THE VALUE OF AIR QUALITY IN CHINESE CITIES: EVIDENCE FROM LABOR AND PROPERTY MARKET OUTCOMES

X. HUANG AND B. LANZ

The value of air quality in Chinese cities: Evidence from labor and property market outcomes[*]

Xuan Huang[†]   Bruno Lanz[‡]

This version: November 2015

Abstract

Using a dual-market sorting model of workers' location decisions, this paper studies the capitalization of air pollution in wages and property prices across Chinese cities. We exploit quasi-experimental variations in particulate matter (PM_{10}) concentration induced by a policy subsidizing coal-based winter heating in northern China, specifying a regression discontinuity design based on cities' location relative to the policy boundary. We estimate that the elasticity of wages and house prices with respect to PM_{10} concentration is 0.41 and -0.71 respectively. Our results are robust to the use of an alternative source of exogenous variation in PM_{10} concentration (sandstorms), supporting the view that the local effect we measure provides policy-relevant information on the value of air quality improvements in China.

Keywords: Hedonic model; Air pollution; Labor market; Housing prices; Local public goods.

JEL Codes: H41; J31; R31; Q53.

---

1 Introduction

Along with the rapid industrialization and urbanization of China in the past decades, severely polluted air is now smothering most Chinese cities. In 2011, seven of the world's ten cities with worse air quality were located in China (Zhang and Crooks, 2012), and daily concentration of suspended particulate matter (PM_{10}) in 84 major Chinese cities averaged 118 $\mu g/m^3$ (MEP, 2012), more than five times the level at which mortality risks have been shown to increase (20 $\mu g/m^3$).[1] In spite of that, commonly proposed measures to mitigate air pollution involve reshaping the composition of primary energy use and sectoral economic activities (Guan and Liu, 2014), which are associated with concerns about high costs and adverse effects on economic growth. Providing evidence about the economic value associated with air quality improvements is thus of great importance to gauge the net welfare impact of urban air quality regulation.

Since the liberalization of the Chinese economy in the 1990s, and in particular that of the labor and land markets (see Zheng et al., 2006, for a review), labor mobility has led to spatial differentiation across Chinese cities (Zheng and Kahn, 2013). Given this, a number of empirical studies have applied the hedonic property value model of Rosen (1974) to measure the capitalization of air pollution in house prices. For example, in a study covering 35 major Chinese cities from 2003 to 2006, Zheng et al. (2010) show that housing values are lower in cities with higher levels of PM_{10}. This is confirmed by Zheng et al. (2014) using data for 85 large Chinese cities from 2006 to 2009 and a study of house prices across neighbourhood in Beijing by Zheng and Kahn (2008). These papers provide supporting evidence that residents 'vote with their feet' (Tiebout, 1956), with market prices adjusting to reflect the prevailing supply of local amenities in different urban areas.

In a system of cities in which labor is mobile, however, location decisions by workers involve both employment and housing conditions. In this paper we apply the dual-market sorting model developed by Rosen (1979) and Roback (1982) in which labor and property markets clear simulta-

---

[1] See World Health Organization (WHO, 2006). PM_{10} represents the part of suspended particles with a diameter smaller than 10 μm. These particles can enter the lungs which, among other things, can cause an increase in the prevalence of cardiovascular diseases (e.g. D'Ippoliti et al., 2003; Tsai et al., 2003). More generally, empirical evidence suggests that air pollution, and in particular PM_{10}, has significant welfare consequences, among which a negative impact on the overall level of public health (e.g. Chay and Greenstone, 2003; Pope III and Dockery, 2006; Chen et al., 2013; Qiao et al., 2014). Growing evidence also suggests that air pollution affects wider economic outcomes such as worker productivity (Graff Zivin and Neidell, 2012), labor supply (Hanna and Oliva, 2015), as well as school attendance (Currie et al., 2009).

neously. Cities that offer relatively high levels of amenities will attract workers, affecting both the supply of labor and the demand for housing, and regional differences in amenities will be compensated by variations in both wages and housing prices. Applications of this approach have shown that a large share of amenity values are accounted for by variations in wages (e.g. Hoehn et al., 1987; Blomquist et al., 1988; Graves and Waldman, 1991; Albouy, 2015). In China the study by Zheng et al. (2009) using 2004 data on the 85 large cities shows that both wages and house prices reflect variations in local amenity indicators (although they do not consider air quality as a local amenity).[2]

Relative to existing studies this paper makes two main contributions. First, we estimate spatially compensating wages and house prices differentials using a unique dataset that includes the full set of 288 Chinese prefectural-level cities (MEP, 2012).[3] More specifically, while existing studies employ a subsample of "large" cities, which experience on average higher wages, housing values and air pollution, we assemble 2011 data for all 288 cities on PM_{10} concentration (MEP, 2012), wages and house prices (NBSC, 2012a), and use data from the 2010 census (NBSC, 2010) to obtain information about local labor and property markets. This allows us to generate a comprehensive picture of intercity differences in China and provides a counterpart to the use of metropolitan statistical area (MSA) in U.S. studies (e.g. Glaeser and Tobio, 2008; Gyourko et al., 2010; Albouy, 2015).

The second contribution of this paper is to use quasi-experimental variations in air pollution in an attempt to identify a causal relationship between PM_{10} concentrations and outcomes on labor and property markets. Specifically, we exploit variations in air quality induced by the River Huai policy which subsidizes winter heating in cities that are located in northern China, where the administrative delimitation between north and south is defined by the River Huai and Qinling Mountains. Because heating systems are, for a large part, inefficient and predominantly use coal

---

[2] Interestingly, Zheng et al. (2009) find little evidence of capitalization effects for the year 1998, but that compensating differentials increased markedly when comparing data from 1998 and 2004. This may reflect both an higher demand for local amenities associated with growth in per capita income, as well as the emergence of a system of open cities towards the end of 1990s (Zheng et al., 2010).

[3] In the country's administrative structure, a prefectural city is an administrative division ranking below a province and above a county/town. A prefectural city emerges if the administrative subdivision meets the following three criteria: (i) the non-agricultural population in the central urban area is over 250 thousand; (ii) the gross value of industrial output is more than CNY 200 million (around USD 31 million); (iii) tertiary industry's share is higher than the primary industry, accounting for over 35% of the gross regional production (GRP).

as a fuel source, the policy heavily contributes to air pollution in the northern part of China. As initially suggested by Almond et al. (2009), the River Huai policy provides a setting to implement a regression discontinuity design, using the discrete jump of air quality in nearby cities across the policy boundary. For example, Chen et al. (2013) employ a similar strategy to document a causal impact of total suspended particles on health outcomes.

Our paper is also germane to the growing hedonic pricing literature that uses quasi-experimental sources of variations to motivate instrumental variable (IV) strategies.[4] This approach has two key advantages. First, since it is practically impossible to control for all possible determinants of wages and house prices, the development of IVs can alleviate recurring concerns about omitted variable bias (Kuminoff et al., 2010). Second, using exogenous sources of variations for amenities of interest (here air pollution) explicitly addresses the issue that amenities are not randomly assigned across location, but are themselves the outcome of an equilibrium sorting process (see Kuminoff et al., 2013, for a review). For example, if households have heterogeneous preferences for air quality, and if they spatially sort according to their preferences, then the level of air quality is an endogenous outcome of the sorting process. In turn, standard OLS estimates of implicit prices are likely to be severely biased towards zero (Chay and Greenstone, 2005).

In the present context, an other benefit of using a regression discontinuity design is that it mitigates concerns about the quality of Chinese air pollution data. While we discuss this issue further below, our identification strategy essentially compares cities that are near to the policy boundary, so that our results would be affected by data quality only insofar as cities included or not under the River Huai policy differ in how they report the data (Almond et al., 2009). On the more negative side, one drawback of our identification strategy is that we discard information from cities located away from the policy boundary, and our estimate represents only a local effect. Since this potentially reduces the policy-relevance of our results, we provide further evidence derived from a distinct set of instruments suggested by Zheng et al. (2014), namely the distance to desert areas. Particulate matter transported by wind from deserts is a major source of PM_{10}, so that this provides an opportunity to check the consistency of our local estimates across a wider set of cities.

Our main findings confirm the existence of compensating differentials in both labor and prop-

---

[4] See for example Black (1999), Epple and Sieg (1999), Davis (2004), Chay and Greenstone (2005), Linden and Rockoff (2008), and Pope (2008).

erty markets across Chinese cities. Weak-instruments tests are passed by both sets of instruments, and we find that they provide very similar local estimates of implicit prices. In our preferred specification, elasticity estimates for wages and house prices with respect to PM_{10} concentration are 0.41 and -0.71 respectively. We also carry out a set of calculations to illustrate the magnitude of our findings. At the individual level, a reduction of PM_{10} concentration by 1 $\mu g/m^3$ is associated with an average value of CNY 208.90 ($\simeq$ USD 32.37).[5] At the city level, this amounts to an aggregate change in wages of around CNY 85.7 million per year on average across cities ($\simeq$USD 13.3 million), and an increase in the value of housing stock by CNY 761.7 million ($\simeq$USD 118 million). While these are large numbers, it only represents a fraction of recent estimates associated with the 1.28 million premature deaths that can be attributed to ambient air pollution each year in China(OECD, 2014).[6]

The remaining of this paper proceeds as follows. In Section 2 we briefly review the dual-market hedonic model. In Section 3 we provide evidence about wages, house prices, and air pollution across Chinese cities, and discuss our identification strategy for implicit prices. Section 4 reports our main estimation results. Section 5 concludes.

