Surface Water Quality and Infant Mortality in China

GUOJUN HE
Hong Kong University of Science and Technology

JEFFREY M. PERLOFF
University of California, Berkeley, and the Gianini Foundation

I. Introduction

China's rapid industrialization led to severe water pollution in many areas due to massive industrial wastewater discharges and extensive use of agricultural fertilizer. Even though about 60%–70% of river water is unsafe for human consumption (World Bank 2006), many people in poor areas rely on the surface water systems for daily use including drinking. Consequently, this water pollution endangers the health of the rural population, especially that of infants. Using data on surface water quality from monitoring sites in 19 provinces and county-level infant mortality rate data from the 2000 census, we estimate the effects of surface water quality on the infant mortality rate in China.

We hypothesize that if surface water becomes slightly degraded, people do not notice the pollution and continue consuming it. Consequently, the infant mortality rate initially rises as water pollution increases. However, as the pollution gets worse, people begin to notice the pollution using visual and other clues and reduce their consumption of surface water. As the degree of pollution becomes very pronounced and hence very apparent, the infant mortality rate falls.

We discuss the literature on the effects of surface water quality on health in Section II and Chinese surface water quality in Section III. In Section IV, we describe the data set, define the key variables, and provide summary statistics. We analyze whether we are likely to face a sample selection problem and describe an ordered-probit model to deal with that issue in Section V. In Sec-

We are extremely grateful to Michael Anderson and William Dow for extensive comments on an earlier version of this article when it was part of the first author's dissertation. We also thank Elisabeth Sadoulet, Meredith Fowlie, David Roland-Holst, Maximilian Auffhammer, Jeremy Magruder, Yuanli Liu, Jing Ma, and Maoyong Fan for their constructive suggestions. We are grateful to the seminar participants in the Harvard School of Public Health and the Department of Agricultural and Resource Economics at University of California, Berkeley. We thank Shuming Bao, Gordon Liu, and Avraham Ebenstein for their help in the process of collecting the data. Qilun Zhou provided superior research assistance. Any errors are the authors'.

Electronically published August 11, 2016

tion VI, we present our estimates of the effects of surface water quality on the infant mortality rate using an ordered-probit selection model. We draw conclusions in Section VII.

II. The Effects of Water Pollution on Health

The health effects of water pollution have been an important research topic in environmental science, epidemiology, environmental economics, and health economics for centuries. The earliest research in this area dates to the 1850s (see Freedman 1991). Using a natural experiment based on the distribution of water in London, John Snow demonstrated that unsanitary water caused cholera outbreaks. He found that the death rate for dirty-water users was over eight times higher than that for the clean-water users.

Many recent studies reported various connections between water pollution and diseases and other public health measures. Some studies focused on water pollution and waterborne diseases, such as typhoid (Cutler and Miller 2005) and diarrhea (Jalan and Ravallion 2003). Other studies explored the relationship between water pollution and cancer, such as Cantor (1997), Davis and Masten (2004), Chen et al. (2005), and Ebenstein (2012).

Much of the research focused on infant and child mortality. Galiani, Gertler, and Schargrodsky (2005) found that privatizing water service improves water quality and reduces child mortality. Merrick (1985), Lavy et al. (1996), and Lee, Rosenzweig, and Pitt (1997) found positive associations between water quality and infant health. Greenstone and Hanna (2014) investigated the effects of air quality and water quality on the infant mortality rate in India and found no significant relationship between the infant mortality rate and water quality. Brainerd and Menon (2012) reported a negative effect of fertilizer agrichemicals in water on infant and child health in India. Currie et al. (2013) found that contaminated drinking water has large and statistically significant effects on birth weight and gestation of infants born to less educated mothers.

Several studies have examined people's pollution avoidance behavior (Mansfield, Johnson, and Van Houtven 2006; Shimshack, Ward, and Beatty 2007; Neidell 2009; Moretti and Neidell 2011; Zivin, Neidell, and Schlenker 2011). For example, Neidell (2009) found that when smog alerts were issued, attendance at major outdoor facilities in Los Angeles decreased significantly. Zivin et al. (2011) found that bottled water sales significantly increased when consumers were informed about tap water quality deteriorations (or violations) in northern California and Nevada. These findings suggest that the provision and public dissemination of information about pollution can encourage the public to engage in avoidance behaviors to decrease exposure and minimize health risks.

