Effects of air quality on housing prices: evidence from China’s Huai River policy

Xinghua Liu, Qiang Li, Satish Chand

School of Business, University of New South Wales (Canberra)
© Copyright 2018 by Authors’ names.

All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Effects of air quality on housing prices: evidence from China’s Huai River policy

Abstract: Estimating the economic value of clean air is of significance to both policymakers and private individuals but its quantification has proved difficult. Of the different valuation approaches used, the classic hedonic theory predicts a negative relationship between air quality and housing prices. Existing attempts to quantify this nexus is plagued by problems of endogeneity, mainly arising from omitted variables that confound air pollution with other determinants of housing prices. We employ a regression discontinuity (RD) design to estimate the impact of air pollution on house prices across a river that demarcates regions with and without coal-fired heating emanating from the Huai River Policy. This policy was decreed by the Chinese government in the 1950s that allowed burning of coal at subsidised prices for indoor heating to only the north of the Huai River. Employing quasi-experimental variation in particulate matter of 10 micrometres or less in aerodynamic diameter (PM10) generated by this arbitrary policy and regression discontinuity (RD) design based on distance from Huai River, we estimate the local average treatment effect (LATE) to provide new evidence on the capitalization of PM10 air pollution into housing values. By using panel data of 30 large cities on either side of the river for the period 2006 to 2015, we found that 1 μg/m^3 reduction in average PM10 results in an approximately 1 percent increase in housing prices. The results are robust to using parametric and nonparametric estimation methods and adjustment to a rich set of covariates.

Key words: air quality; housing prices; PM10; Huai River Policy; regression discontinuity design; China

I. Introduction

It is widely acknowledged that air pollution has substantial and wide-ranging adverse effects on human health (Brunekreef and Holgate, 2002, Pope III and Dockery, 2006). Many developed nations, in response, have implemented air quality regulations such as the ‘Clean Air Act’ in the US and the UK. Similar actions by developing nations however are absent even when a few such as China have recently been suffering from worsening air quality. As such, the Chinese government and its people have turned their attention on improving air quality. The Chinese government declared a “war on pollution” in 2014, announcing a national emissions trading scheme aimed at using market forces to cut greenhouse gas emissions (Jotzo et al., 2018). Although these regulations enacted by different countries seem to have an effect on improving air quality, the costs and benefits of clean air on governments and individuals remain largely unclear. Therefore, the assessment of the economic costs and benefits of the clean air is of increasing significance for economists, individual citizens, and policymakers.

The economic value of clean air can be estimated using hedonic method, whose main idea is to use private market to infer the implicit price function for public goods. The most commonly used private market is the real estate market. In this case, information on demand for public goods is then capitalised into the prices of housing. Nevertheless, estimating the housing prices-air quality nexus using hedonic housing approach is severely plagued by the issues of endogeneity. First is the omitted variables bias, i.e., other determinants of housing prices may also be correlated with our explanatory variable of interest--air quality. Second, measurement error of
the air quality indicators may also threaten the conventional approach. As a consequence, previous evidences suggest that estimations using hedonic method tend to show weak results and largely underestimate the true effects of air pollution on housing prices (Bayer et al., 2009; Smith and Huang, 1995).

Although there is a small body of literature investigating the housing price response to air quality using particular identification strategies to specifically address problems of endogeneity, the evidence is mainly drawn from the experience of developed economies. For example, Chay and Greenstone (2005) exploit the reductions in air pollution induced by the Clean Air Act Amendments as a quasi-experiment and use the instrumental variable estimator to examine the effects of TSP\(^1\) on housing prices in the U.S. Their findings indicate that individuals place higher value on clean air than the estimates presented by the previous literature.

Our study is relevant to two strands of this literature. The first strand is the body of literature that values air quality using various approaches. Besides the hedonic approach mentioned above, another way to estimate the economic value of clean air is the life satisfaction arising from non-market valuation approach. For example, Luechinger (2009) use the IV approach exploiting the natural experiment created by the mandated scrubber installation at power plants, with wind directions dividing counties into treatment and control groups. And the results suggest that air pollution has a negative effect on welling-being, and again, the effect is larger for IV than conventional estimates. Similarly, Ambrey et al. (2014) employs the life satisfaction approach to estimate the cost of PM\(_{10}\) exceedances in south east Queensland and found a negative relationship between air pollution and life satisfaction.

Another strand of study related to our analysis looks at how (dis)amenities are capitalised into housing prices. There are a number of existing studies investigating the impact of environmental externalities on housing price including air quality (Chay and Greenstone, 2005), water quality (Leggett and Bockstael, 2000), urban green space (Jim and Chen, 2006), etc. Apart from this, the relationship between school quality and housing prices has been widely investigated upon (Kane et al., 2006). Econometric identification issues, again, often plague this area of study. Black (1999) is an excellent example in addressing the potential for reverse causality between the variables included in the model. His analysis, by identifying houses located on attendance district boundary, suggests a 5 percent improvement in test scores leads to a 2.5 percent increase in housing prices. In terms of the dis-amenities, a growing number of studies look at the impact of crime and violence on housing prices. A recent example is Besley and Mueller (2012), which estimates the peace dividend as capitalised into housing price changes in North Ireland.