2   Theory: Compensating differentials and local amenities

This section presents the basic theoretical framework underlying our estimation. It builds on Rosen's (1979) idea that individuals' decisions to locate in a given city involves both employment and housing. Intuitively, locations with higher levels of local amenities will attract relatively more workers, which induces an increase in the supply of labor and in the demand for housing. This, in turn, will drive wages down and property prices up. Thus as compared to the traditional hedonic property value model (Rosen, 1974), the dual-sorting equilibrium implies that the economic value of local amenities will be reflected in both wages and house prices, and spatially compensating price differentials on both markets can be identified by jointly estimating a hedonic wage function and a hedonic property value function.

---

[5] Throughout the paper we use an average exchange rate for 2011 of CNY 1 $\simeq$ USD 0.155.

[6] In particular OECD (2014) calculates that, in 2010, deaths attributed to ambient air pollution translate into 24.58 million years of life lost. Based on this, they estimate that the economic cost of outdoor air pollution and associated health impacts is about USD 1.4 trillion.

Equilibrium interactions between wages, house prices, and the provision of local amenities is formalized in the model by Roback (1982). In this framework, implicit prices of local amenities is an outcome of workers and firms' location decisions considering (respectively) the attributes and production costs associated with different locations. Specifically, a representative individual maximizes a utility function $U(X, H; Q)$ by making choices over a numeraire good X (with price normalized to one) and a non-traded good H (i.e. housing), as well as a location-specific amenity, here a measure of air quality Q, subject to a budget constraint:[7]

$$\max_{X,H,Q} \quad U(X, H; Q) \quad s.t. \quad w + I = X + Hr \tag{1}$$

where w is the wage rate, I is non-labor income (independent of location), and r is the price of housing.

Market equilibrium implies that wages and house prices adjust to make individuals indifferent between locations. Thus in equilibrium individuals achieve the same utility level $\overline{V}$ in different locations (otherwise they would move), so that indirect utility is constant across locations: $V(w, r; Q) = \overline{V}$. Here I and the price of good X are assumed constant and omitted for simplicity. Denoting partial derivatives by $\partial V / \partial w = V_w$, we have $V_w > 0$ and $V_r < 0$, and since air quality can a priori be assumed to have a positive impact on household's welfare, $V_Q > 0$.

Firms' location decision follows a similar logic. Consider a representative cost-minimizing firm which admits a constant return to scale production technology. Profit maximization requires that unit cost $C(w, r; Q)$ of producing X is equal to the price of X (here normalized to one), and in equilibrium production costs have to be the same in all locations, so that: $C(w, r; Q) = 1$. While we have that $C_w > 0$ and $C_r > 0$, the impact of air pollution on production costs depends on technology. Specifically, if pollution is a by-product of the production process, then pollution abatement is likely to be costly to the firm, so that $C_Q > 0$. It is also possible, however, that pollution negatively affects production possibilities (e.g. through reduced labor productivity), which would imply that $C_Q < 0$.

Market interactions between workers and firms imply that the impact of air pollution on wages and housing costs depends on both preferences and technology. This can be seen by totally differ-

---

[7] As in Roback (1982), the use of a representative agent implicitly imposes that all individuals chose their location simultaneously, and that they are identical in tastes and skills.

entiating equilibrium conditions for individuals and firms and collecting terms:

$$\frac{dw}{dQ} = \frac{1}{\Phi}\left(-V_Q C_r + V_r C_Q\right) \tag{2}$$

$$\frac{dr}{dQ} = \frac{1}{\Phi}\left(-V_w C_Q + V_Q C_w\right) \tag{3}$$

where $\Phi = V_w C_r - V_r C_w > 0$. When $V_Q > 0$, increasing air quality Q while maintaining utility at level $\overline{V}$ would require either higher house prices r or lower wages w. If in addition $C_Q > 0$ increasing Q while keeping production costs constant would require lower expenditures on wages or housing costs. Taken together, this implies that locations with better air quality command higher equilibrium wages (i.e. equation (2) is unambiguously positive). However, the equilibrium impact of pollution on housing price (equation 3) is indeterminate, as it reflects conflicting interests of households and firms. Symmetrically, when $C_Q < 0$ the sign of (2) is indeterminate, while equation (3) is negative.

These equations can be used to derive an expression for the implicit price for local amenities P_Q^*, which measures the change in income required to compensate individuals for marginal changes in Q. In particular, solving for the ratio V_Q/V_w we have that:

$$P_Q^* = \frac{V_Q}{V_w} = -\frac{dw}{dQ} - \frac{V_r}{V_w}\frac{dr}{dQ}, \tag{4}$$

From Roy's identify $-\frac{V_Q}{V_w}$ can be interpreted as the inverse demand curve or marginal willingness to pay (MWTP) for air quality, and similarly the demand for housing is $-\frac{V_r}{V_w} = H^*$. Hence we have:

$$MWTP_Q = -\frac{dw}{dQ} + H^* \cdot \frac{dr}{dQ}. \tag{5}$$

MWTP for air quality improvement (i.e. the change in income that would leave individuals indifferent to marginal air pollution reduction) can be estimated though a hedonic wages equation and a hedonic property value model, respectively identifying dw/dQ and dr/dQ.[8]

One fundamental issue with the identification of implicit prices using spatially compensat-

---

[8] As suggested in the initial Roback (1982) paper, this approach can be used to construct an index-ranking of cities capturing variations in the quality of life. See for example Gyourko et al. (1999). In this paper we rather focus on estimating the implicit price of air pollution as we were not able to locate plausibly exogenous variations for the set of amenities that would make a city-ranking exercise meaningful.

ing variations is the endogeneity of local amenities. If individuals have heterogeneous tastes for amenities (here air quality), then Tiebout's (1956) sorting mechanism implies self-selection of individuals with specific tastes, making the level of air quality endogenous (see Kuminoff et al., 2013, for a review). Similarly, firms with inefficient technologies will be attracted by location with lax regulation, and since air quality problems in China are mainly driven by the use of energy (coal), this may affect labor productivity and wages through firms' technology. This endogeneity problem is also related to recurring concerns about omitted variables in hedonic regressions (Cropper et al., 1988; Kuminoff et al., 2010), since air pollution is likely correlated with unobserved characteristics of different locations. Using exogenous sources of variation in air quality is thus crucial to identify (local) MWTP estimates, and we discuss two different sets of potential instruments in next section.

A second noteworthy issue with the Roback (1982) framework is the absence of migration costs. In particular, if location decisions of individuals and firms are constrained, then adjustments in wages and house prices will not fully reflect the value of amenities. Intuitively, this implies that ignoring migration costs biases MWTP estimates downwards (see Bayer et al., 2009; Kennan and Walker, 2011). China has notoriously implemented mobility restrictions through the *Hukou* policy to regulate migration of rural workers, and although this policy has been phased out since 1990 and completely abolished in some provinces (see Zheng et al., 2014), it should be kept in mind that our implicit assumption of free labor mobility is a simplification.[9] In turn, our MWTP estimates represent a lower bound of true benefits associated with air quality improvements.

3 Data and empirical strategy

In this section we discuss identification of the implicit price of air quality. We first present descriptive evidence about variations in inter-city wages and housing prices as well as PM_{10} concentrations. We then present our identification strategy and discuss our sources of exogenous variation for air pollution.

---

[9] Starting in 1958, the Hukou policy classifies people into local and non-local residents as well as agricultural and non-agricultural status. This classification was used mainly to grant access to local public services (such as children' access to public schools), and has thus restricted labor mobility for decades. Importantly however, the Hukou system did not impose restrictions for housing market transactions (Zheng et al., 2010).

3.1 Descriptive evidence: Wages, housing prices and air quality in China

The baseline year for our analysis is 2011, being the only year in which data on air pollution is available for *all* 288 prefecture-level cities (MEP, 2012). Data on regional wages and house prices are taken from NBSC (2012a) and also refer to 2011. Moreover, referring to 2011 is convenient because detailed census data on regional labor and property markets is available from the 2010 census (NBSC, 2010). Data sources and summary statistics for all the variables used in the analysis are provided Appendix A.

Observed variations in wages and housing prices across Chinese cities are substantial, as illustrated by Figure 1. Yearly wages is on average CNY 35,957 ($\simeq$USD 5,566) and house prices, which measures annual average transaction price for residential dwellings (in CNY/m^2), is on average CNY 4,117 ($\simeq$USD 637). Across cities, we observe a positive correlation between wages and housing values (ρ=0.63, p-val.<0.01).

Our main measure of air quality is yearly average PM_{10} concentration in $\mu g/m^3$. This data is mapped in Figure 2 together with the distribution of cities across air quality thresholds defined by WHO (2006).[10] The reported concentrations of PM_{10} is consistent with the widespread concerns about air pollution in Chinese cities. In 2011, among all the 288 prefectural cities, none of them achieved the lowest threshold level (20 $\mu g/m^3$), and almost 70% of them do not meet the laxest standard (70 $\mu g/m^3$).

An alternative widely used measure of air pollution in China is the air pollution index (API). As we detail in Appendix B, the API is computed by each city (NBSC, 2012a) and combines data on the concentration of PM_{10}, SO_2 and nitrogen dioxide (NO_2). As an alternative measure of air quality we use the number of days during which air pollution in a given city area is slightly polluted or worse (i.e. API is above 100). The value of 100 is selected because it is the first threshold at which health issues are known to appear.

Importantly, some researchers have raised questions regarding the manipulation of API data reported by cities. Recently China's central government has used air quality as one of the criteria

---

[10] According to the air quality guidelines from WHO (2006), the lowest level of annual average PM_{10} concentrations at which total, cardiopulmonary and lung cancer mortality have been shown to increase is $20\mu g/m^3$. There are also three intermediate targets of PM_{10}. At 30 $\mu g/m^3$ the long-term risk of premature mortality increases by around three percent, at 50 $\mu g/m^3$ the risk increases by around 10 percent, and at 70 $\mu g/m^3$ the risk is around 15 percent higher.