III. Surface Water Quality in China

In China, the overall surface water quality is graded on the basis of chemical pollutant indicators, including the pH value and the concentrations (measured by milligrams/liter) of dissolved oxygen, biochemical oxygen demand, ammonia, and nitrogen. The overall surface water quality is graded on a 6-degree scale, where type I water is the best quality water, and type VI is the worst.

According to the China Ministry of Water Resources, type I water is an "excellent" source of potable water. Type II water is a "good" source of potable water. Type III water is "fair." Because type II and III water may have pathogenic bacteria and parasites ova, drinking that water may cause diseases. Type II and III water should be purified and treated (such as by boiling) before drinking. Type IV water is polluted and unsafe to drink without advanced treatment, which is only possible at water supply plants. Type V is seriously polluted and can never been used for human consumption. Type VI water is called "worse than Type V water," and any direct contact with it is harmful to humans.[1]

Individuals can easily distinguish clean water from heavily polluted water, such as types V and VI. Very polluted water is murky and smelly and may have algal blooms on the surface. Distinguishing "excellent" water from "good" or "fair" water is more challenging.

Not all the harms from consuming polluted water are immediately apparent. Some effects occur at once, while others appear only after toxins have accumulated in the body. For example, consuming contaminated water can cause malaria outbreaks within days, but it may take decades for water pollution to cause cancers. In this study, we focus on infant deaths that occur within a year.

IV. Data and Summary Statistics

China established a nationwide water quality monitoring system in the 1980s. Each year, the Ministry of Water Resources publishes the China National Water Resource Yearbook, which provides water quality information for major lakes, rivers, and reservoirs. Many provinces also publish province-level water-body quality measures. The national and provisional publications are the sources of water quality data used in this article.

We identified each water quality monitoring site's location and matched it with the 2000 census data at the county level.[2] Our sample includes 461 coun-

[1] See table A1 (tables A1–A4 available online) for the limit values of water quality indicators.

[2] There are five administrative levels in China. The highest level is the provincial level, which include provinces, autonomous regions, direct-controlled municipalities, and special administrative regions (Hong Kong and Macao). The second highest level of government is the prefecture level that includes prefectures, autonomous prefectures, prefecture-level cities, and leagues. Next is the county level, which includes counties, autonomous counties, county-level cities, and city districts. Below

ties in 19 provinces. The sampled provinces include Anhui, Beijing, Chongqing, Fujian, Guangxi, Guizhou, Hianan, Hebei, He'nan, Jiangsu, Jiangxi, Liaoning, Ningxia, Shandong, Shanghai, Sichuan, Tianjin, Yunnan, and Zhejiang. Guangxi and Ningxia are in the Autonomous Region; Beijing, Chongqing, Shanghai, and Tianjin are directly controlled municipalities; and the rest are governed at the provincial level. Northern provinces are relatively undersampled because we could not find water quality data for several.

The distribution of the six types of water quality is presented in table 1. Nearly half (47%) of the rivers and other bodies of water in our sample are seriously polluted (types IV, V, and VI).

The infant mortality rate is the annual number of deaths among infants (less than 1 year old) per thousand live births. The overall Chinese infant mortality rate has been decreasing at a rapid pace over the past 40 years, from about 150 per thousand in the 1960s to around 20 per thousand in the 2000s.

Neonatal death is the major component of infant deaths. About 60%–70% of infant deaths occurred within the first month after they were born. Most of the deaths during the neonatal period are due to endogenous causes (inherited defects), such as congenital anomalies, gestational immaturity, birth complications, and other physiological problems. Lack of proper care during the pregnancy can cause neonatal death. For example, if pregnant women drink polluted water, the neonatal mortality may rise. The postneonatal mortality rate describes the death rate of infants age 1 month to 1 year. The vast majority of postneonatal deaths are from exogenous causes, such as injuries and environmental and nutritional factors, especially as they interact with infectious disease like gastroenteritis and pneumonia. Water pollution may cause both neonatal and postneonatal mortality.