China’s air quality is notoriously poor. According to World Health Organisation (WHO), China is the world’s deadliest country for outdoor air pollution: more than 1 million people died from dirty air in China in 2012\(^2\). China daily also reports that nearly 90 percent of China’s large cities failed to meet air quality standards in 2014\(^3\). Thus, providing estimates on the economic value of clean air at pollution levels far exceeding those ever recorded in any other country has policy significance. The Huai River Policy generates an exogenous variation in air quality within China that can be used to investigate individuals’ valuations of clean air (Almond et al.

---

1. Total suspended particles (TSP) measures the mass concentration of particulate matter in the air. Within TSP, PM\(_{10}\) stands for particles with a diameter of 10 μm or less, and PM\(_{2.5}\) stands for those with a diameter of 2.5 μm or less. Particulates that are 10 μm or greater are filtered and generally do not enter the lungs. Particulates smaller than 10 μm are likely to enter the lungs. Particulate matter that is smaller than 2.5 μm (PM\(_{2.5}\)) can enter into the Alveoli where gas exchange occurs. Throughout the world, ambient monitoring now focuses on PM\(_{10}\) and PM\(_{2.5}\).
2. https://www.theguardian.com/environment/2016/sep/27/more-than-million-died-due-air-pollution-china-one-year#img-1
The goal of this paper is to estimate the economic value of clean air and empirically test whether the improvement of air quality leads to a rise in housing prices. Specifically, we employ a regression discontinuity (RD) design to estimate the impact of air pollution on housing prices across a river that demarcates regions with and without coal-fired heating emanating from the Huai River Policy. This policy was decreed by the Chinese government in the 1950s that allowed for free to heavily subsidised coal for indoor heating only to the north of the Huai River. One consequence of this policy is the sustained differences in the exposed PM$_{10}$ levels across the river, noting the fact that PM$_{10}$ being heavy does not get blown across space and its effects are visible. Both sharp and fuzzy RD design is employed in the empirical analysis based on the discrete difference in air quality around the boundary.

The rest of the paper is organised as follows. The next section discusses the institutional background. Section 3 describes the data and summary statistics. Section 4 explains identification and empirical strategy. Section 5 presents the empirical results and section 6 concludes.

II. Background

The home heating policy is a product of the Chinese central planning regime during the 1950s to 1980s. During this period, winter heating was considered as a basic right and free heating was provided for homes and offices by the government via offering free or substantially subsidised coal fuel. One legacy of this system is that many homes still receive free to highly subsidised heating to this day. However, budget restrictions imply that the provision of free heating to every household is unrealistic in winter is unrealistic. Therefore, residents of only areas located in northern part of China were entitled to this welfare support. The border that used to divide northern and southern China (China’s north-south divide) is the Huai River/Qinling Mountain Range. The line is also overlapped with the 0$^\circ$ January Celsius of the average temperature.

The Huai River/Qinling Mountain Range is ambiguous due to the significant change of the geographical conditions of Huai River basin over the past few centuries. The Huai River entered the Yellow Sea directly before the Northern Song Dynasty. Nevertheless, due to the Yellow River’s repeated change of its course southwards, it met the Huai River, resulting in the substantial change of the geography of the Huai River basin. This ambiguity of the river leads to difficulty in drawing an accurate line. Following the method adopted by several studies (Chen et al., 2013, Ebenstein et al., 2017, Almond et al., 2009), we draw the Huai River line first based on its main waterway$^4$. Cities to the north of this line are covered by the policy of free winter heating provided from November 15 to March 15. Those to the south, in contrast, lack centralised heating infrastructure and there are no private providers of winter heating to these areas until now.

It is worth noting that the Chinese heating system is predominantly coal-fired, and pollution-intensive partly because of dated technology. Compared to heating generated using electricity, gas and oil heating systems that prevail in most developed countries, coal fired boilers are inefficient and environmentally unfriendly. the norm is for a boiler located in each residential building that generates heating for the whole complex. The function of home heating is that the water is heated by burning coal in the boiler and then sent through iron pipes to each household in that building. Typically, the water has to travel long distances before arriving at the destination. As

---

$^4$ It originates in Tongbai Mountain in Henan Province, flows through Henan, Anhui and Jiangsu Province, and finally enters the Yangtze River at Yangzhou in Jiangsu Province.
such, the energy loss is common through this heating process. This process is also accompanied with the
incomplete combustion of coal, resulting in the emissions of the several toxic pollutants, including mercury,
sulfur dioxide (SO₂), nitrogen oxides (NOₓ) and particulate matter (Chen et al., 2005).