Figure 1: Average wages and house prices in 288 Chinese cities, 2011 (source: NBSC, 2012a)
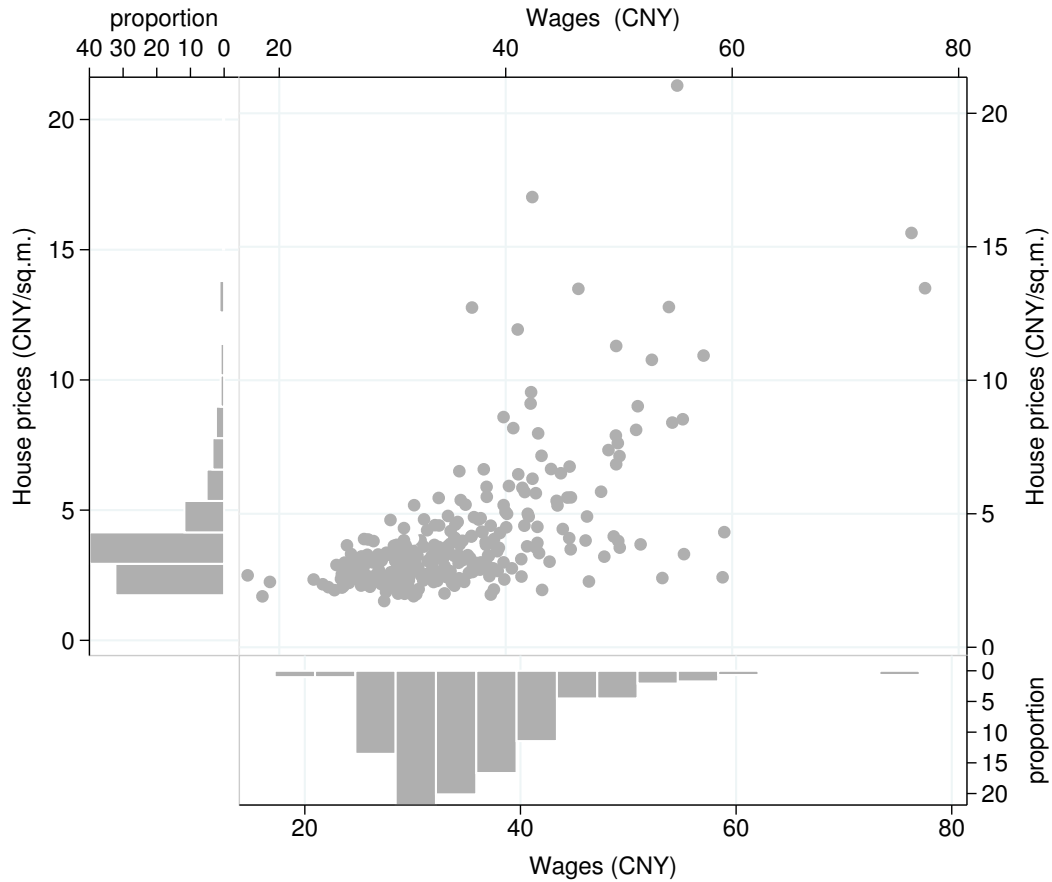
to evaluate the performance of local governments, and the "National environmental protection model city" award requires more than 85% days per year with daily API lower than 100. Therefore, the local government has incentives to manipulate reported API, likely inducing reported pollution levels below real values. For example, Andrews (2008) suggests the potential for a systematic under-reporting of API readings around the threshold of 100 in Beijing, which generates inconsistent metrics between API and PM_{10} concentrations. Similarly, Ghanem and Zhang (2014) and Ravetti (2015) provide evidence about discontinuities at different thresholds, which may be interpreted as data manipulation.

However, API data has also been shown to provide useful information about the variation of air pollution among different regions. Chen et al. (2012) compare API data with visibility and

Figure 2: Air quality in 288 Chinese cities, 2011 (source: MEP, 2012)
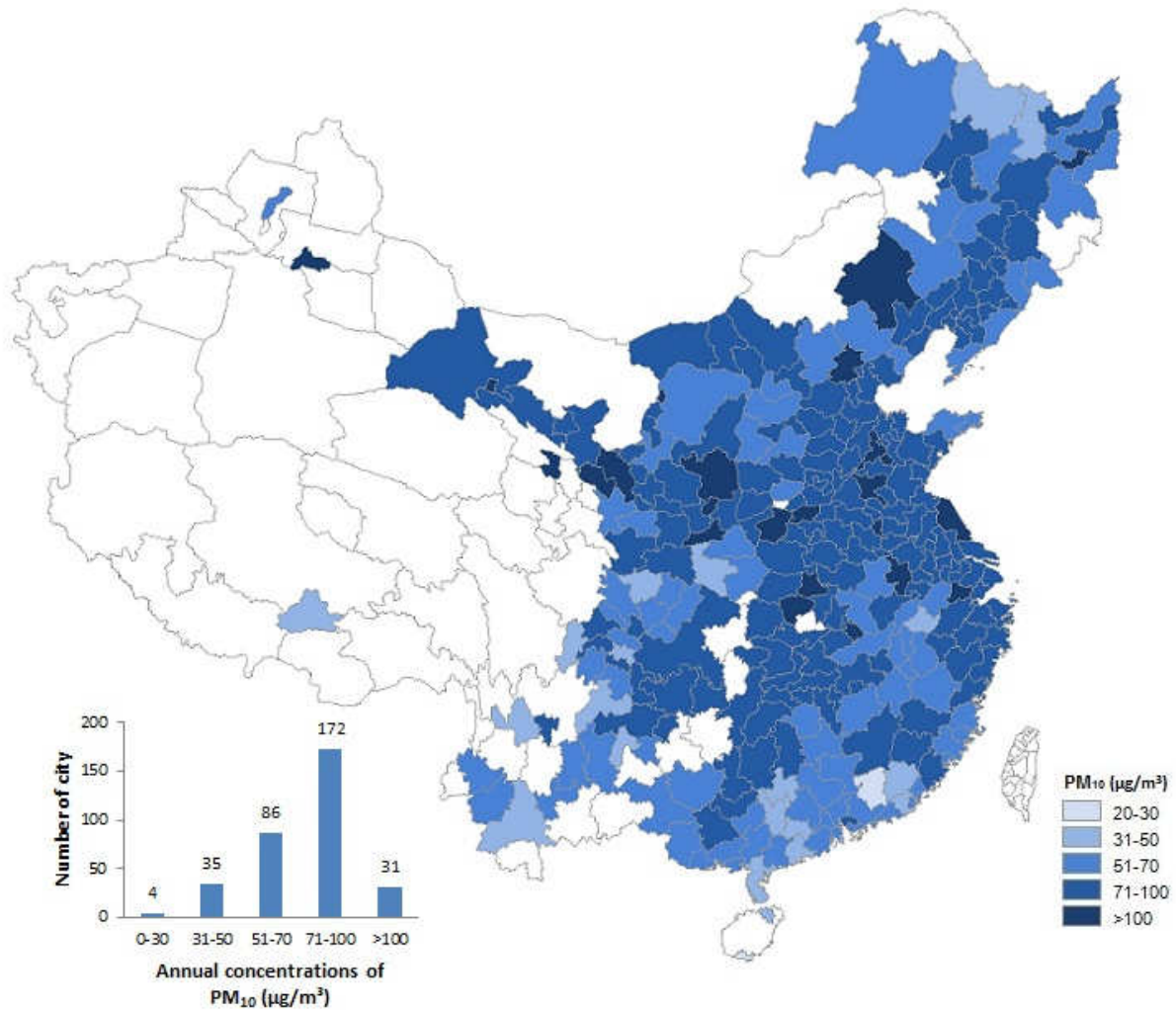
aerosol optical depth measurement to examine potential data manipulation, showing that both measures are essentially consistent with each other. Moreover, the correlation between API and aerosol optical depth does not change significantly when the API is closely below or above 100. Nevertheless, we will focus on PM_{10} concentrations as our primary measure of air quality, which is likely to be more difficult to manipulate. Moreover, because PM_{10} concentration is related to visibility, it is also expected to be more directly related to household behavior and thus reflected in market outcomes.

3.2 Estimation and identification: Exogenous sources of air quality variations

In the framework laid out by Rosen (1979) and Roback (1982), the implicit price of air pollution can be identified from inter-city wages and house prices differentials. Formally, indexing cities by i, we write the hedonic wages equation as:

$$Wages_i = \alpha_0 + \alpha_1 Air\ pollution_i + \alpha_2 House\ prices_i + \alpha_3 X_i^{wages} + \varepsilon_i^{wages} \qquad (6)$$

where the α's are parameters to be estimated from the data, X^{wages} is a vector of covariates, and ε is an error term. Similarly the hedonic house prices equation is given by:

$$House\ prices_i = \beta_0 + \beta_1 Air\ pollution_i + \beta_2 Wages_i + \beta_3 X_i^{house\ prices} + \varepsilon_i^{house\ prices} \qquad (7)$$

where notation follows the same logic as above.

Simultaneity in the determination of wages and housing prices calls for an instrumental variable (IV) approach. In particular, in the wage equation (6) we instrument housing prices with the set of determinants included in $X^{house\ prices}$ ("price shifters"), while in the house prices equation (7) we instrument wages with variables included in X^{wages}. Thus identification of the partial effects of housing prices on wages requires that variables in the vector X_2 can plausibly be excluded from X_1, and vice versa.

The main coefficients of interest, α_1 and β_1, are the two components of the MWTP for air quality (see equation 5). As discussed above however, air pollution is likely endogenous, and we use two sources of exogenous variations in PM_{10} concentration to identify partial MWTP estimates in equations (6) and (7). Using different sets of IVs is important because they capture local effects at different points in the sample. Thus consistency in the results would provide more credence that our MWTP estimates represent a policy-relevant measure of the value of air quality improvements.

