0.16)	(0.19)
	[0.652]	[0.243]	[0.003]	[0.747]	[0.696]
	{102.3}	{101.5}	{91.92}	{93.34}	{94.62}
ln(Non-Cardiorespiratory Mortality Rate)	0.09	0.05	0.08	0.07	0.19
	(0.10)	(0.09)	(0.13)	(0.12)	(0.14)
	[0.135]	[0.151]	[0.933]	[0.973]	[0.35]
	{43.04}	{41.27}	{43.13}	{45.07}	{44.97}
Life Expectancy (years)	-1.62	-1.29	-5.52**	-5.67**	-5.43*
	(1.66)	(1.68)	(2.39)	(2.36)	(2.94)
	[0.101]	[0.6]	[0.001]	[0.737]	[0.984]
	{757.1}	{758}	{746.8}	{748.2}	{750.2}
Panel 2: Impact of TSP (100 μg/m^3) on the Listed Variable, Two-stage Least Squares					
ln(All Cause Mortality Rate)	0.04	0.03	0.14**	0.13**	0.24*
	(0.03)	(0.04)	(0.07)	(0.06)	(0.13)
ln(Cardiorespiratory Mortality Rate)	0.05	0.03	0.21**	0.20**	0.31*
	(0.04)	(0.05)	(0.09)	(0.08)	(0.17)
ln(Non-Cardiorespiratory Mortality Rate)	0.03	0.02	0.04	0.04	0.13
	(0.03)	(0.03)	(0.07)	(0.06)	(0.10)
Life Expectancy (years)	-0.56	-0.49	-3.00**	-2.90**	-3.56
	(0.54)	(0.62)	(1.33)	(1.24)	(2.34)

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : N=125. Each cell in the table represents the coefficient from a separate regression. All models have demographic controls from China's 2000 census, wealth controls from the DSP, and climate controls. Heteroskedastic-consistent standard errors are reported in parentheses. In panel 1, we also report in brackets the p-value associated with the addition of an nth term in latitude. This is calculated using a likelihood ratio test derived by comparing the log likelihood of a model estimated with the extra term restricted to be zero versus an unrestricted model where the extra term is estimated without restriction. The value of Akaike's Information Criterion (AIC) statistic is reported in braces.

Table S10

Robustness checks of choice of functional form for latitude, DSP locations within 5^0 Latitude of Huai River

	Linear & Controls	Quadratic & Controls	Cubic & Controls
	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
TSP (100 µg/m^3)	2.17***	1.18*	0.60
	(0.66)	(0.69)	(0.55)
	[0.576]	[0.0001]	[0.28]
	{232}	{216.4}	{215.8}
ln(All Cause Mortality Rate)	0.30*	0.24	0.33
	(0.15)	(0.17)	(0.20)
	[0.171]	[0.587]	[0.577]
	{21.91}	{22.84}	{23.74}
ln(Cardiorespiratory Mortality Rate)	0.50***	0.40*	0.51**
	(0.19)	(0.22)	(0.24)
	[0.042]	[0.359]	[0.652]
	{56.1}	{56.05}	{57.19}
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.02	0.09
	(0.13)	(0.15)	(0.18)
	[0.999]	[0.931]	[0.66]
	{8.704}	{10.56}	{11.73}
Life Expectancy (years)	-5.30*	-4.04	-5.74
	(2.85)	(3.48)	(3.84)
	[0.036]	[0.333]	[0.463]
	{400.3}	{400.1}	{400.6}
Panel 2: Impact of TSP (100 µg/m^3) on the Listed Variable, Two-stage Least Squares			
ln(All Cause Mortality Rate)	0.14*	0.20	0.56
	(0.08)	(0.17)	(0.56)
ln(Cardiorespiratory Mortality Rate)	0.23**	0.34	0.85
	(0.10)	(0.21)	(0.76)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.02	0.16
	(0.06)	(0.13)	(0.34)
Life Expectancy (years)	-2.44	-3.44	-9.57
	(1.50)	(3.26)	(10.03)

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : N=69. Each cell in the table represents the coefficient from a separate regression. Heteroskedastic-standard errors are reported in parentheses. All models have demographic controls from China's 2000 census, wealth controls from the DSP, and climate controls. In panel 1, we also report in brackets the p-value associated with the addition of an n^{th} term in latitude. This is calculated using a likelihood ratio test derived by comparing the log likelihood of a model estimated with the extra term restricted to be zero versus an unrestricted model where the extra term is estimated without restriction. The value of Akaike's Information Criterion (AIC) statistic is reported in braces.