This study takes advantage of the stark differences in the air quality across Chinese cities generated by this
arbitrary policy, and uses this variation to estimate the impact of this environmental dis-amenity on housing
prices. This is done is two steps. First, we estimate the impact of Huai River Policy on air pollution and housing
prices separately to test any discontinuity across the river. Then, we examine the effects of air quality on housing
prices using fuzzy regression discontinuity (fuzzy RD) approach.

III. Data sources and summary statistics

A. Data sources

To implement the analysis, we compiled a comprehensive panel dataset that provide city-specific information on
pollution levels, city location relative to Huai River and housing prices for 30 cities between the year 2006 and
2015. The unit of observation in our data is the city. The practical reason for the use of city level data is that the
published housing price and air pollution data are all at city level, thereby enabling us to match them only at this
level. The housing prices data come from China Real Estate Statistical Yearbook (2006-2015). The yearbook is
compiled by National Bureau of Statistics of China (NBS) and China Index Academy. It reports economics data
of real estate all over the country and 35 large and medium cities in China for each year. In terms of housing
prices, it reports average prices for commercialised housing as well the prices for its four components:
residential housing, office building, commercial housing and housing for other purpose. We focus only on the
residential housing for our research purpose for two reasons. First, residential housing is considered the most
important part of the housing market in urban China. Second, air quality is more likely to be capitalised into
prices of residential housing, since this share of housing is where people choose to live for their lifetime.

The air pollution data were obtained from the China Environment yearbook compiled by the Ministry
of Environmental Protection (MEP). The data files contain annual information on the ambient air quality in 31
major cities and include measures of several pollutants. These air quality measures include concentration of
particulate matters, sulphur dioxide (SO₂), nitrogen dioxide (NO₂), as well as the days of air quality equal to or
above Grade II and proportion of days of air quality equal to or above Grade II in the whole year.

We focus on particulate matter smaller than 10 µm in diameter (PM_{10}) pollutant for two reasons. First, PM_{10},
compared to other forms of air pollutants (PM_{2.5}, SO₂, NO₂) is heavier and therefore is less likely to travel long
distances which can contaminate the measurement of air pollution across the river. Second, an increasing body
of evidence suggests that PM_{10} is by far the most important local air pollutant in terms of health effects (Bayer et
al., 2009). PM_{10} was first tracked by the Chinese monitoring system in 2003 and therefore our analysis draws on
the readings from 2003 to 2015.

Besides the outcome variable and the explanatory variable of our interest, to adjust the estimates in the
subsequent statistical analysis, we also collected a range of potential confounding factors that are used as control
variables. We obtained the city level data related to both housing supply and demand which in turn will exert an

---

5 We exclude Lhasa (capital city of Tibet) in our sample and restrict the sample to the year 2006-2015 due to availability of certain
covariates.
effect on local housing prices of each city. This series of variables were compiled from the China Statistical Yearbook, China Real Estate Statistical Yearbook and online database of NBS. These city-level variables include: per capita floor space of buildings started each year by enterprises for real estate development, per capita real GDP, real wage, population density, fiscal surplus, dummy for coastal city, number of students enrolled at higher education and number of hospitals per capita.

B. Summary statistics

Figure 1 presents trends from 2006-2015 in the average annual ambient concentration of PM$_{10}$ and average housing prices of cities north or south of the Huai River separately. As revealed by panel A in Figure 1, air quality in cities south of the river is continuously better than those to the north despite their parallel trend. Further, to better display the spatial difference in PM$_{10}$ and housing prices in cities north and south of the Huai River, these data are drawn in the provincial map of China as shown in Figure 2. Noticeably, cities in the north are generally exposed to substantially higher level of PM$_{10}$ concentrations than those in the south. This is consistent with the claim that air pollution is more severe in northern Chinese cities. However, the figure fails to reveal some of the systematic patterns of the spatial distribution of housing price across the river line.

Panel A. Changes of PM$_{10}$ over time

Panel B. Changes of housing price over time

Figure 1. Changes of PM$_{10}$ in northern and southern Chinese cities

Notes: The line in blue represents the southern Chinese cities while the line in orange denotes the northern Chinese cities. The unit for PM$_{10}$ and housing prices is µg/m$^2$ and yuan/m$^2$, respectively. Data on PM$_{10}$ are from China Environment yearbook and data on housing prices are from China Real Estate Statistical Yearbook.
Figure 2. The comparison of PM10 and housing prices in northern and southern Chinese cities

Notes: The PM10 levels and housing prices in each city are shown in the provincial map of China. The 30 cities are matched with the provinces in which they are located except for Tibet due to unavailability of data. The unit for PM10 and housing prices is µg/m³ and Yuan, respectively. Data on PM10 are from China Environment yearbook and data on housing prices are from China Real Estate Statistical Yearbook.

Table 1 reports summary statistics for the key variables in this analysis. China’s mean value of PM10 concentrations was 102.99 µg/m³ over the period 2006-2015. For the purpose of comparison, we calculate the US PM10 concentrations for the same period based on data reported by United States Environmental Protection Agency (EPA). The figure is 68.58 µg/m³ for the US, making China’s PM10 concentrations more than 1.5 times US levels.