Our main set of instruments is borrowed from Almond et al. (2009) and exploits variation induced by the Huai River policy. The aim of this policy intervention is to subsidize winter heating in northern Chinese cities, where the boundary of the policy is defined by the River Huai and Qinling Mountains, near the $33°$ latitude. Cities in the north face on average colder winter temperatures than those in the south, although the boundary of the policy purely reflects an administrative decision as the climate of cities located along the policy boundary is very similar.

Because heating systems are coal-based and typically quite inefficient, the policy has been shown to be associated with substantial amount of particulate matter emissions (Almond et al., 2009). In turn, based on the boundary of the policy, Chen et al. (2013) have used a regression discontinuity design to evaluate a causal impact of suspended particulate matter concentration on health and mortality. This approach effectively compares cities that were included or not in the Huai River policy, exploiting the discrete change in use of coal-based heating systems and associated variations in air pollution.

In the present paper we follow the same approach as Chen et al. (2013), and specify the following equation for the level of PM_{10} concentration:

$$Air\ pollution_i = \gamma_0 + \gamma_1 I_i^{Huai\ policy} + \sum_n \gamma_n (Latitude_i)^n + \varepsilon_i^{Huai\ policy} \qquad (8)$$

where the γ's are parameters to be estimated, $I_i^{Huai\ Policy}$ is an indicator variable which is equal to one if city i is included in the River Huai policy, the variable *Latitude* measures the distance (in degrees) between the city and latitude $33°$ (comuputed from NGCC, 2012), and ε is an error term. In equation (8) the indicator variable captures the discrete increase in air pollution among northern and southern cities caused by the River Huai policy (after adjusting for a flexible polynomial function in latitude).

The second set of instruments is inspired by the work of Zheng et al. (2014), who uses information about sandstorms as an other source of exogenous variations in air quality. Sandstorms represent an important source of suspended particles, including PM_{10}, with impacts on local concentrations being measured as far as the North America's west coast (McKendry et al., 2001). Following Zheng et al. (2014), we use a measure of the proximity to deserts in Inner Mongolia to capture exogenous contribution of sandstorms to city-level PM_{10} concentration.[11] We specify a polynomial specification to model the impact of sandstorms on city-level PM_{10} concentration:

$$Air\ pollution_i = \delta_0 + \sum_n \delta_n (Sandstorm_i)^n + \varepsilon_i^{Sandstorm} \qquad (9)$$

---

[11] More specifically, we compute the average distance to the four largest sand storm origins: Lashan (256 days with sandstorms), Xilingol (163 days with sandstorms), Bayannur (142 days with sandstorms), and Erdos (138 days with sandstorms). The fifth largest sandstorm origin, Xingan (66 days with sandstorms), experiences a significantly lower number of days with sandstorm, and excluding it from the analysis does not affect our results.

Table 1: Estimation results for alternative specifications of the pollution equation

|	PM_{10} (1)	PM_{10} (2)	PM_{10} (3)	PM_{10} (4)	PM_{10} (5)	PM_{10} (6)	Polluted days (7)	Polluted days (8)	Polluted days (9)
Huai Policy	6.33* (3.34)	6.38** (2.93)	7.69** (3.23)	–	–	–	15.93*** (3.62)	15.97*** (3.44)	15.95*** (3.80)
Latitude	0.62** (0.25)	0.84*** (0.22)	0.52 (0.40)	–	–	–	0.32 (0.27)	0.47* (0.26)	0.48 (0.47)
$Latitude^2$	–	-0.17*** (0.02)	-0.18*** (0.02)	–	–	–	–	-0.12*** (0.02)	-0.12*** (0.02)
$Latitude^3$	–	–	0.002 (0.003)	–	–	–	–	–	-0.00002 (0.003)
Sandstorm	–	–	–	-18.69*** (2.06)	31.21*** (10.28)	19.28 (34.43)	–	–	–
$Sandstorm^2$	–	–	–	–	-18.82*** (3.80)	-8.85 (27.74)	–	–	–
$Sandstorm^3$	–	–	–	–	–	-2.51 (6.92)	–	–	–
Constant term	74.77*** (1.78)	82.19*** (1.76)	81.86*** (1.79)	101.82*** (2.86)	73.15*** (6.42)	77.30*** (13.13)	15.16*** (1.93)	20.42*** (2.06)	20.42*** (2.11)
Adjusted R^2	0.123	0.324	0.324	0.221	0.280	0.278	0.193	0.270	0.268
F-statistic	21.16	46.93	35.42	82.57	56.92	37.87	35.23	36.42	27.22

Notes: 288 observations for each regression. In columns (1) to (3) and (7) to (9) the dependent variable is average PM_{10} concentration in 2011 (in $\mu g/m^3$). In columns (4) to (6) the dependent variable is the number of days with API $\geq$ 100 in 2011. OLS estimates reported. Standard errors in parentheses. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

where the δ's are parameters to be estimated, *Sandstorm* measures the distance to the sandstorm origins (, in thousand km, computed from NGCC, 2012), and ε is an error term.

Initial evidence about our instruments for air pollution is presented in Table 1. In columns 1 to 3, we report results for equation (8) with average PM_{10} concentration as the dependent variable and linear, quadratic and cubic specifications in each respective column. Columns 4 to 6 report similar specifications for equation (9). Finally columns 7 to 9 report results for equation (8) with the number of days in which API is greater than 100 as the dependent variable.

For equation (8) the discontinuity indicator is positive and highly statistically significant, and relatively similar across specifications. This confirms evidence from Almond et al. (2009) and Chen et al. (2013) that the River Huai policy generate significant variations in air quality nearby the policy boundary, as measured by PM_{10} and the number of polluted days. For equation (9), we find that the distance to sandstorm origins is negatively related to PM_{10} concentrations, and that the marginal effect declines with distance, which is consistent with Zheng et al. (2014). For the three specifications we report, our results further suggest that quadratic specifications are generally preferred, and in the remaining of the analysis we thus use quadratic polynomial functions for both equations (8) and (9). This is illustrated graphically in Figure 3 for PM_{10} concentrations.

Figure 3: Summary measure of average PM_{10} concentration around the River Huai
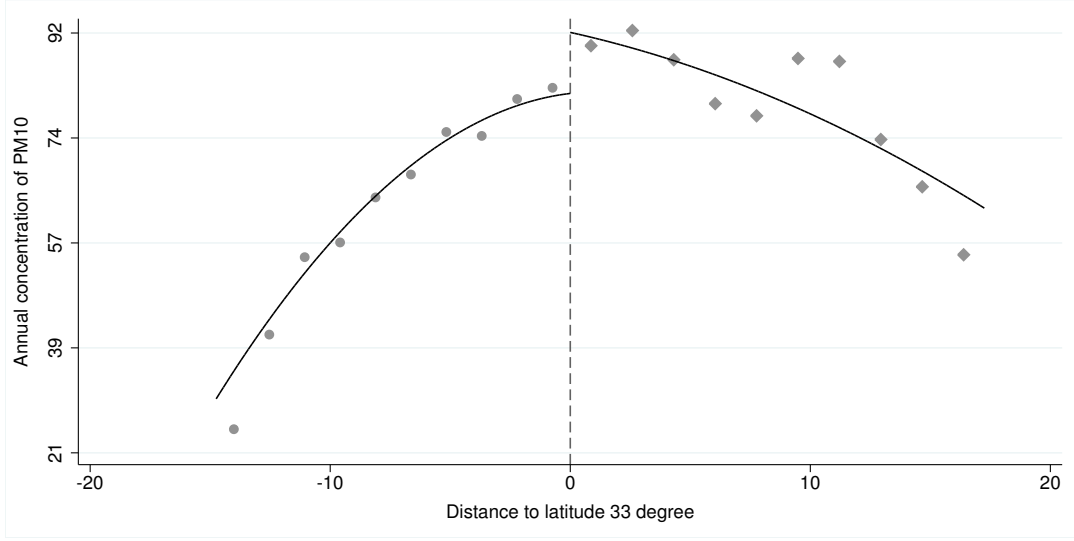
Notes: Yearly average PM_{10} concentration (in $\mu g/m^3$) as a function of distance to the River Huai policy threshold (in latitude). Points represent average value across cities, with cities grouped in ten bins according to the distance to policy threshold (N=288). Solid line shows fitted values from a quadratic function.

4 Estimation results

This section first reports regression results for the hedonic wages and house prices equations. We then illustrate the magnitude of the results by computing the aggregate value of air quality improvements across Chinese cities as well as individual-level MWTP.