Table S11

Robustness checks of choice of functional form for latitude allowing the effect to vary North and South of the Huai River

	Linear & Controls	Quadratic & Controls	Cubic & Controls
	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
TSP (100 μg/m^3)	2.87***	1.18*	0.52
	(0.55)	(0.64)	(0.81)
	[0.557]	[0.0003]	[0.517]
	{493.3}	{480.9}	{483.6}
ln(All Cause Mortality Rate)	0.11	0.33**	0.38**
	(0.10)	(0.15)	(0.19)
	[0.05]	[0.029]	[0.919]
	{38.45}	{35.33}	{39.16}
ln(Cardiorespiratory Mortality Rate)	0.13	0.49***	0.50**
	(0.13)	(0.18)	(0.24)
	[0.19]	[0.001]	[0.71]
	{101.9}	{91.8}	{95.12}
ln(Non-Cardiorespiratory Mortality Rate)	0.08	0.12	0.20
	(0.09)	(0.15)	(0.17)
	[0.008]	[0.764]	[0.456]
	{39.51}	{42.98}	{45.4}
Life Expectancy (years)	-1.57	-6.78**	-4.61
	(1.67)	(2.89)	(4.08)
	[0.059]	[0.002]	[0.575]
	{758}	{749.1}	{752}
Panel 2: Impact of TSP (100 μg/m^3) on the Listed Variable, Two-stage Least Squares			
ln(Total Death Rate)	0.04	0.28*	0.72
	(0.03)	(0.16)	(1.05)
ln(Cardiorespiratory Mortality Rate)	0.04	0.41*	0.95
	(0.04)	(0.22)	(1.41)
ln(Non-Cardiorespiratory Mortality Rate)	0.03	0.10	0.38
	(0.03)	(0.12)	(0.59)
Life Expectancy (years)	-0.55	-5.74*	-8.79
	(0.55)	(3.14)	(13.38)

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000).

Note : N=125. Each cell in the table represents the coefficient from a separate regression. All models have demographic controls from China's 2000 census, wealth controls from the DSP, and climate controls. Heteroskedastic-consistent standard errors are reported in parentheses. In panel 1, we also report in brackets the p-value associated with the addition of an n^{th} term in latitude and it interacted with a "North" dummy. This is calculated using a likelihood ratio test derived by comparing the log likelihood of a model estimated with the extra terms restricted to be zero versus an unrestricted model where the extra term is estimated without restriction. The value of Akaike's Information Criteria (AIC) statistic is reported in braces. This table differs from Table S9 in which latitude is restricted to take an equivalent functional form in the South and the North.

Table S12

Patterns in Health-Related Government Policies South and North of the Huai River

	(1)	(2)	(3)
Number of Hospitals	0.01	-0.02	-0.03*
(000s)	(0.01)	(0.02)	(0.02)
Number of Beds in Hospitals	1.29	-2.39	-2.89
(000s)	(0.85)	(1.59)	(2.13)
Number of Doctors	0.63	-1.70	-2.34
(000s)	(0.63)	(1.17)	(1.58)
Social Services Workers	0.14	-2.17*	-2.46
(10,000 persons)	(0.66)	(1.24)	(1.75)
Health Care, Sports and Social Welfare Workers	0.04	-0.61*	-0.73
(10,000 persons)	(0.19)	(0.36)	(0.51)
Budgetary Expenditure of Local Government	-338	-1,233	-1,817
(million yuan)	(420)	(932)	(1413)
Polynomial in Latitude	None	Cubic	Linear
Only Cities within +/- 5^0 degrees latitude of Huai	N	N	Y

* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China City Statistical Yearbooks (1996-2000)

Note : Each cell in the table represents the coefficient from a separate regression. Heteroskedastic-consistent standard errors are reported in parentheses. In each regression, we regress the listed variable on a "North" dummy variable. The sample in columns (1) and (2) includes all cities with information on social spending in the listed categories. The sample in column (3) is restricted to cities near the Huai River (27.50-37.50). The larger sample in columns (1) and (2) includes 261 cities and the restricted sample in column (3) includes 138 cities.