The infant mortality rate summary statistics are presented in table 2. In our sample, the average the infant mortality rate across counties is 19.2 per thousand. The variation in infant mortality rates is large, with a standard deviation of 15.8. The distribution of infant mortality rate is skewed to the left, with a few counties out in the tail with considerably higher infant mortality rates than the average. The lowest infant mortality rate in our data is less than 1 per thousand, while the highest infant mortality rate is 81.2 per thousand.

In general in China, female infants are more likely to die than are male infants. The female infant mortality rate is about 22 per thousand, and the male infant mortality rate is about 17 per thousand. The difference is striking because male infants are usually more vulnerable and thus more likely to die than

---

the county is the township level and the (informal) village level. The census data are not available for these last two levels.

TABLE 1
WATER QUALITY DISTRIBUTION IN THE SAMPLE

	Type I	Type II	Type III	Type IV	Type V	Type VI	Total
Frequency	23	132	90	60	59	97	461
Percentage	4.99	28.63	19.52	13.02	12.8	21.04	100

Source. Ministry of Water Resources of the People's Republic of China (2010).

female infants on purely medical grounds. This difference is consistent with the popular argument that rural Chinese people prefer boys to girls, so they invest more on male infants' health. The variance of the female infant mortality rate is also higher than that of the male infant mortality rate.

To isolate the effect of water pollution on health, we include a set of control variables from the 2000 census of China: the percentage of the population that is nonagricultural, illiteracy rate, average rooms per home, and per capita housing area (square meters). We also include a few socioeconomic variables in the 2000 China Statistical Yearbook: per capita GDP (Chinese yuan), per capita government expenditure, and the number of beds in medical institutions per 10,000 people. We show the summary statistics of these variables in table 3.

The nonagricultural share is the population that lives in nonagricultural areas (including the mobile population) of the total population. Urbanization, which reduces the nonagricultural share, has both positive and negative health effects, and the net impact on population is not obvious (see, e.g., Van de Poel, O'Donnell, and Van Doorslaer 2009). It is generally believed that the positive consequences of urbanization outweigh the negative ones for infants' health. Urbanization is associated with better sanitation and medical treatments and easier access to tap water and infant care, all of which play important roles in improving infant health. So, we expect a negative association between the percentage of the nonagricultural population and the infant mortality rate.

The illiteracy rate variable is the share of the total population over 15 who are illiterate (have not completed an elementary school education). Many stud-

TABLE 2
CHINESE INFANT MORTALITY RATES (PER THOUSAND)

			Quantile		
	Mean	SD	25%	50%	75%
All	19.23	15.82	8.64	14.62	25.31
Male	16.97	13.44	7.89	13.63	21.81
Female	21.97	20.42	8.61	15.17	29.47

Source. China 2000 census.

TABLE 3
SUMMARY STATISTICS OF THE EXPLANATORY VARIABLES

Variable	Mean	SD	Min	Max
Precipitation (100 mm)	11.08	5.07	1.15	24.60
Wastewater dumping	19.29	11.34	.42	47.38
Nonagricultural population (%)	30.16	26.55	4.38	97.40
Illiteracy rate	9.52	6.15	1.57	48.34
Rooms per household	2.60	.58	1.50	4.92
Housing area per capita (m^2)	23.47	6.48	9.47	45.64
GDP per capita (¥1,000)	9.25	8.20	.82	55.84
Government expenditure per capita	.78	.93	.02	9.49
Hospital beds per 10,000 people	32.32	23.39	5.98	161.58

Sources. China 2000 census and National Bureau of Statistics of China (2000).

ies have found that infants are less likely to die, the more educated are their parents, especially their mothers (Delgado et al. 2002; Behrman et al. 2003; Basu and Stephenson 2005; Frost, Forste, and Haas 2005; Miguel 2005; Boyle 2006; Cowell 2006; Cutler and Lleras-Muney 2006).

Poor housing and overcrowding are negatively associated with infant health (Martin 1967; Brennan and Lancashire 1978; Victora et al. 1988). We expect that as either of our measures of living conditions—the average number of rooms per household and the per capita housing area (square meters)—increases, the infant mortality rate falls.

We expect per capita county GDP to be negatively correlated with the infant mortality rate. We use per capital government expenditure to approximate the local government's investments on social welfare programs, such as public health insurance, sanitation maintenance, tap water provision, waste management, and pollution treatment. We expect it to be negatively correlated with the infant mortality rate.