Table 1. Summary statistics for key variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM10 (µg/m³)</td>
<td>300</td>
<td>102.99</td>
<td>31.14</td>
<td>30</td>
<td>305</td>
</tr>
<tr>
<td>Real housing prices(yuan/m²)</td>
<td>300</td>
<td>5916.31</td>
<td>4243.11</td>
<td>1541</td>
<td>49659.82</td>
</tr>
<tr>
<td>Per capita real GDP(10,000 yuan)</td>
<td>300</td>
<td>6.537922</td>
<td>5.270577</td>
<td>1.221441</td>
<td>70.41043</td>
</tr>
<tr>
<td>Real wage(yuan)</td>
<td>297</td>
<td>45112.18</td>
<td>27029.21</td>
<td>16911</td>
<td>398203.3</td>
</tr>
<tr>
<td>Fiscal surplus (100 million yuan)</td>
<td>299</td>
<td>-152.18</td>
<td>248.85</td>
<td>-2315.93</td>
<td>601.83</td>
</tr>
<tr>
<td>Population density (people/km²)</td>
<td>299</td>
<td>1617.66</td>
<td>1373.97</td>
<td>223.31</td>
<td>11449.3</td>
</tr>
<tr>
<td>Per capita floor space (m²)</td>
<td>300</td>
<td>8.06</td>
<td>4.32</td>
<td>0.97</td>
<td>24.81</td>
</tr>
<tr>
<td>1 (Coastal city)</td>
<td>300</td>
<td>0.13</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
IV. Econometric model

We use both the conventional and the RD approach to estimate the association between PM$_{10}$ and housing prices. The basic idea and structure of this section follows closely to that of Chen et al. (2013) and Ebenstein et al. (2017). The conventional strategy uses the panel data fixed effect to estimate the following equation:

$$\text{HP}_{it} = \beta_0 + \beta_1 \text{PM}_{it} + X_{it}\Gamma + \epsilon_{it}$$ \hspace{1cm} [1]$$

Where $i$ denotes a city and $t$ indexes a year. $\text{PM}_{it}$ is the PM$_{10}$ concentration in city $i$ in year $t$, $X_{it}$ is a vector of observed covariates used to control other determinants of housing prices other than air quality and $\epsilon_{it}$ is an error term. The outcome variable is the natural logarithm of the average residential housing prices in city $i$ in year $t$. $\beta_1$ is the coefficient of interest for Eq. [1].

Undoubtedly, this estimation of Eq. [1] is plagued by the endogeneity problem and therefore it is difficult to get a consistent $\beta_1$. First is the omitted variables bias. The validity of the conventional approach depends on the assumption that the adjustment of covariates rules out all the confounding factors. Although we have controlled for a set of observed characteristics that are available, unobserved determinants of housing prices may also covary with $\text{PM}_{it}$. Admittedly, compared to cross sectional analysis, omitted variable bias is of lesser concern for panel data analysis because the panel structure of the data allows us to eliminate the effects of any time invariant unobservables by inclusion of city fixed effects. Besides this, measurement error of pollution concentrations is another endogeneity concern, since there is evidence that air quality readings tend to be underreported by the officials. Consequently, the conventional estimates of the nexus between air quality and housing prices may be biased and unable to reveal the causal linkages because of these two econometric identification issues.

The second approach leverages the regression discontinuity (RD) design where we exploit the exogenous variation caused by Huai River policy to estimate its impact on both PM$_{10}$ concentrations and housing prices. The development of RD identification is based on the idea that the treatment status is determined by a continuous variable crossing a cut off due to imposition of some arbitrary rule. This feature, thus, provides good experiments and a local randomization for estimating treatment effect. The RD design has emerged almost simultaneously in a wide range of disciplines but has only recently become important in applied econometrics. In this study we exploit a RD design that has a break in spatial dimension where a treatment is applied along a geographical boundary. Specifically, the RD design used in this research is based on the discrete increase in the provision of free to subsidized coal for indoor heating north of the Huai River. This is done in two steps.

First, we use sharp regression discontinuity (sharp RD) design to test any discontinuity in PM$_{10}$ and housing prices across the river. This approach assumes the Huai River policy as a treatment variable and estimate its impact on the two endogenous variables. In this case, the treatment status is a deterministic and discontinuous function of the running variable (the degree distance north of the Huai River boundary), making the estimates a sharp RD design. Second, the study exploits a fuzzy regression discontinuity (fuzzy RD) design to estimate the impact of PM$_{10}$ on housing prices. A fuzzy RD is where the treatment status is no longer deterministically
related to the threshold-crossing rule. In our text, the PM$_{10}$ pollutant exists on both sides of the river but the probability of PM$_{10}$ exposure increases considerably across the river, making our study a fuzzy RD in nature. Consequently, the discontinuity becomes an instrumental variable for PM$_{10}$. It is worth noting that the validity of our RD design rests on the assumption that any unobserved determinants of PM$_{10}$ and housing prices change smoothly as they cross the river line. If this identification assumption is true, adjustment for flexible polynomial terms in degree distance north of the river or local linear regressions (LLR) will eliminate all sources of bias and hence allow us to draw casual inference.