Table S13

Using the Huai River Policy to Estimate the Impact of Total Suspended Particulates (100 μg/m^3) on Health Outcomes Controlling for Health-Related Government Policies

	(1)	(2)	(3)
Panel 1: Impact of "North" on the Listed Variable, Ordinary Least Squares			
TSP (100 μg/m^3)	2.45***	1.93***	2.06***
	(0.66)	(0.67)	(0.65)
ln(All Cause Mortality Rate)	0.22*	0.30**	0.38**
	(0.13)	(0.12)	(0.16)
ln(Cardiorespiratory Mortality Rate)	0.37**	0.44***	0.60***
	(0.16)	(0.15)	(0.21)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.10	0.06
	(0.13)	(0.12)	(0.13)
Life Expectancy (years)	-5.04**	-6.65***	-6.08**
	(2.47)	(2.34)	(3.03)
Panel 2: Impact of TSP (100 μg/m^3) on the Listed Variable, Two-stage Least Squares			
ln(All Cause Mortality Rate)	0.09*	0.16**	0.18**
	(0.05)	(0.07)	(0.09)
ln(Cardiorespiratory Mortality Rate)	0.15**	0.23**	0.29**
	(0.06)	(0.10)	(0.12)
ln(Non-Cardiorespiratory Mortality Rate)	0.00	0.05	0.03
	(0.05)	(0.06)	(0.07)
Life Expectancy (years)	-2.06**	-3.44**	-2.95*
	(0.95)	(1.45)	(1.69)
Climate Controls	N	Y	Y
Census and DSP Controls	N	Y	Y
Health-Related Government Policy Controls	N	Y	Y
Polynomial in Latitude	Cubic	Cubic	Linear
Only DSP locations within 5^0 latitude	N	N	Y