We use the number of beds per 10,000 people in medical institutions as a measure of the availability of medical treatment. In regions where hospitals are readily available, the infant mortality rate should be lower.

The reliability of the data collected by the Chinese government is often questioned by academic researchers. Local governments in China have been criticized for hiding news from the public about water pollution accidents. Consequently, we examined the consistency of the data in three ways.

In China, multiple agencies (e.g., the Ministry of Water Resource, the Environment Protection Agency, and the Center for Disease Control and Prevention) collected data on surface water quality. Since we do not know the raw data sources of different water quality reports, we first checked for consistency between reports of water quality data between national water resources and provincial water resources. We found that 98% of these paired reports are consistent, and the few differences are all within one water quality level. In the few cases in which they differ, we rely on the data from the provincial water resource reports.[3]

Second, we compared a subsample of the our 2004 water quality data with 2004 water quality data provided by the World Bank, which are used in Ebenstein (2012). The data are almost always identical for comparable monitoring sites.

Third, several monitoring sites' data are provided by River Basin Water Quality Reports. Again, we found no substantial difference between these and other sources.

Thus, the various reports on water quality are consistent. In the past several years, the government has trumpeted its efforts to promote the transparency of surface water quality information. For example, the central government started to release weekly water quality reports for 100 national monitoring sites to the public in 2004, and it started to publicize real-time water quality data in 2009.

V. Sample Selection

To estimate the effects of water pollution on health consistently, we need to avoid sample selection bias due to endogenous migration. Ebenstein (2012) suggests that China provides an ideal context to estimate the health effects of water pollution because of the household registration system (Hukou). This system prevents people from moving from rural to urban areas. It also makes it relatively difficult to move within rural and urban areas. A variety of benefits, such as health care and social security, are associated with the household registration system. If the system were to prevent people from migrating, we could treat people's location and thus the water quality they face as exogenous.

Although this argument may have held in the past, the share of people migrating increased by more than an order of magnitude between 1982, when it was 0.66% according to the census, and 2000 when it was 7.9%. Given that many rural migrant workers are of childbearing age, if the migration decision is correlated with water pollution levels, we may face a selection problem.

We address the potential sample selection problem using a Heckman-type selection model. In the first step, we estimate the relationship between the endogenous water pollution on the instrument variables. Conditional on the estimates from the first step, we estimate the effects of water pollution on the infant mortality rate in the second step.

We use wastewater dumping and precipitation as instruments for surface water quality. Industrial wastewater dumping degrades surface water quality. We use the amount of untreated wastewater discharged into the water system at the prefectural level as an instrument for the water quality in a given county. We treat the prefecture-level wastewater dumping as an exogenous variable at

[3] If we drop these observations, we obtain the same qualitative results reported below.

the county level because a county takes the total amount of discharged wastewater as given.[4]

Precipitation is one of the most important factors that influence surface water quality. Where rainfall is heavy, surface water quality is typically good. Precipitation affects surface water quality through two primary channels. First, rain directly dilutes the concentration of water pollutants and thus improves water quality. Second, rain causes the river to flow faster, which carries away water pollutants quickly and makes the river less prone to eutrophication (Zhong et al. 2005).

China has seven major river basins: Changjiang (Yangtze), Huang, Zhujiang, Huai, Songhuajiang, Liao, and Hai. Surface water quality in the southern river basins is better than that in the northern river basins because southern China has more rain. In our sample, the average yearly precipitation in the northern provinces (Beijing, Tianjin, Hebei, Liaoning, and Ningxia) is 458 millimeters, with a standard deviation of 148 millimeters. In contrast, the average precipitation in the rest of the provinces is 1,286 millimeters with a standard deviation of 417 millimeters.

In figure 1, we show the proportion of the river segments that were seriously polluted (types IV, V, and VI) from 1991 to 2005. Roughly 60%–70% of rivers in northern China (Huai River, Hai River, Songhuajiang River, Liao River) were severely polluted. In contrast, in southern river basins, such as Changjiang and Zhujiang, the proportion of polluted river segments was much lower.