4.1 Regression results

Table 2 presents the results for the wages equation across several specifications, regressing average yearly wages (in 1000 CNY/Year) on average house prices (in 1000 CNY/m^2), air pollution, and a number of controls that characterize the labor market across cities. In order to assess the importance of endogeneity bias of house prices and pollution, we start by reporting OLS estimates that treat house prices and air pollution as exogenous (column 1). Columns (2) and (3) then provide separate evidence about endogeneity of house prices and air pollution respectively. In column (2) we instrument house prices with characteristics of regional housing markets, and in column (3) we instrument PM_{10} concentrations using a quadratic specification defined in equation (8). Column

Table 2: Estimation results for the hedonic wages equation

	Wages OLS1 (1)	Wages IV1 (2)	Wages IV2 (3)	Wages IV3 (4)	Wages IV4 (5)	Wages IV5 (6)	Wages IV6 (7)
$\ln(PM_{10})$	0.07*	0.07*	0.32***	0.41***	0.36***	–	–
	(0.04)	(0.04)	(0.07)	(0.08)	(0.08)		
$\ln$(Polluted days)	–	–	–	–	–	0.09***	–
						(0.02)	
PM_{10}	–	–	–	–	–	–	0.23***
							(0.04)
$\ln$(House prices)	0.18***	0.10	0.18***	0.17***	0.09	0.26***	–
	(0.03)	(0.07)	(0.03)	(0.07)	(0.07)	(0.06)	
House prices	–	–	–	–	–	–	1.34***
							(0.43)
Workforce	-0.01**	-0.01**	-0.01***	-0.01***	-0.01**	-0.01***	-0.32**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.14)
Male	-0.002	-0.002	-0.002	-0.003	-0.003	-0.002	-0.10
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.09)
College	0.02***	0.02***	0.01***	0.01**	0.01***	0.01***	0.38***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.15)
Service share	-0.001	-0.0003	0.0004	0.001	0.002	-0.002	0.01
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.06)
Unemployment	-0.01	-0.02	-0.03*	-0.03*	-0.04**	-0.04**	-1.02*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.60)
Migrants	0.001**	0.001**	0.001**	0.001**	0.001***	0.001**	0.03
	(0.0004)	(0.0005)	(0.0004)	(0.001)	(0.0005)	(0.001)	(0.02)
Constant term	3.66***	3.71***	2.70***	2.37***	2.63***	4.13***	41.73***
	(0.34)	(0.34)	(0.42)	(0.46)	(0.45)	(0.38)	(13.45)
IV: House price	N	Y	N	Y	Y	Y	Y
IV: Pollution	N	N	Y	Y	(sandstorms)	Y	Y
R^2	0.529	–	–	–	–	–	–
Cragg-Donald	–	15.01	40.46	11.71	12.28	9.258	11.17
1st part. F-stat: House prices	–	15.01	–	13.23	13.99	10.99	11.83
1st Shea part. R^2: House prices	–	0.333	–	0.371	0.363	0.346	0.343
1st part. F-stat: PM_{10}	–	–	40.46	12.38	13.04	14.19	13.84
1st Shea part. R^2: PM_{10}	–	–	0.305	0.355	0.346	0.404	0.379

Notes: 288 observations for each regression. The dependent variable is the log of average yearly wages (in 1000 CNY/Year), except for column (7) where it is in levels. Standard errors in parentheses. In the IV panel, 'Y' indicates that the variable is instrumented, 'N' that it is not. LIML estimates reported. Critical values for the Cragg-Donald statistics (Stock and Yogo, 2005) for a maximal LIML size (5% significance level) are: 3.8 for Wages IV1, 6.5 for Wages IV2, and 3.6 for Wages IV3-IV6. Estimation results of the first stage for IV estimation are reported in Appendix C Tables C1 and C2. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

(4) provide our main results, jointly instrumenting house prices and pollution. Results for the first stage regressions of IV specifications are reported in Tables C1 and C2 of Appendix C.[12]

While the positive and statistically significant impact of house prices on wages is only modestly affected by the use of instruments, the impact of PM_{10} concentration is only of importance in specifications where it is instrumented (columns 3 and 4). The specification in which both house prices and pollution are instrumented (column 4) suggests the elasticity of wages with respect to PM_{10} concentration is positive and highly statistically significant, with a point estimate of 0.41. At

---

[12] The Cragg-Donald test statistics are mostly above the value of 10, considered as a rule of thumb for instruments not to be weak, although they are close to the critical values reported in Stock and Yogo (2005). We thus report Limited Information Maximum Likelihood (LIML) estimates, which are more reliable than two stage least square estimates (TSLS), and report critical values of the Cragg-Donald statistics in the notes of the result tables. Note that TSLS estimates are, however, very similar to LIML estimates reported here.

the mean of the sample, this implies that a unit increase in PM_{10} concentration is associated with an increase of average yearly wages by CNY 192.40 ($\simeq$USD 29.82).

Other determinants of the labor market outcomes generally have the expected effect on average wages.[13] In particular, a larger workforce and higher unemployment rates are associated with lower wages, while a more educated workforce and a larger share of immigrants in the city is associated with higher wages. The share of services in GDP is also positively correlated with wages, although the effect is not statistically significant and mostly picked up by our measure of education.

Columns (5) to (7) provide further evidence about the relationship between air pollution and wages under alternative assumptions. First, in column (5) we employ our alternative set of instruments, which measure the impact of sandstorm on inter-city variations of PM_{10} concentration. We use a quadratic specification for the average distance to sandstorm origins (see equation 9). Results suggest a slightly lower elasticity of wages with respect to air pollution, although the difference between alternative sets of IVs is small (around 10%).[14]

The specification reported in column (6) is based on an alternative measure of air pollution, namely the number of days in which the API is no less than 100. Recall that this measure captures average daily API values, and combines information on the concentration of PM_{10}, SO_2 and NO_2 (see Appendix B). Instruments again exploit variations induced by the River Huai policy, although results are virtually identical if we use the other set of instruments. We find an elasticity of around 0.1, so that a marginal increase in the number of polluted days is associated with an increase of wages by CNY 147.60 ($\simeq$USD 22.88) at the mean of the sample.[15]

The final specification for the wages equation, reported in column (7), is based on level-coded wages, house prices and PM_{10} concentrations. We find that the results are broadly consistent with

---

[13] It should be emphasized that these estimates are not meant to represent causal relationships, but mainly capture structural determinants of the labor market. That is, given the simultaneous equation framework considered, we mainly use them as shifters for the identification how wages impact house prices (see Table 3 below).

[14] Also note that for all the IV specifications instruments pass the weak identification test and in general estimates from the first stage regressions displayed in Appendix C have the expected sign (although interpretation is complicated by the large number of controls included).

[15] We also investigated the use of yearly API average, although publicly-available data for this indicator only covers 83 large prefectural cities (the same subsample used in Zheng et al., 2009, 2014, for example). Results suggest that a marginal increase in average API is associated with a CNY 470 increase in wage (p-val.=0.013), or an elasticity of 0.76 at the mean of the sample. Importantly, the subsample of larger cities features an average wage of around CNY 41,000 relative to about CNY 36,000 for full sample of cities.

the corresponding log-log specification in terms of sign and statistical significance. The elasticity of wages with respect to PM_{10} concentration is around 0.49 at the mean of the sample, which is quite close to the 0.41 elasticity reported in column 4.

Turning to the hedonic equation for the housing market, Table 3 reports results for regressions of average house prices on average yearly wages, air pollution, and key characteristics of local property markets. Column (1) to (4) again provide evidence about the impact of endogeneity on implicit price estimates. In column (1) OLS estimates are reported, in column (2) only the wages variable is instrumented, in column (3) pollution is instrumented with a quadratic specification defined in equation (8), and in column (4) both wages and pollution are instrumented. Results for the first stage regression of IV specifications are reported in Tables C3 and C4 of Appendix C.[16]

We find strong evidence that endogeneity significantly biases marginal effects of both wages and PM_{10} concentration towards zero. When instrumenting both variables (column 4), the elasticity of house prices with respect to PM_{10} is negative and highly statistically significant, and estimated at -0.71. At the mean of the sample, this implies that a marginal increase in average PM_{10} concentration is associated with reduction of house prices by around CNY 37.70/m^2 ($\simeq$USD 5.84). At the individual level, given an average per capita living space of 30.63m^2 (NBSC, 2010), this corresponds to CNY 1154.60 ($\simeq$USD 178.97).

Other determinants of house prices have again the expected sign. Appliances (in particular fitted kitchen) and population density (a measure of housing demand) have a positive impact on house prices, while increasing the supply of land by local governments depresses house prices. Moreover, while cities with more universities per resident have higher house prices, the number of hospital beds per resident, has a negative sign and is highly statistically significant (although the effect is quantitatively small).

Column (5) to (7) provide further evidence on the relationship between air pollution and house prices. First, column (5) shows that using instruments capturing the distance to sandstorm origins leads to estimates that are virtually identical. Similarly, column (6) uses the number of days with API greater or equal to 100 as a measure of air quality, leading to qualitatively similar results; at the average of the sample, an increase in the number of days with API$\geq$100 is associated with a

---

[16] Table 3 again reports LIML estimates (with critical values for the Cragg-Donald test statistics reported in the notes of the result tables), although TSLS estimates are very similar.

Table 3: Estimation results for the hedonic property market equation

|	Housing OLS1 (1)	Housing IV1 (2)	Housing IV2 (3)	Housing IV3 (4)	Housing IV4 (5)	Housing IV5 (6)	Housing IV6 (7)
$\ln(PM_{10})$	-0.06	-0.14*	-0.33***	-0.71***	-0.70***	–	–
	(0.06)	(0.07)	(0.10)	(0.16)	(0.17)		
ln(Polluted days)	–	–	–	–	–	-0.14***	–
						(0.04)	
PM_{10}	–	–	–	–	–	–	-0.08***
							(0.02)
ln(Wages)	0.75***	1.95***	0.80***	2.06***	2.13***	1.72***	–
	(0.09)	(0.30)	(0.09)	(0.31)	(0.37)	(0.30)	
Wages	–	–	–	–	–	–	0.34***
							(0.05)
Piped water	0.003***	-0.001	0.001***	0.0001	-0.0001	0.001	-0.0003
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.01)
Fitted kitchen	0.003	0.01***	0.003	0.01***	0.01***	0.01**	0.04**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.02)
Fitted toilets	0.001	-0.002	0.001	-0.001	-0.001	-0.001	0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.01)
Fitted shower	0.004***	0.003**	0.003***	0.002	0.001	0.001	-0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.01)
Density	0.16***	0.15***	0.18***	0.18***	0.18***	0.19***	1.31***
	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.32)
Land supply	-0.06	-0.07	-0.06*	-0.08	-0.08	0.14	-0.78**
	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	(0.23)	(0.33)
Coastal	0.27***	0.22***	0.21***	0.09	0.09	0.11	0.62
	(0.06)	(0.08)	(0.06)	(0.09)	(0.09)	(0.09)	(0.59)
Universities	10.17***	10.24***	10.30***	10.51**	10.51**	9.46**	44.84*
	(2.92)	(3.65)	(2.98)	(4.15)	(4.21)	(3.94)	(26.82)
Hospitals	-0.02	-0.08***	-0.02	-0.07***	-0.08***	-0.06***	-0.42***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.16)
Constant term	-1.89***	-5.66***	-0.92*	-3.70***	-3.97***	-5.09***	-4.86***
	(0.39)	(1.01)	(0.49)	(0.90)	(1.04)	(1.02)	(1.58)
IV: Wages	N	Y	N	Y	Y	Y	Y
IV: Pollution	N	N	Y	Y	(sandstorms)	Y	Y
R^2	0.663	–	–	–	–	–	–
Cragg-Donald	–	12.34	49.46	6.85	7.59	8.15	8.03
1st part. F-stat: Wages	–	13.84	–	13.84	22.15	14.52	18.71
1st Shea part. R^2: Wages	–	0.379	–	0.379	0.319	0.351	0.303
1st part. F-stat: PM_{10}	–	–	49.46	19.05	13.84	13.84	13.84
1st Shea part. R^2: PM_{10}	–	–	0.351	0.301	0.379	0.379	0.379