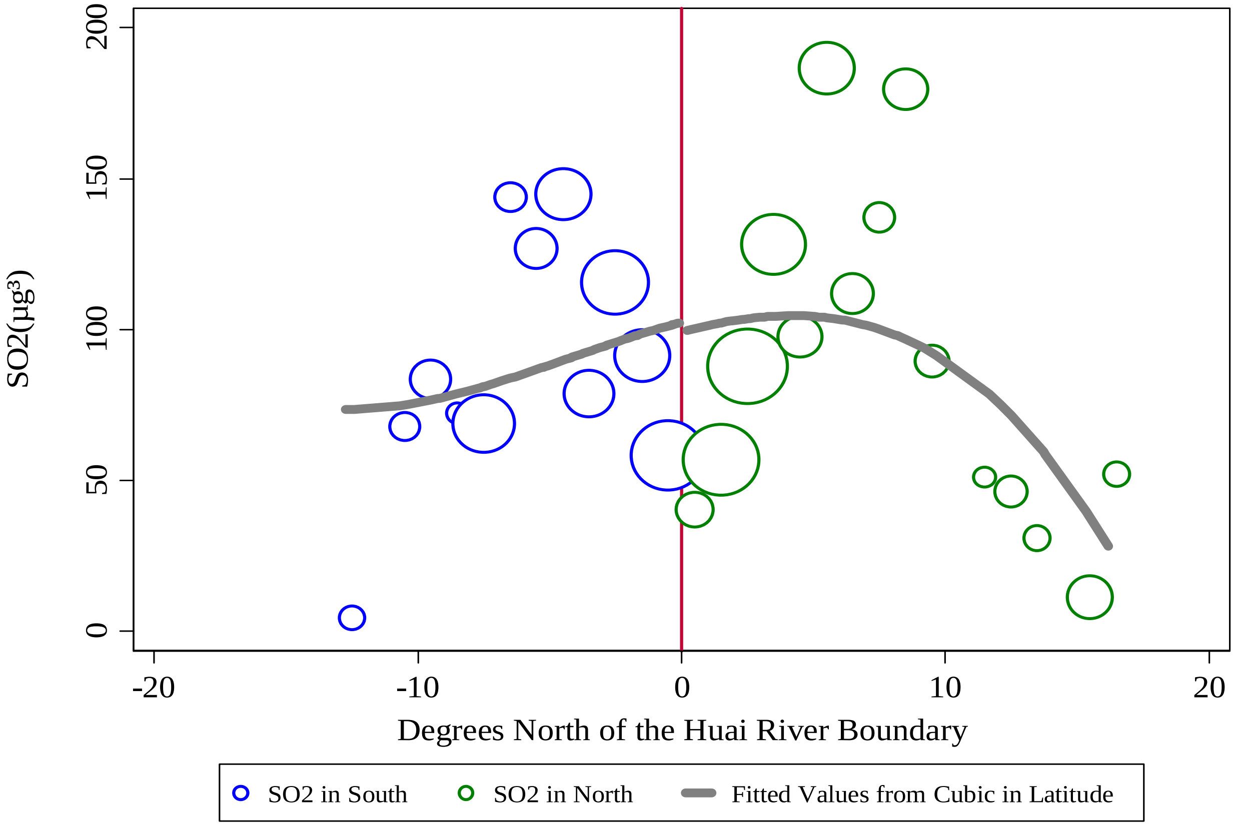
* significant at 10% ** significant at 5%. *** significant at 1%.

Source : China Disease Surveillance Points (1991-2000), China Environmental Yearbooks (1981-2000), World Meteorological Association (1980-2000), China City Statistical Yearbooks (1996-2000).

Note : The regressions are estimated in a manner identical to those presented in Table 3, but in columns 2 and 3 we include the additional health-related government policies listed in Table S12.

Figure S1

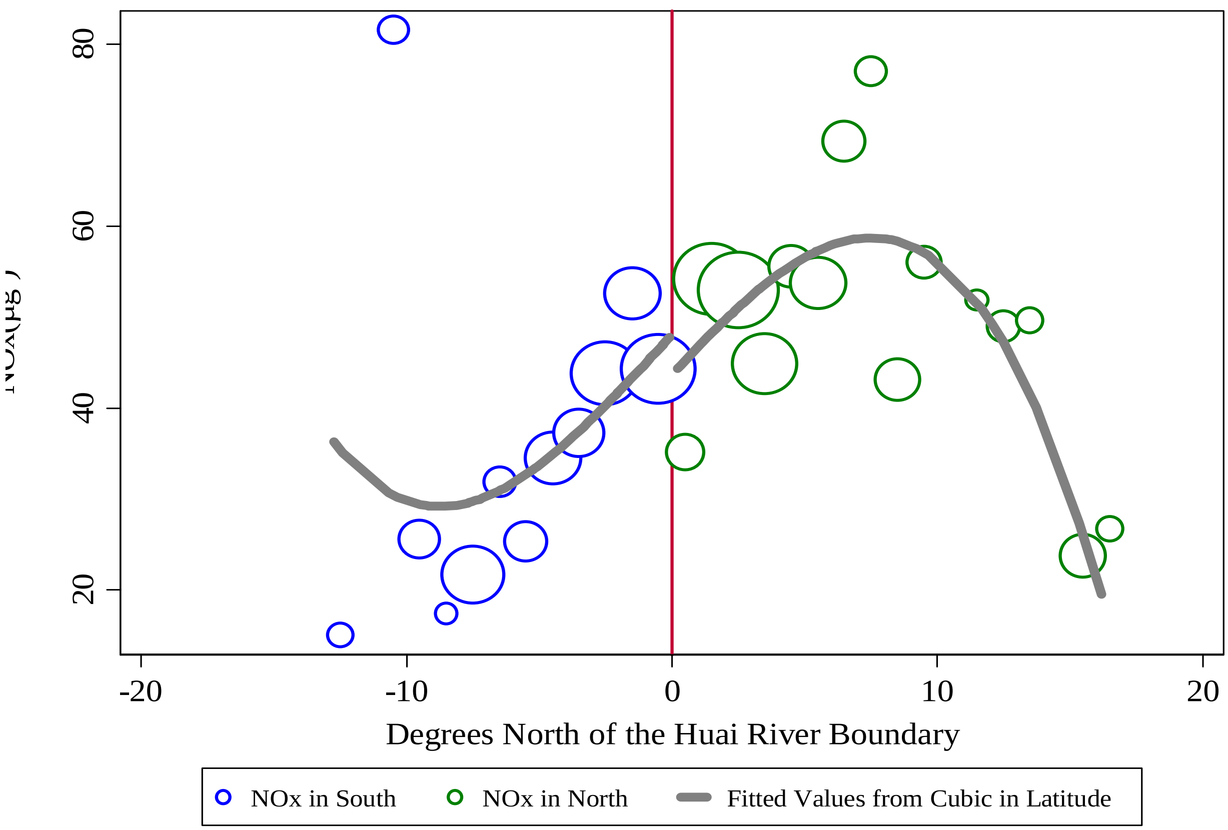
Sulphur Dioxide Concentration Exposure among Chinese Disease Surveillnace Points by Distance from Huai River Boundary