Although precipitation has a large effect on surface water quality, for it to be a valid instrument, it must affect surface water quality but not be correlated with the error term in our infant mortality rate equation. Our primary concern is that rainfall fluctuations may affect infant health through other channels apart from surface water effects. For example, some studies have argued that an increase in rainfall increases agricultural production, lowering food prices, and increasing nutrient intake and hence health.[5] Although this argument sounds plausible, it is unlikely to be a problem in our study for two reasons.

First, small shocks in agricultural production induced by rainfall variations do not cause infants to die, unless the households are extremely poor and heavily depend on the food or the resulting income to survive. Recent studies that investigate the relationship between rainfall and health outcomes found a mixture

[4] A prefecture usually includes dozens of counties. Ideally, we would like to use the total wastewater discharged in all other counties within the same prefecture as the instrument variable for a particular county. Unfortunately, we cannot obtain county-level wastewater data.

[5] The link between rainfall and agricultural income has been investigated in the literature. For example, Levine and Yang (2006) showed that more rainfall increases rice output in Indonesia. Duflo and Udry (2004) looked at how men and women's income and spending change when the yields of different crops vary because of their different sensitivity to rainfall.

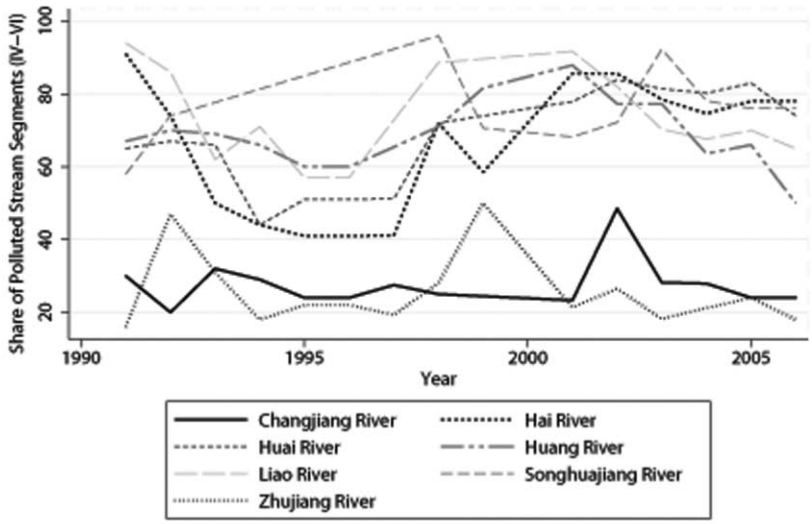

Figure 1. Polluted segment of the main water system in China

of positive, negative, and nonsignificant results.[6] Studies that found a significant effect of rainfall fluctuations (within its usual range) on the infant mortality rate always focused on poor and arid or semiarid regions. In contrast, all the counties in our sample are located in relatively water-abundant areas, where the irrigation process mainly relies on the surface water system instead of rainfall. So it is unlikely that rainfall affects the infant mortality rate through its impact on agricultural production.

When we regress agriculture production per capita on precipitation, the estimated coefficient on precipitation is -24.5 (a 100 millimeter increase in precipitation is associated with a ¥24.5 decrease in per capital agricultural production) with a *t*-statistic of -0.99.[7] Thus, we conclude that rainfall is not correlated with agriculture production in our sample.

Second, even in arid and semiarid regions, where water scarcity reduces agricultural production, the primary channel by which rainfall affects the infant mortality rate is not through its impact on agriculture production. For example, Bhalotra (2010) found that in rural areas of India income shocks have significant negative effects on the infant mortality rate. However, when he controlled for rainfall, the income effect did not change, suggesting that the effect of aggregate income on the rural infant mortality rate is not driven by agriculture income. In Rocha and Soares's (2011) study, irrespective of how they introduced

[6] The details of these studies are in table A2.
[7] Regression results are reported in table A3.

agricultural production into the regression, the impact of rainfall variations on health at birth was not affected, also suggesting that agricultural income does not affect the infant mortality rate. Instead, they found that the negative impact of rainfall on the infant mortality rate would be greatly reduced by using piped water. Thus, if people no longer use contaminated water, rainfall itself would not harm infant health.