Based on the sharp RD of the second approach, the following parametric equations are estimated to test the effect of Huai River policy on PM$_{10}$ concentrations and housing prices respectively:

$$\begin{align*}
PM_{it} &= \alpha_0 + \alpha_1 N_i + f(D_i) + N_i f(D_i) + X_{it} \gamma + u_{it} \quad [2a] \\
HP_{it} &= \delta_0 + \delta_1 N_i + f(D_i) + N_i f(D_i) + X_{it} \varphi + \varepsilon_{it} \quad [2b]
\end{align*}$$

Where $i$, and $t$ references city and year, respectively. PM$_{it}$ represents the PM$_{10}$ concentration in city $i$ in year $t$. HP$_{it}$ is the log term of the average residential housing prices in city $i$ in year $t$. $N_i$ is a dummy variable with 1 indicating city is located in the north of the Huai River and 0 otherwise. $f(D_i)$ is a polynomial term in degrees north of the Huai River and $N_i f(D_i)$ is its interaction with north Huai River dummy. $X_{it}$ is a vector of observed covariates.

An alternative estimation strategy for the sharp RD involves nonparametric estimate of Eq. [2a] and [2b]. The nonparametric approach of RD requires good estimates of the mean of outcome variable in small neighbourhoods to the right and left of the cut-off. Local linear regression (LLR) is a commonly used nonparametric version of regression to obtain such estimates(Hahn et al., 2001). The LLR setup is as follows:

$$\begin{align*}
PM_{it} &= \alpha_0 + \alpha_1 N_i + \alpha_2 D_i + \alpha_3 N_i D_i + u_{it} \quad [3a] \\
HP_{it} &= \delta_0 + \delta_1 N_i + \delta_2 D_i + \delta_3 N_i D_i + \varepsilon_{it} \quad [3b]
\end{align*}$$

The LLR uses restricted samples within $h$ latitude degrees of the Huai River ($|D| \leq h$). A fully nonparametric approach requires data-driven rules for selection of the width of the sample window, also known as “bandwidth”. Throughout this paper, $h$ is chosen according to bandwidth selection method proposed by Imbens and Kalyanaraman (2012).

Alternatively, we use what is called a "discontinuity sample" by (Angrist and Lavy, 1999) to implement the local linear methods. Specifically, we also report results from the parametric RD approach described in Eq. [2a] and [2b] that limits the sample to cities within 7° latitude of the Huai River. This sample restriction provides us an informal way of estimating the local average treatment effect (LATE) that use bandwidths and kernels to focus on observations near the cut-off value.

As mentioned earlier, we use Fuzzy RD to estimate the impact of PM$_{10}$ on housing prices. Fuzzy RD leads naturally to an instrumental variable or two-stage least squares (2SLS) estimation strategy. Obviously, the dummy variable $N$ is the instrument variable exploited. Therefore, the two-stage regression specifications are as follows:

$$\begin{align*}
PM_{it} &= \alpha_0 + \alpha_1 N_i + f(D_i) + N_i f(D_i) + X_{it} \gamma + u_{it} \quad [2a] \quad \text{first stage} \\
HP_{it} &= \beta_0 + \beta_1 PM_{it} + f(D_i) + N_i f(D_i) + X_{it} \varphi + \varepsilon_{it} \quad [2c] \quad \text{second stage}
\end{align*}$$
where $\text{PM}_{10}$ represents the fitted values from estimating Eq. [2a]. The other variables are defined as above.

Compared to the conventional approach described in Eq. [1], the Fuzzy RD design addresses the endogeneity concerns arising from omitted variables and the measurement error of air quality. In relation to estimates based on Eq. [2a], [2b], [3a], [3b], 2SLS is advantageous in terms of external validity since it estimates the impact of air quality, rather than the Huai River policy, on housing price; thus the empirical results can be applied for other settings.

Hahn et al. (2001) develop a nonparametric IV procedure using local linear regression to estimate the top and bottom of the Wald estimator with less bias. Wald estimator for fuzzy RD captures the causal effect on compliers defined as unit of observations whose treatment status changes as we move just below or just above the threshold (Angrist and Pischke, 2008). The IV Wald estimator is the ratio of two sharp discontinuities, namely the ratio of the estimated discontinuity in housing prices to the estimated discontinuity in $\text{PM}_{10}$, with both the denominator and the nominator estimated by LLR.