Notes: 288 observations for each regression. The dependent variable is the log of average house prices (in 1000 CNY/m^2), except for column (7) where it is in levels. Standard errors in parentheses. In the IV panel, 'Y' indicates that the variable is instrumented, 'N' that it is not. LIML estimates reported. Critical values for the Cragg-Donald statistics (Stock and Yogo, 2005) for a maximal LIML size (5% significance level) are: 4.5 for Housing IV1, 6.5 for Housing IV2, and 3.7 for Housing IV3, 3.8 for Housing IV4, and 3.7 for Housing IV5-IV6. Estimation results of the first stage for IV estimation are reported in Appendix C Tables C1 and C2. *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

reduction of house prices by CNY 25.92 ($\simeq$USD 4.02). Finally, column (7) reports results from a linear specification, which again confirms that results are preserved, although the elasticity of house prices with respect to PM_{10} concentration is -1.42 at the mean of the sample, which is significantly larger than in the corresponding log-log specification.

4.2 Marginal willingness to pay estimates

We now combine evidence derived from intercity wages and house prices differential to illustrate the magnitude of our estimates. First, we provide about the aggregate wages impact across the working population in each city. Second, we compute the total change in value for the stock of

Figure 4: Aggregate marginal impact ($\mu g/m^3$) across Chinese cities (CNY 1 $\simeq$ USD 0.155, 2011)

(a) Aggregate change in wages across cities

(b) Aggregate change in house prices across cities
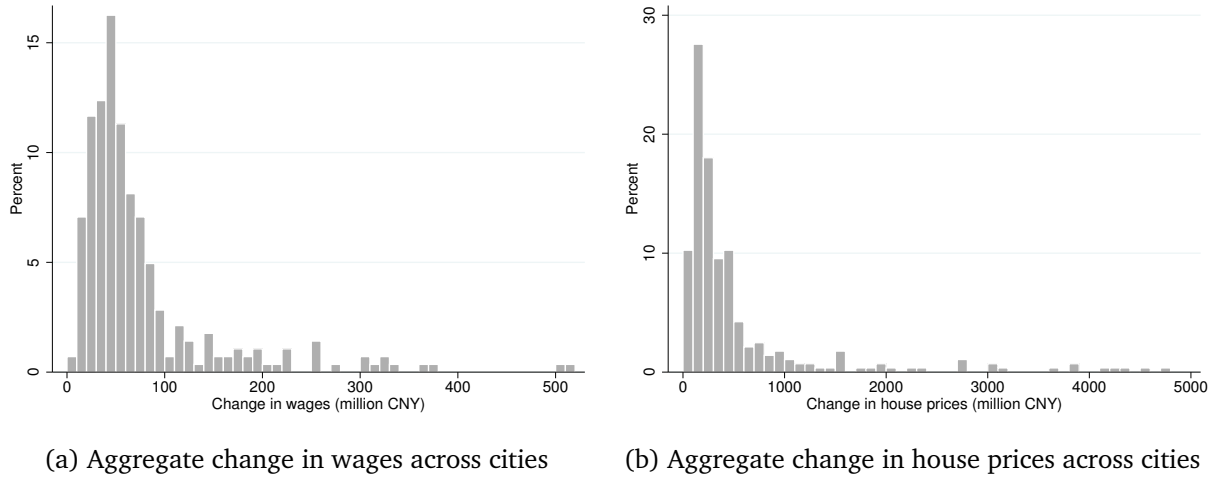
dwellings of each city. Finally, we compute MWTP estimates at the individual level using equation 5. We stress that this represents some 'back-of-the-envelope' calculations with the objective of assessing the relative importance of each component across cities.

Results are displayed in Figure 4. Panel (a) shows the distribution of aggregate changes in wages across cities, multiplying our elasticity estimate reported in Table 2 (column 4) by the ratio of pollution to wages in each city and the number of workers. The mean of the distribution is CNY 85.67 million ($\simeq$USD 13.28 million), while the median is CNY 53.29 million ($\simeq$USD 8.26 million). Panel (b) shows the distribution for the aggregate change in house prices, multiplying the elasticity estimate reported in Table 3 (column 4) by the ratio of pollution to wages in each city and total living area. The mean of the resulting distribution is CNY 761.71 million ($\simeq$USD 118.06 million) and the median is CNY 53.29 million ($\simeq$USD 40.81 million). An implication is that a large share of benefits associated with improved air quality are reflected in the labor market. As discussed in Roback (1982) however, aggregate changes in wages involves a transfer from firms to workers, so that it should not be counted as part of aggregate benefits associated with changes in air quality.

An alternative approach to illustrate the magnitude of our results is to compute MWTP for a reduction of PM$_{10}$ concentration at the individual level (see equation 5), evaluated at the mean of the sample. Specifically, the coefficient α_1 captures the partial MWTP associated with the change in wages of each individual, while the impact on housing price measured by β_1 needs to be multiplied by a measure of the per capita living space and then spread over the tenure length of the property.

Again for the purpose of illustration, we use an average per capita living area of of 30.63 m^2 (NBSC, 2010) and a tenure length of 70 years, which is in line with the fact that property law in China effectively grants residential land-use rights only for 70 years.[17] This suggest an average individual MWTP of CNY 208.90 (p-val.<0.01; 95% confidence interval: 137.85 – 279.93) for a unit reduction in PM$_{10}$ concentration.

5   Summary and conclusion

This paper has exploited variations in air quality induced by the River Huai policy to study compensating price differentials in China's labor and property markets. Using a regression discontinuity design, we estimate that the elasticity of wages with respect to PM$_{10}$ concentration is around 0.41, while that of house prices is -0.71. Estimates based on an alternative set of instruments capturing exogenous variation in PM$_{10}$ induced by sandstorms are almost identical, supporting the view that our results hold policy-relevant information on the benefits associated with air quality improvements.

Combining evidence from labor and property markets, we estimate average individual MWTP to be around CNY 208.90 for 1 μg/m^3 reduction in PM$_{10}$ concentration, or about USD 32.37. A key contribution of our work is to show that a significant share of economic value associated with air pollution is capitalized in the labor market. We note, however, that comparing economic value capitalized in wages and house prices is inherently difficult and requires assumptions about how these spread over time. Moreover, at the aggregate level changes in wages represent a transfer from firms to individuals, so that they should not be part of aggregate welfare calculations. Nevertheless, air pollution matters to both households and firms in China, and our work provides information towards the evaluation of air quality regulation in Chinese cities.

Aside from providing novel evidence about the economic value of air quality improvements, a key contribution of this paper has been the use of quasi-experimental variations to address endogeneity of local air pollution. Specifically, we have shown that endogeneity in air pollution is quantitatively more important than addressing endogeneity implied by simultaneity in wages

---

[17] According to the 1982 Constitution of China, land is publicly owned and private ownership is legitimately prohibited.

and house prices. Moreover, instrumenting for air pollution alleviates concerns about omitted variables in the hedonic equations, which is a recurrent issue for hedonic studies because it is virtually impossible to measure and control for all possible factors affecting market outcomes.

We close by highlighting the need for further research on the relationship between markets and environmental outcomes in China. While we provide one of the first attempts to relate air pollution to both labor and property markets, we have essentially ignored costs and constraints associated with migration, which implies that our MWTP estimate represent a lower bound to the welfare effect of air quality improvements (Bayer et al., 2009). Future research should addresses the broader issue of location decisions by firms and households in relation to local amenities and public policies. Moreover, cheap labor and booming real-estate market have fueled rapid economic growth in China for the past two decades, while at the same time awareness and concerns about environmental problems have been growing. Thus much work remains to be done on the dynamics of these processes and interactions with natural resource use and urban amenities in China.