Ebenstein (2012) investigated the effects of water pollution on digestive cancer in China and found that precipitation has only a very weak relationship with other disease mortality rates except for digestive and lung cancer rates, and he found almost no relationship between rainfall and cancer rates in areas with high usage of tap water. His results also showed that the effect of rainfall variations on people's health is primarily due to its impact on the surface water.

In this study, if we include per capita agriculture production as a control variable, we find that it is not statistically significant and that the estimates of the effect of water quality on the infant mortality rate are unchanged. Thus, we believe that precipitation is a valid instrument and that rainfall affects infant health through its impact on surface water quality. We also verified that in 1999 and 2000 no catastrophic natural disaster such as severe droughts or floods occurred in the sampled counties.

We use a two-step sample selection model to estimate the infant mortality rate conditional on surface water quality. In the first step, we estimate the water quality using an ordered probit model. If people move in response to the local water quality, the water pollution level is endogenous.

Because the number of observations of type I water is relatively small (table 1), we aggregate type I and type II water into a single group, which leaves us with five water quality categories. Each county i has water quality in category 1 (type I and type II), category 2 (type III), category 3 (type IV), category 4 (type V), or category 5 (type VI), ranging from good quality to bad. We estimate the water quality using an ordered probit:

$$\text{water}_i^* = \alpha' \boldsymbol{z}_i + u_i \tag{1}$$

$$\text{water}_i = \begin{cases} 1 & \text{if } -\infty < \text{water}_i^* \le \mu_1, \\ 2 & \text{if } \mu_1 < \text{water}_i^* \le \mu_2, \\ 3 & \text{if } \mu_2 < \text{water}_i^* \le \mu_3, \\ 4 & \text{if } \mu_3 < \text{water}_i^* \le \mu_4, \\ 5 & \text{if } \mu_4 < \text{water}_i^* < \infty, \end{cases} \tag{2}$$

where water_i^* is the unobserved latent selection variable (actual water quality), $\boldsymbol{z}_i$ is a set of variables that affect water quality, u_i is a normal disturbance; water_i is the observed 5-degree water quality scale, and the unobserved cutoffs satisfy $\mu_1 < \mu_2 < \mu_3 < \mu_4$.

The infant mortality rate IMR_i is a linear function of the independent variables $\boldsymbol{x}_i$ and the demographic and socioeconomic variables, conditional on the water quality level. We estimate separate coefficients of $\boldsymbol{x}_i$ for each category water_i:

$$\text{IMR}_i = \begin{cases} \beta_1'\boldsymbol{x}_i + \epsilon_{i1} & \text{if } \text{water}_i = 1, \\ \beta_2'\boldsymbol{x}_i + \epsilon_{i2} & \text{if } \text{water}_i = 2, \\ \beta_3'\boldsymbol{x}_i + \epsilon_{i3} & \text{if } \text{water}_i = 3, \\ \beta_4'\boldsymbol{x}_i + \epsilon_{i4} & \text{if } \text{water}_i = 4, \\ \beta_5'\boldsymbol{x}_i + \epsilon_{i5} & \text{if } \text{water}_i = 5, \end{cases} \tag{3}$$

where for each water quality category j, ϵ_{ij} has mean 0 and variance σ_j^2 and is bivariate normal with u_i. The correlation between ϵ_{ij} and u_i is ρ_j for group j. We assume that ϵ_{ij} and u_i are independently and identically distributed across observations.

We estimate this model using a two-step estimation procedure (Greene 2002) that is a generalization of Heckman's (1979) binary-model estimator.[8] Define

$$\begin{aligned} \lambda_i \equiv E(u_i|\text{water}_i, \boldsymbol{z}_i) &= \frac{\int_{\mu_j}^{\mu_{j+1}} (\text{water}_i^* - \alpha'\boldsymbol{z}_i)\phi(\text{water}_i^* - \alpha'\boldsymbol{z}_i)d\text{water}_i^*}{\Phi(\mu_{j+1}-\alpha'\boldsymbol{z}_i)-\Phi(\mu_j-\alpha'\boldsymbol{z}_i)} \\ &= \frac{\phi(\mu_j-\alpha'\boldsymbol{z}_i)-\phi(\mu_{j+1}-\alpha'\boldsymbol{z}_i)}{\Phi(\mu_{j+1}-\alpha'\boldsymbol{z}_i)-\Phi(\mu_j-\alpha'\boldsymbol{z}_i)}, \end{aligned} \tag{4}$$

where ϕ is the standard normal density function, and Φ is the standard normal cumulative distribution function. Then the expected infant mortality rate, conditional on all the observed factors, is

$$E[\text{IMR}_i|\text{water}_i, \boldsymbol{z}_i, \boldsymbol{x}_i] = \beta_j'\boldsymbol{x}_i + E\left(\epsilon_{ij}|\text{water}_i = j, \boldsymbol{z}_i\right) = \beta_j'\boldsymbol{x}_i + \rho_j\sigma_j\lambda_i\,. \tag{5}$$