V. Empirical Results

Table 2 reports the summary statistics for several key determinants ($\text{PM}_{10}$ and control variables) of housing prices and lends support to the validity of our RD design. The first two columns report the means and the standard derivations in cities located north or south of the Huai River. Column [3] presents the mean difference between the south and the north along with the standard errors. The table shows that $\text{PM}_{10}$ exposure differs significantly among north and south cities. Concentrations of sulphur dioxide ($\text{SO}_2$) and nitrogen dioxide ($\text{NO}_2$) are also reported in the China Environmental Yearbook. Indeed, Table 2 demonstrates that the differences among these pollutants are also statistically significant among southern and northern cities. Due to the fact that sulphur dioxide and nitrogen dioxide are lighter and can travel in greater distance than $\text{PM}_{10}$, we focus on $\text{PM}_{10}$ to distinguish it from other air pollutants as a potential environmental dis-amenity on housing prices caused by living north of the river.

<table>
<thead>
<tr>
<th>Variable</th>
<th>North</th>
<th>South</th>
<th>Difference in means</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{PM}_{10}$ ($\mu g/\text{m}^3$)</td>
<td>119.4133</td>
<td>86.56</td>
<td>-32.8533*** (3.058324)</td>
</tr>
<tr>
<td>$\text{SO}_2$ (mg/m$^3$)</td>
<td>0.0508667</td>
<td>0.039867</td>
<td>-0.01688** (0.0020688)</td>
</tr>
<tr>
<td>$\text{NO}_2$ (mg/m$^3$)</td>
<td>0.04354</td>
<td>0.0416533</td>
<td>-0.00018867 (0.0013518)</td>
</tr>
<tr>
<td>Real_hprice(yuan/$\text{m}^2$)</td>
<td>5442.61</td>
<td>6390.02</td>
<td>-947.41** (487.70)</td>
</tr>
<tr>
<td>Log(real_hprice)</td>
<td>8.43</td>
<td>8.62</td>
<td>0.19*** (0.06)</td>
</tr>
<tr>
<td>Per capita real GDP(10,000 yuan)</td>
<td>6.449367</td>
<td>6.626477</td>
<td>0.17711 (0.6095278)</td>
</tr>
<tr>
<td>Real wage(yuan)</td>
<td>45047.3</td>
<td>45176.63</td>
<td>129.3319 (3142.108)</td>
</tr>
<tr>
<td>Fiscal surplus (100 million yuan)</td>
<td>-148.8867</td>
<td>-155.4967</td>
<td>-6.610044 (28.82844)</td>
</tr>
<tr>
<td>Population density(people/km$^2$)</td>
<td>1652.998</td>
<td>1582.079</td>
<td>-70.91001 (159.132)</td>
</tr>
<tr>
<td>Per capita floor space (m$^2$)</td>
<td>7.937815</td>
<td>8.188631</td>
<td>0.2508162 (0.4990873)</td>
</tr>
<tr>
<td>Students enrolled at higher education</td>
<td>379014.2</td>
<td>488100.1</td>
<td>109085.9*** (25285.29)</td>
</tr>
<tr>
<td>No. of hospitals owned per 10,000 people</td>
<td>0.598874</td>
<td>0.4537101</td>
<td>-0.1451639*** (0.0477961)</td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level
A. Estimates of the effect of the Huai River Policy on PM$_{10}$ and housing prices

An appealing feature of any RD design is that it can be illustrated visually. First, we graphically show the effect of the Huai River Policy on air pollution. Figure 3 shows PM$_{10}$ at various cities against their degree distance north of the river. The dots in Figure 3 represent the mean of PM$_{10}$ concentrations falling within each bin. The fitted values represented by the solid line are from a kernel weighted 2nd-order polynomial regression fit of the PM$_{10}$ on the score. All control and treated observations are included in the control and treated fit, respectively. The plot reveals a positive jump at the cut-off: the PM$_{10}$ concentration seems to be higher in cities north of the river than those to the south. In contrast, the plot in Figure 4 shows a negative jump at the border, making the plotted line in Figure 4 almost a mirror image of that in Figure 3.

![Figure 3. Fitted value from regression of PM$_{10}$ exposure on distance from the Huai River](image)

Besides the graphical evidence, Table 3 reports the empirical results from estimating Eq. [2a] and [2b]. All the estimates are adjusted for the full set of covariates reported in the data sources section. Column [1] reports the fixed effect estimates of the coefficient on a dummy variable indicating whether located to the north of the river after controlling for a third-order polynomial term in distance from the river as well as its interaction with the north river dummy using the full sample. Column [2] and Column [3] apply the
parametric and non-parametric RD approach to estimate these two equations. Column [2] use the subsample within 7° latitude of the Huai River which was imposed manually as our discontinuity sample. On the contrary, in Column [3], the estimates are based on the local linear regression with a triangular kernel and optimal bandwidth (h) selected by the method prescribed by Imbens and Kalyanaraman (2012).

Overall, Table 3 presents statistical evidence of considerable rise in PM10 and decrease in housing price at the Huai River, which is similar to what we have presented in Figure 3 and Figure 4. At the threshold, the PM10 increase by 19 μg/m³ as indicated by conventional method, whereas in the RD design, the increase is substantially larger, by 41 to 47 μg/m³, respectively. In comparison, the log term of housing prices decline as we move across the border by 27% and 42% for the conventional and parametric RD approach. The estimates using nonparametric approach display coefficient of slightly larger magnitude: the estimated decline in log (housing prices) of cities north of the river is 43%.