Note : Each observation is generated by averaging SO2 across the Disease Surveillance Point locations within a 1^0 latitude range. The size of the point is in proportion to the number of sites within the 1^0 latitude range. The plotted line reports the fitted values from a regression of S02 on a cubic polynomial in latitude using the sample of DSP locations. See the main text for a discussion of these results.

Figure S2

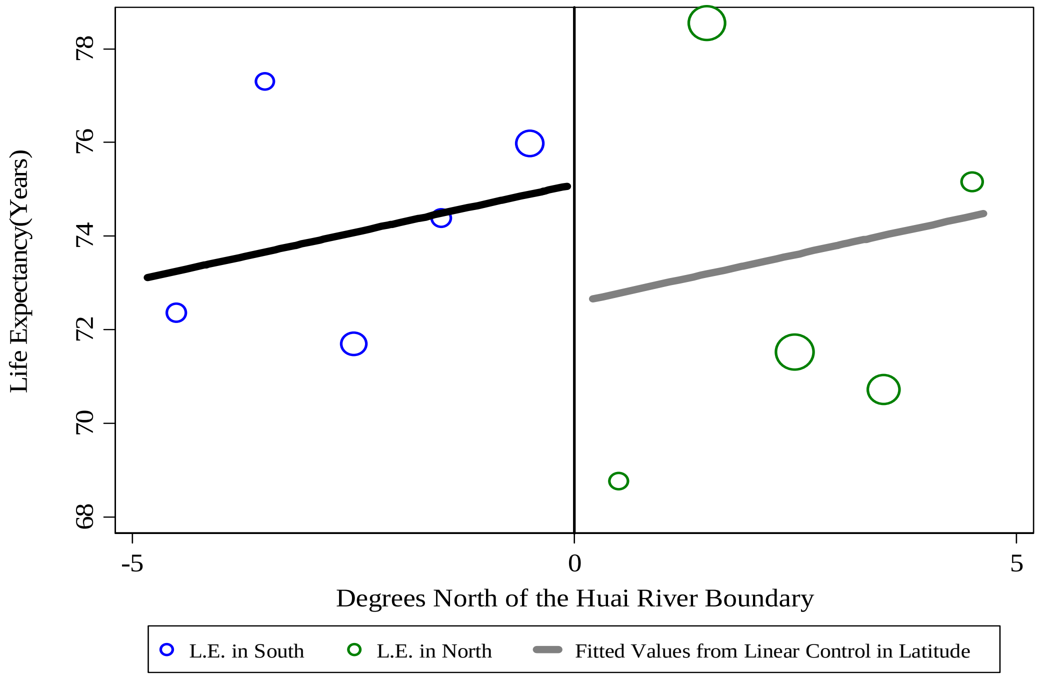
Nitrous Oxide Concentration Exposure among Chinese Disease Surveillance Points by Distance from Huai River Boundary