[8] There are two popular approaches in estimating the probit selection model: the full information maximum likelihood (FIML) approach and the two-step approach. In a binary selection case, Puhani (2000) found that FIML is usually more efficient than the two-step estimator. However, in an ordered probit selection model, such as we use, Chiburis and Lokshin (2007) found that the two-step estimator is more robust and is the better choice for almost all practical applications.

Thus, if we only regress IMR_i on $\boldsymbol{x}_i$ over the subsample $\{i: \text{water}_i = j\}$, without taking into account λ, the estimation will be inconsistent if $\rho_j \neq 0$.

For the ordered-probit selection model to be identified, $\boldsymbol{z}$ must contain at least one variable that is not in $\boldsymbol{x}$. That is, we must have at least one instrument z for the selection variable water (observed water quality) that is a significant determinant of water quality yet satisfies the exclusion restriction $\text{Cov}(\boldsymbol{z}, \epsilon_j) = 0$ for all j. We use wastewater dumping, rainfall, and their squares as instruments for water quality level.

In the first step, we estimate (2) by an ordered probit of water on $\boldsymbol{z}$, yielding the consistent estimates $\hat{\alpha}$ and $\hat{\mu}_j$. Define $\widehat{\text{water}}_i^* = \hat{\alpha}'\boldsymbol{z}_i$. Using (4), we consistently estimate λ_i by

$$\hat{\lambda}_i = \frac{\phi\left(\hat{\mu}_j - \widehat{\text{water}}_i^*\right) - \phi\left(\hat{\mu}_{j+1} - \widehat{\text{water}}_i^*\right)}{\Phi\left(\hat{\mu}_{j+1} - \widehat{\text{water}}_i^*\right) - \Phi\left(\hat{\mu}_j - \widehat{\text{water}}_i^*\right)}, \tag{6}$$

for $j = \text{water}_i$.

By using the observations i for which $j = z_i$, an ordinary least squares (OLS) regression of IMR on $\boldsymbol{x}$ and $\hat{\lambda}$ provides a consistent estimate of β_i'.

Moreover, σ_j can be estimated by

$$\begin{aligned}\hat{\sigma}_j &\equiv \tfrac{1}{n_j}\left(\text{RSS}_j - \hat{C}_j^2 \sum_{i:j=j} \frac{\partial \hat{\lambda}_i}{\partial \widehat{\text{water}}_i^*}\right) \\ &= \frac{\text{RSS}_j}{n_j} - \frac{\hat{C}_j^2}{n_j}\left\{\frac{\left(\hat{\mu}_j - \widehat{\text{water}}_i^*\right)\phi\left(\hat{\mu}_j - \widehat{\text{water}}_i^*\right) - \left(\hat{\mu}_{j+1} - \widehat{\text{water}}_i^*\right)\phi\left(\hat{\mu}_{j+1} - \widehat{\text{water}}_i^*\right)}{\Phi\left(\hat{\mu}_{j+1} - \widehat{\text{water}}_i^*\right) - \Phi\left(\hat{\mu}_j - \widehat{\text{water}}_i^*\right)}\right\},\end{aligned} \tag{7}$$

where n_j is the number of observations in which equation j is observed, $\hat{C}_j$ is the coefficient on $\hat{\lambda}$, and RSS_j is the residual sum of squares for the regression. Because $\hat{C}_j$ is a consistent estimator of $\rho_j\sigma_j$, we have a consistent estimator for ρ_j:

$$\hat{\rho}_j \equiv \frac{\hat{C}_j}{\hat{\sigma}_j}. \tag{8}$$

VI. Estimations

We report the regression results in table 4. In the first step, we regress water quality on the explanatory variables $\boldsymbol{z}_i$, which include our four instrument variables (precipitation, wastewater dumping, and their squares) and all the demographic and socioeconomic $\boldsymbol{x}_i$. The estimated coefficients of the four in-