Table 3. RD estimates of the impact of the Huai River Policy

<table>
<thead>
<tr>
<th>Outcome</th>
<th>[1]</th>
<th>[2]</th>
<th>[3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM10</td>
<td>19.1157***</td>
<td>41.251***</td>
<td>47.47836*</td>
</tr>
<tr>
<td></td>
<td>(8.64)</td>
<td>(7.7153)</td>
<td>(28.10 87)</td>
</tr>
<tr>
<td>Housing price(log)</td>
<td>-0.272**</td>
<td>-0.42026***</td>
<td>-0.4306474***</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.05896 )</td>
<td>(0.1575027)</td>
</tr>
<tr>
<td>RD type</td>
<td>polynomial</td>
<td>polynomial</td>
<td>Local Linear Regression(LLR)</td>
</tr>
<tr>
<td>Polynomial function</td>
<td>Third</td>
<td>Linear</td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>All</td>
<td>7°</td>
<td></td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

B. Estimates of the effect of PM10 on housing prices

Table 4 reports the estimated effect of PM10 on housing prices based on a variety of approaches. Column [1] reports the results from application of the conventional fixed effect method, which acted as a baseline model to compare the estimates with those based on RD and 2SLS approaches. Specifically, the fixed effect estimator of the coefficient of PM10 exposure on housing prices is statistically insignificant and marginal in size: a 1 unit increase in PM10 is correlated with a 0.1% increase in housing price, but as noted above the parameter estimate is insignificantly different from zero. Additionally, fuzzy RD using restricted sample based on the data-driven approach and bandwidth selected manually are presented in Column [2] and Column [3], respectively. They are statistically significant at 10% significant level. Overall, the results indicate that the fuzzy RD approach suggests a larger coefficient in magnitude compared to conventional estimates. This lends support to the fact that omitted variables bias plagues the conventional approach and tend to underestimate the true value of clean air.

Table 4. Comparing fixed effect and RD estimates of PM10’s impact on housing prices

<table>
<thead>
<tr>
<th>Outcome</th>
<th>[1]</th>
<th>[2]</th>
<th>[3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing price(log)</td>
<td>-0.001</td>
<td>-0.0101651*</td>
<td>-0.0083*</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0058542)</td>
<td>(0.00467)</td>
</tr>
<tr>
<td>Estimation method</td>
<td>Fixed effect</td>
<td>IV</td>
<td>IV</td>
</tr>
<tr>
<td>RD Type</td>
<td></td>
<td>Local Linear Regression(LLR)</td>
<td>Local Linear Regression(LLR)</td>
</tr>
<tr>
<td>Sample</td>
<td>All</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 1 percent level.
VI. Robustness

We now present results for several robustness checks to assess the significance and robustness of our findings and the validity of our RD design. In this section, we examine the robustness of our results by (i) using fake cut-off lines; and (ii) testing whether our results were affected qualitatively by the decisions made in our paper along several dimensions, such as adjustment for covariates, sample selection, functional forms of our models, and kernel weighting methods.

First, to test the robustness of our results, we repeat the estimation of fuzzy RD using the alternative cut-off lines. While we know the Huai River line is the true cut-off of interest due to the indoor heating policy in the 1950s, we assume other lines around the Huai River as the cut-offs for the placebo tests. We redraw a window of four displacements at 0.5°-latitude intervals north and south of the Huai River across China as well as at the actual Huai River (which is equivalent to the 0° displacement) and re-estimate our key coefficient of interest. The results from these falsification exercises are shown in Figure 5. According to Figure 5, the coefficient is only statistically significant at the actual Huai River. All other instances fail to yield significant coefficients for housing price changes, suggesting that the previous findings reported in this paper are a result of the Huai River policy rather than an artifact of this application of the RD approach.

Figure 5. Placebo test of RD design on the impact of PM10 on housing prices

In addition, we examine the robustness of our results based on tests from several dimensions, including adjustment for covariates, sample selection, functional forms of our models, and kernel weighting methods to confirm whether our results are qualitatively the same. In Table 5, we expand the results of Table 4 by including

** Significant at the 5 percent level.
* Significant at the 10 percent level.
a full set of parametric estimates with and without adjusting for the covariates, and local linear estimates using a variety of kernel weighting methods. The entries in columns (1) and (4) are identical to those in Table 4, which are adjusted using the full set of covariates and a linear polynomial in distance to the Huai River. Columns (3), (5) and (7) report the estimated coefficients excluding the covariates. Using a triangular kernel, the entries in column (4) match the local linear results in Table 4, while estimates using an epanechnikov and uniform kernel are presented in columns (5) - (8), respectively.