Note : Each observation is generated by averaging NOx across the Disease Surveillance Point locations within a 1^0 latitude range. The size of the point is in proportion to the number of sites within the 1^0 latitude range. The plotted line reports the fitted values from a regression of S02 on a cubic polynomial in latitude using the sample of DSP locations. See the main text for a discussion of these results.

Figure S3

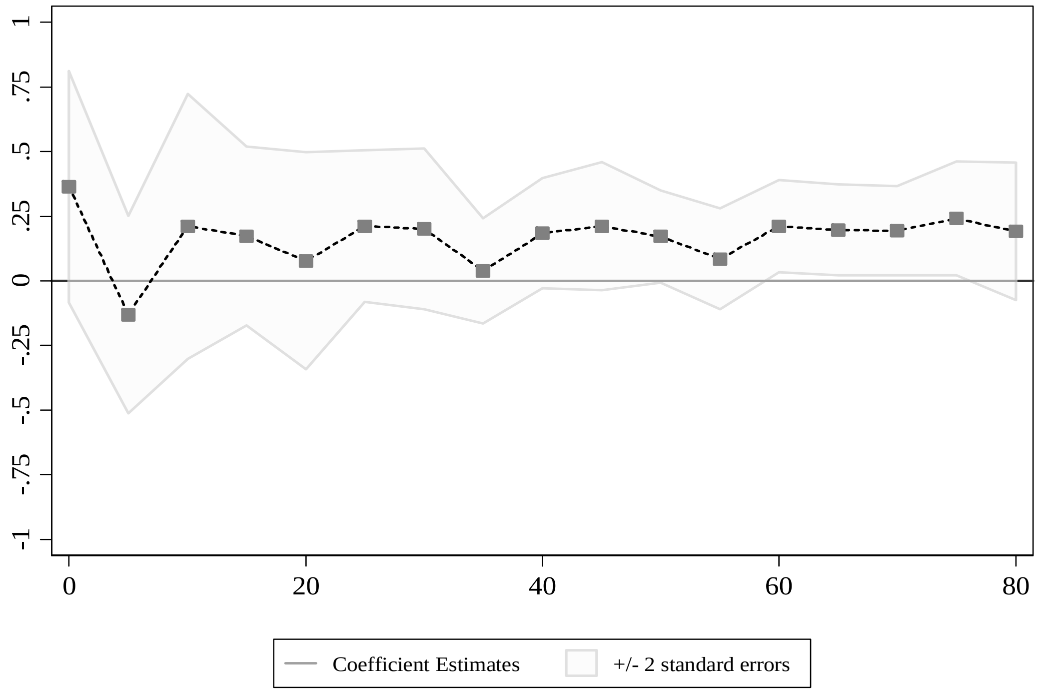
Life Expectancy among Chinese Disease Surveillance Points by Distance from Huai River Boundary in Sample within 5^0 Latitude



Note : Each observation is generated by averaging predicted life expectancy across the Disease Surveillance Point locations within a 1^0 latitude range. The size of the point is in proportion to the number of sites within the 1^0 latitude range. The plotted line reports the fitted values from a regression of life expectancy on a linear term in latitude using the sample of DSP locations.

Figure S4

Estimated Impact of Long-Term TSP exposure (1981-2000) on Cardiorespiratory Mortality by Age using the Huai River Policy (Two-stage Least Squares)



Note : Each point in the plot represents the coefficient from a separate regression, where we examine how the log of age-specific death rates are affected by long-term exposure to TSP (μg/m^3). The IV regressions are estimated in the manner described in Table 3, where the log of the death rate from cardiorespiratory mortality causes is the dependent variable and long-term TSP average is the independent variable. The models are estimated using 1(North) as the instrument for average TSP.