Appendix A   Data sources and summary statistics

Variable	Definition	Mean	Std. dev.	p10	p25	p50	p75	p90	Source
Wages	Average yearly wages (1000 CNY/Year)	35.96	8.24	27.02	30.48	34.62	39.81	47.05	NBSC (2012a)
Workforce	Share of population: Age 18 to 64 (%)	74.44	4.36	68.59	71.38	74.72	77.55	79.56	NBSC (2010)
Male	Share of population: Male (%)	52.77	8.80	42.48	50.20	52.43	57.93	62.51	NBSC (2012b)
College	Share of population: College graduates (%)	8.10	4.85	3.93	5.02	6.73	8.98	15.07	NBSC (2010)
Service share	Share of GDP from services activities (%)	35.01	9.04	24.75	29.44	34.35	39.61	45.50	NBSC (2012a)
Unemployment	Unemployment rate (%)	3.30	0.74	2.20	2.80	3.50	3.90	4.20	NBSC (2012a)
Migrants	Share of population: Immigration (%)	21.64	32.78	6.40	8.62	13.24	22.39	42.46	NBSC (2010)
House prices	Average house prices (1000 CNY/m^2)	4.12	2.47	2.40	2.73	3.38	4.37	6.68	NBSC (2012a)
Piped water	Share of houses with piped water (%)	63.25	21.46	35.00	45.00	61.50	83.00	93.00	NBSC (2010)
Fitted kitchen	Share of houses with kitchen (%)	85.07	11.88	71.00	80.00	88.00	93.50	97.00	NBSC (2010)
Fitted toilets	Share of houses with toilet (%)	70.68	16.50	49.00	58.00	73.00	84.00	90.00	NBSC (2010)
Fitted shower	Share of houses with shower (%)	51.19	22.79	18.00	32.50	52.00	67.50	81.00	NBSC (2010)
Density	Population density (1000 persons /km^2)	0.47	0.54	0.08	0.18	0.34	0.62	0.87	NBSC (2012a)
Land supply	Traded land area (100 km^2)	0.13	0.42	0.02	0.04	0.07	0.12	0.22	CLRA (2012)
Coastal	Indicator equal to 1 if city is on the coast	0.09	0.28	0.00	0.00	0.00	0.00	0.00	NGCC (2012)
Universities	Number of universities per 1000 residents	0.01	0.01	0.00	0.01	0.01	0.01	0.02	NBSC (2012a)
Hospitals	Number of hospital beds per 1000 residents	3.71	1.46	2.26	2.68	3.43	4.41	5.65	NBSC (2012a)
PM$_{10}$	Average PM$_{10}$ concentration (μg/m^3)	77.44	19.11	53.00	64.00	78.00	91.00	100.00	MEP (2012)
Polluted days*	Number of days with API $\geq$ 100	21.94	21.59	0.00	4.50	17.00	34.00	48.00	NBSC (2012a)
Huai Policy	Indicator equal to 1 if city is included in the River Huai policy	0.43	0.50	0.00	0.00	0.00	1.00	1.00	NGCC (2012)
Latitude	Distance to latitude 33° (degree °)	-0.06	6.67	-9.27	-5.17	-0.56	4.71	8.80	NGCC (2012)
Sandstorm	Average distance to sandstorm origins (1000 km)	1.30	0.48	0.68	0.93	1.31	1.67	2.02	NGCC (2012)

Notes: 288 observations. * Definition of API is in Appendix C. Exchange rate for 2011: CNY 1 $\simeq$ USD 0.155.

Appendix B   Definition of the air pollution index (API)

The conversion method from pollutant concentrations to API values is available from the website of China's Ministry of Environmental Protection. First, average concentration of air pollutants are

Table B1: Air pollution index (API)

API Level	API Score	Description	Thresholds[1]		
			PM_{10}	SO_2	NO_2
I	$0 - 50$	Excellent	0.05	0.05	0.08
II	$51 - 100$	Good	0.15	0.15	0.12
III-1	$101 - 150$	Slight	0.35	0.8	0.28
III-2	$151 - 200$	Light			
IV-1	$201 - 250$	Moderate	0.42	1.6	0.565
IV-2	$251 - 300$	Heavy			
V	301-400	Severe	0.52	2.1	0.75
	401-500		0.6	2.62	0.94

Notes: [1] PM_{10}, SO_2 and NO_2 concentration are measured as average mg/m^3 per day. *Source*: China's Ministry of Environmental Protection (MEP).

reported from the monitoring stations in each city. According to the Automated Methods for Ambient Air Quality Monitoring (HJ/T $193-2005$), cities with a population of over 3 million people are required to use at least eight monitoring stations to measure urban air quality. Specifically, PM_{10}, SO_2 and NO_2 are measured as average per day.

The API is constructed from piecewise linear indexes computed for each pollutant. First concentrations are converted to pollutant-specific indexes:

$$I_k = \frac{I_{high} - I_{low}}{C_{high} - C_{low}} (C_k - C_{low}) + I_{low} \tag{B1}$$

where I_k is the pollution index of pollutant k, C_k is pollutant k's concentration, C_{low} is the concentration threshold (defined in Table B1), C_{high} is the concentration threshold that is $\geqslant C_k$, I_{low} is the index threshold corresponding to C_{low}, and I_{high} is the index threshold corresponding to C_{high}. The API then represents the highest value of those indexes:

$$API = max(I_{PM_{10}}, I_{SO_2}, I_{NO_2}) \tag{B2}$$

Note that the scale for each pollutant is non-linear, as is the final API value. In other words, an API of 100 is not equivalent to a doubling of pollution associated with an API of 50. Moreover, at the

boundary between API categories, there is a discontinuous jump of one API unit.

As shown in Table B1, an API of 100 is regarded as the threshold of air pollution that can lead to non-negligible health consequences. Thus days with an API under 100 are defined as a "Blue Sky" days (i.e. days in which air pollution poses little or no risk to human health). When the API is between 101 and 200, air pollution may generate slight irritations to breathing and lead to heart diseases. When API is between 201 and 300, air pollution will aggravate the symptoms of patients with cardiac and lung diseases, even healthy people will be noticeably affected. When API is above 300 air is severely polluted and may cause irritations and other symptoms. People are advised to avoid outdoor activities.

Appendix C   First stage regression results

Table C1: First stage regressions for pollution in the wages equation

	Wages IV2 (1)	Wages IV3 (2)	Wages IV4 (3)	Wages IV5 (4)	Wages IV6 (5)
Huai Policy	0.07* (0.04)	0.08* (0.04)	–	0.62*** (0.17)	7.99*** (2.88)
Latitude	0.01*** (0.004)	0.01*** (0.004)	–	0.06*** (0.02)	0.83*** (0.30)
Latitude^2	-0.002** (0.0003)	-0.002*** (0.0003)	–	-0.01*** (0.001)	-0.14*** (0.02)
Sandstorm	–	–	0.68*** (0.15)	–	–
Sandstorm^2	–	–	-0.36*** (0.06)	–	–
Constant term	4.28*** (0.50)	4.00*** (0.51)	4.30*** (0.48)	-0.93 (2.19)	43.43 (35.68)
Housing controls	N	Y	Y	Y	Y
Wages controls	Y	Y	Y	Y	Y
R^2	0.42	0.46	0.47	0.45	0.46

Notes: 288 observations for each regression. Standard errors in parentheses. In the panel indicating the set of controls, 'Y' means that the set of controls is included, 'N' that it is not. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

Table C2: First stage regressions for house prices in the wages equation

	Wages IV1 (1)	Wages IV3 (2)	Wages IV4 (3)	Wages IV5 (4)	Wages IV6 (5)
Piped water	0.002**	0.003***	0.003**	0.004***	0.01**
	(0.001)	(0.003)	(0.001)	(0.001)	(0.01)
Fitted kitchen	-0.0005	0.003	-0.002	0.002	0.01
	(0.002)	(0.002)	(0.002)	(0.002)	(0.01)
Fitted toilets	0.003**	0.003*	0.003**	0.003*	0.01*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.01)
Fitted shower	0.004***	0.001	0.003***	0.001	-0.01*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.01)
Density	0.04	0.06	0.04	0.06	0.63***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.24)
Land supply	-0.03	-0.02	-0.03	-0.02	-0.42**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.20)
Coastal	0.23***	0.23***	0.23***	0.23***	1.62***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.33)
Universities	-4.51	-5.40	-4.48	-5.40	-62.23***
	(3.46)	(3.39)	(3.43)	(3.39)	(18.86)
Hospitals	-0.06***	-0.06***	-0.06***	-0.06***	-0.50***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.11)
Constant term	-0.38	-0.38	-0.43	-0.38	-4.27
	(0.60)	(0.58)	(0.57)	(0.58)	(3.21)
Pollution controls	N	Y	Y	Y	Y
Wages controls	Y	Y	Y	Y	Y
R^2	0.70	0.72	0.71	0.72	0.74

Notes: 288 observations for each regression. Standard errors in parentheses. In the panel indicating the set of controls, 'Y' means that the set of controls is included, 'N' that it is not. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

Table C3: First stage regressions for pollution in the house prices equation

	Housing IV2 (6)	Housing IV3 (7)	Housing IV4 (8)	Housing IV5 (9)	Housing IV6 (10)
Huai Policy	0.07*	0.08*	–	0.62***	7.99***
	(0.04)	(0.04)		(0.17)	(2.88)
Latitude	0.02***	0.01***	–	0.06***	0.83***
	(0.004)	(0.004)		(0.02)	(0.30)
Latitude^2	-0.002***	-0.002***	–	-0.01***	-0.14***
	(0.0003)	(0.0003)		(0.001)	(0.02)
Sandstorm	–	–	0.68***	–	–
			(0.15)		
Sandstorm^2	–	–	-0.36***	–	–
			(0.06)		
Constant term	4.54***	4.00***	4.30***	-0.93	43.43
	(0.30)	(0.51)	(0.48)	(2.19)	(35.68)
Housing controls	Y	Y	Y	Y	Y
Wages controls	N	Y	Y	Y	Y
R^2	0.44	0.46	0.47	0.45	0.46

Notes: 288 observations for each regression. Standard errors in parentheses. In the panel indicating the set of controls, 'Y' means that the set of controls is included, 'N' that it is not. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

Table C4: First stage regressions for wages in the house prices equation

	Housing IV1 (1)	Housing IV3 (2)	Housing IV4 (3)	Housing IV5 (4)	Housing IV6 (5)
Workforce	-0.005	-0.01*	-0.01*	-0.01*	-0.26*
	(0.004)	(0.004)	(0.004)	(0.005)	(0.15)
Male	0.0002	-0.001	-0.0004	-0.002	-0.05
	(0.002)	(0.002)	(0.002)	(0.002)	(0.07)
College	0.03***	0.03***	0.03***	0.03***	1.15***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.13)
Service share	0.001	0.002	0.002	0.001	0.07
	(0.001)	(0.001)	(0.001)	(0.001)	(0.04)
Unemployment	-0.005	-0.01	-0.004	-0.01	-0.40
	(0.01)	(0.01)	(0.01)	(0.01)	(0.50)
Migrants	0.001**	0.002***	0.002***	0.002***	0.06***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.02)
Constant term	3.67***	4.11***	4.01***	4.11***	55.83***
	(0.36)	(0.35)	(0.34)	(0.35)	(12.58)
Pollution controls	N	Y	Y	Y	Y
Housing controls	Y	Y	Y	Y	Y
R^2	0.57	0.60	0.58	0.60	0.64

Notes: 288 observations for each regression. Standard errors in parentheses. In the panel indicating the set of controls, 'Y' means that the set of controls is included, 'N' that it is not. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

References

Albouy, D. (2015) "What are cities worth? Land rents, local productivity, and the total value of amenities," *Review of Economics and Statistics*, forthcoming.