Table 5. Comparing the fixed effect and regression discontinuity estimates of PM$_{10}$'s impact on housing prices

<table>
<thead>
<tr>
<th>Housing prices</th>
<th>Fixed effect (within 7 degrees)</th>
<th>Fixed effect (within 7 degrees)</th>
<th>2SLS Polynomial RD (within 7 degrees)</th>
<th>Local Linear Regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Housing prices</td>
<td>-0.001</td>
<td>-0.0011</td>
<td>-0.0197***</td>
<td>-0.0204***</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0008)</td>
<td>(0.00557)</td>
<td>(0.00467)</td>
</tr>
<tr>
<td>Obs.</td>
<td>294</td>
<td>200</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Kernel</td>
<td>Triangle</td>
<td>Triangle</td>
<td>Epanech.</td>
<td>Epanech.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Uniform</td>
<td>Uniform</td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level

Moreover, we report a number of specifications to see to what extent results are sensitive to the order of the polynomial. In Table 6, to further explore the robustness of the results, we adopt alternative approaches to implement the parametric RD design by using various functional form, varying from a linear polynomial interacted with a North dummy to a quartic polynomial subsequently as one moves from left to right. We consider the restricted sample of cities within 7 degrees latitude of the Huai River in columns (1)-(5). For example, columns (1), (2) and (3) include degree distance to the Huai River, its square, and its cube, and each term's interaction with a North dummy in the discontinuity sample. The results show that except for column (4), the figures are all qualitatively similar to those presented in the baseline specifications.

Table 6. Robustness checks of choice of functional form for latitude

<table>
<thead>
<tr>
<th>Housing price</th>
<th>Linear (within 7 degrees)</th>
<th>Quadratic (within 7 degrees)</th>
<th>Cubic (within 7 degrees)</th>
<th>Quartic (within 7 degrees)</th>
<th>Quintic (within 7 degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>-0.00833**</td>
<td>-0.0081**</td>
<td>-0.01204***</td>
<td>0.08037***</td>
<td>-0.01249***</td>
</tr>
<tr>
<td></td>
<td>(0.00467)</td>
<td>(0.00333)</td>
<td>(0.00284)</td>
<td>(0.01939)</td>
<td>(0.00081)</td>
</tr>
</tbody>
</table>

Notes: ***Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level
VII. Conclusion

The value of clean air on human health is widely acknowledged but quantitative impact of such an amenity are largely absent. This absence is due to problems of endogeneity, and those relating to the the attribution of the costs of pollution. Estimates of the value of clean air, however, are critical to the cost-benefit analysis of any policy intervention regarding environmental protection. In this paper, we draw on a unique intervention with respect to use of coal-fired heating amongst households in China to estimate the elasticity of house prices to concentrations of particulate matter of 10 micrometres or less in aerodynamic diameter (PM$_{10}$). China is a fascinating case study of the impact of pollution for at least three reasons: (i) it is the largest developing country in the world; (ii) efforts at abating pollution are on an upward trajectory; and, (iii) policies get implemented with few problems of non-compliance. The last of the above is critical to our research design that then leads to the quantitative estimates of the value of clean air.

The Huai River Policy generated a stark difference in air quality in southern and northern China. Our analysis exploits the structure of this unique Chinese Huai River Policy to provide new evidence on the capitalisation of air quality on housing prices. We employ a regression discontinuity (RD) design to estimate the impact of air pollution on house prices across a river that demarcates regions with and without coal-fired heating emanating from the Huai River Policy. This policy was decreed by the Chinese government in the 1950s that allowed burning of coal at subsidised prices for indoor heating to only the north of the Huai River. Employing quasi-experimental variation in particulate matter of 10 micrometres or less in aerodynamic diameter (PM$_{10}$) generated by this arbitrary policy and regression discontinuity (RD) design based on distance from Huai River, we estimate the local average treatment effect (LATE) of PM$_{10}$ air pollution on house prices. By using panel data of 30 large cities on either side of the river for the period 2006 to 2015, we found that 1 μg/m$^3$ reduction in average PM$_{10}$ results in a 1 percent increase in housing prices. The results are robust to using parametric and nonparametric estimation methods and adjustment to a rich set of covariates.

The findings also confirm the importance of instrumenting for local air pollution to reduce the endogeneity problem. Our econometric analysis reveals that conventional fixed effect regressions under-estimate the value of clean air. Accounting for problems of endogeneity improves the estimates, providing statistically significant parameter estimates on the effect of PM$_{10}$ on housing prices. These estimates of the economic value of clean air are robust to adjustment in covariates and are less sensitive to model specification than cross-sectional and fixed-effects estimates. Extrapolation from US-based studies suggests that the costs of poor air are substantial in terms of health effect and welfare in general. The PM$_{10}$ concentrations analysed in our dataset far exceed those used for developed countries but are typical for developing countries such as China and India. We found that the value of clean air far exceed those for the US that is exposed to substantially lower levels of air pollution. A 1 mg/m$^3$ reduction in TSPs results in a 0.2–0.4 percent increase in mean housing values in the US (Chay and Greenstone, 2005), compared to approximately 1 percent increase in housing price in China based on our estimates. Our findings suggest that the value of clean air for China exceed those reported for the rich world.
Reference