Almond, D., Y. Chen, M. Greenstone, and H. Li (2009) "Winter heating or clean air? Unintended impacts of China's Huai River policy," *American Economic Review*, 99 (2), pp. 184 – 190.

Andrews, S. Q. (2008) "Inconsistencies in air quality metrics: 'Blue Sky' days and PM_{10} concentrations in Beijing," *Environmental Research Letters*, 3, p. 4009.

Bayer, P., N. Keohane, and C. Timmins (2009) "Migration and hedonic valuation: The case of air quality," *Journal of Environmental Economics and Management*, 58, pp. 1–14.

Black, S. (1999) "Do better schools matter? Parental valuation of elementary education," *Quarterly Journal of Economics*, 114 (2), pp. 577 – 599.

Blomquist, G., M. Berger, and J. Hoehn (1988) "New estimates of quality of life in urban areas," *American Economic Review*, 78 (1), pp. 87 – 107.

Chay, K. and M. Greenstone (2003) "The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession," *Quarterly Journal of Economics*, 118, 3, pp. 1121–1167.

Chay, K. Y. and M. Greenstone (2005) "Does air quality matter? Evidence from the housing market," *Journal of Political Economy*, 113, 2, pp. 376–424.

Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2012) "Gaming in air pollution data? Lessons from China," *The BE Journal of Economic Analysis & Policy*, 12, 3, p. 424.

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

CLRA (2012) "China land and resources almanac." Editorial office of China Land and Resources Almanac.

Cropper, M., L. Deck, and K. McConnell (1988) "On the choice of functional form for hedonic price functions," *Review of Economics and Statistics*, 70, pp. 668 – 675.

Currie, J., E. A. Hanushek, E. M. Kahn, M. Neidell, and S. G. Rivkin (2009) "Does pollution increase school absences?" *The Review of Economics and Statistics*, 91, 4, pp. 682–694.

Davis, L. (2004) "The effect of health risk on housing values: Evidence from a cancer cluster," *American Economic Review,* 94, pp. 1693 – 1704.

D'Ippoliti, D., F. Forastiere, C. Ancona, N. Agabity, D. Fusco, P. Michelozzi, and C. Perucci (2003) "Air pollution and myocardial infarction in Rome: A case-crossover analysis," *Epidemiology*, 14, pp. 528 – 535.

Epple, D. and H. Sieg (1999) "Estimating equilibrium models of local jurisdictions," *Journal of Political Economy*, 1007 (4), pp. 645 – 681.

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

Glaeser, E. and K. Tobio (2008) "The rise of the sunbelt," *Southern Economic Journal*, 74 (3), pp. 610 – 643.

Graff Zivin, J. and M. Neidell (2012) "The impact of pollution on worker productivity," *American Economic Review*, 102, 7, pp. 3652–73.

Graves, P. and D. Waldman (1991) "Multimarket amenity compensation and the behavior of the elderly," *American Economic Review*, 81 (5), pp. 1374 – 1381.

Guan, D. and Z. Liu (2014) "Tracing back the smog: Source analysis and control strategies for $PM_{2.5}$ pollution in Beijing-Tianjin-Hebei." Beijing: China Environment Press.

Gyourko, J., M. Kahn, and J. Tracy (1999) "Quality of life and environmental comparaisons," in E. Mills and P. Cheshire eds. *The Handbook of Applied Urban Economics*, pp. 1413 – 1454: Amsterdam: North Holland.

Gyourko, J., C. Mayer, and T. Sinai (2010) "Dispersion in house price and income growth across markets: Facts and theories," in E. Glaeser ed. *Agglomeration Economics*, pp. 67 – 104: Cambridge MA: National Bureau of Economic Research.

Hanna, R. and P. Oliva (2015) "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City," *Journal of Public Economics*, 122, pp. 68–79.

Hoehn, J. P., M. C. Berger, and G. C. Blomquist (1987) "A hedonic model of interregional wages, rents and amenity values," *Journal of Regional Science*, 27, 4, pp. 605–620.

Kennan, J. and J. R. Walker (2011) "The effect of expected income on individual migration decisions," *Econometrica*, 79, 1, pp. 211–251.

Kuminoff, N., C. Parmeter, and J. Pope (2010) "Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?" *Journal of Environmental Economics and Management*, 60, pp. 145 – 160.

Kuminoff, N., V. Smith, and C. Timmins (2013) "The new economics of equilibrium sorting and policy evaluation using housing markets," *Journal of Economic Literature , 51(4), 1007İC1062*, 51 (4), pp. 1007 – 1062.

Linden, L. and J. Rockoff (2008) "Estimates of the impact of crime risk on property values from Megan's laws," *American Economic Review*, 98, pp. 1103 – 1127.

McKendry, I., J. Hacker, R. Stull, S. Sakiyama, D. Mignacca, and K. Reid (2001) "Long-range transport of Asian dust to the lower Fraser Valley, British Columbia, Canada," *Journal of Geophysical Research*, 106 (16), pp. 18361 – 18370.

MEP (2012) "2011 report on the state of the environment of China." China's Ministry of Environmental Protection (MEP), China Environmental Science Press.

NBSC (2010) "The sixth national population census of china." National Bureau of Statistics of China.

——— (2012a) "China city statistical yearbook." National Bureau of Statistics of China, China Statistics Press.

——— (2012b) "China statistical yearbook for regional economy." National Bureau of Statistics of China, China Statistics Press.

NGCC (2012) "National geomatics center of china." web: www.ngcc.cn.

OECD (2014) *The Cost of Air Pollution: Health Impacts of Road Transport*: OECD Publishing, Paris.

Pope III, C. A. and D. W. Dockery (2006) "Health effects of fine particulate air pollution: Lines that connect," *Journal of the Air & Waste Management Association*, 56, 6, pp. 709–742.

Pope, J. (2008) "Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise," *Journal of Urban Economics*, 63, pp. 498 – 516.

Qiao, L., J. Cai, H. Wang, W. Wang, M. Zhou, S. Lou, R. Chen, H. Dai, C. Chen, and H. Kan (2014) "$PM_{2.5}$ constituents and hospital emergency-room visits in Shanghai, China," *Environmental Science & Technology*, 48, pp. 10406–10414.

Ravetti, C. (2015) "A dragon eating its own tail: Biased pollution signals and households' response in china." CIES Working paper.

Roback, J. (1982) "Wages, rents, and the quality of life," *Journal of Political Economy*, 90, 6, pp. 1257–1278.

Rosen, S. (1974) "Hedonic prices and implicit markets: Product differentiation in pure competition," *Journal of Political Economy*, 82, pp. 34–55.

——— (1979) "Wage-based indexes of urban quality of life," in P. Mieszkowski and M. Straszheim eds. *Current Issues in Urban Economics*, pp. 74 – 104: Baltimore: Johns Hopkins University Press.

Stock, J. H. and M. Yogo (2005) "Testing for weak instruments in linear IV regression," in D. Andrews and J. Stock eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*: Cambridge: Cambridge University Press, pp. 80 – 108.

Tiebout, C. M. (1956) "A pure theory of local expenditures," *Journal of Political Economy*, 64 (5), pp. 416 – 424.

Tsai, S., W. Goggins, H. Chiu, and C. Yang (2003) "Evidence for an association between air pollution and daily stroke admissions in Kaohsiung, Taiwan," *Stroke*, 34 (11), pp. 2612 – 2616.

WHO (2006) "WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide." Geneva: World Health Organisation (WHO).

Zhang, Q. and R. Crooks (2012) "Toward an environmentally sustainable future: Country environmental analysis of the People's Republic of China." Mandaluyong City, Philippines: Asian Development Bank (ADB).

Zheng, S., Y. Fu, and H. Liu (2006) "Housing-choice hindrances and urban spatial structure: Evidence from matched location and location-preference data in Chinese cities," *Journal of Urban Economics*, 60, pp. 535 – 557.

——— (2009) "Demand for urban quality of living in China: Evolution in compensating land-rent and wage-rate differentials," *The Journal of Real Estate Finance and Economics*, 38, pp. 194–213.

Zheng, S., M. E. Kahn, and H. Liu (2010) "Towards a system of open cities in China: Home prices, FDI flows and air quality in 35 major cities," *Regional Science and Urban Economics*, 40, pp. 1–10.

Zheng, S., J. Cao, M. E. Kahn, and C. Sun (2014) "Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities," *The Journal of Real Estate Finance and Economics*, 48, pp. 398–414.

Zheng, S. and M. Kahn (2008) "Land and residential property markets in a booming economy: New evidence from Beijing," *Journal of Urban Economics*, 63 (2), pp. 743 – 757.

———— (2013) "Understanding China's urban pollution dynamics," *Journal of Economic Literature*, 51 (3), pp. 731 – 772.