TABLE 4
INFANT MORTALITY RATE ORDERED-PROBIT SELECTION MODEL

	First Step (Water Quality)	Second Step (Infant Mortality Rate by Water Type)				
		I or II	III	IV	V	VI
Precipitation (100 mm)	−.271***	. . .	. . .	. . .	. . .	. . .
	(.05)	. . .	. . .	. . .	. . .	. . .
Precipitation^2	.006***	. . .	. . .	. . .	. . .	. . .
	(.00)	. . .	. . .	. . .	. . .	. . .
Dumping	.033**	. . .	. . .	. . .	. . .	. . .
	(.02)	. . .	. . .	. . .	. . .	. . .
Dumping^2	−.001***	. . .	. . .	. . .	. . .	. . .
	(.00)	. . .	. . .	. . .	. . .	. . .
Nonagricultural population (%)	.006**	−.185***	−.007	−.164*	−.243***	−.023
	(.00)	(.07)	(.11)	(.09)	(.07)	(.06)
Illiteracy (%)	.022**	.925***	1.330***	1.053***	1.296***	1.000***
	(.01)	(.20)	(.35)	(.28)	(.37)	(.18)
Rooms	.236**	−3.813*	1.005	.251	−5.964**	−.734
	(.11)	(2.31)	(3.12)	(3.93)	(2.60)	(2.06)
Housing area	−.008	−.401***	−.597**	−.580*	−.586**	−.33
	(.01)	(.15)	(.28)	(.31)	(.26)	(.31)
GDP per capita (¥1,000)	.021*	−.323	−.644*	−.26	−.169	−.463*
	(.01)	(.21)	(.37)	(.33)	(.14)	(.28)
Government expenditure per capita	.004	−2.223	2.134	−1.585	−.934	−.872
	(.08)	(2.45)	(5.50)	(2.64)	(.97)	(2.07)
Hospital beds (per 10,000)	−.002	.04	−.025	.204	.224**	.008
	(.00)	(.05)	(.10)	(.14)	(.11)	(.07)
λ_i	. . .	−.994	−.106	9.552***	3.309	10.20***
		(3.04)	(2.94)	(3.40)	(2.06)	(2.78)
Correlation between ϵ_{ij} and u_i		$\rho_0 = -.09$	$\rho_1 = -.01$	$\rho_2 = .66$	$\rho_4 = .37$	$\rho_5 = .79$

Note. Relationship between infant mortality and surface water quality estimated by an ordered-probit sample-selection process using a two-step consistent estimator. Standard errors in parentheses. $N = 460$.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

strument variables are all statistically significant. Both precipitation and wastewater dumping have strong effects on surface water quality.[9]

In the second step, we estimate the relationship between IMR_i and the independent variables $\boldsymbol{x}_i$, taking into account the first-step estimates $\hat{\lambda}$. The estimated coefficients of $\boldsymbol{x}_i$ vary across different groups. The percentage of the nonagricultural population is statistically significant in type I or II, type IV, and type V areas. It is negatively correlated with the infant mortality rate, as expected. On average, a 10% increase in the nonagricultural population is associated with

[9] The relationship between water quality and the instrumental variables is similar if we treat water quality as a continuous variable and estimate the relationship using OLS. See table A4 for the regression results.

roughly a 1.9 per thousand drop in the infant mortality rate in type I or II areas. The estimated coefficients for type IV and V areas are 1.6 and 2.4 per thousand, respectively.

A higher illiteracy rate statistically significantly increases the infant mortality rate in all regressions. As the illiteracy rate increases by 1%, the infant mortality rate falls by roughly 1 per thousand.

Housing and living conditions are typically negatively correlated with the infant mortality rate. Both the average number of rooms per household and the per capita housing area are statistically significantly associated with the infant mortality rate. For example, if each household in type I or II areas has one more room at home, the infant mortality rate will fall by 3.8. If per capita housing area in type I or II areas increases by 10 square meters, the infant mortality rate will decrease by 4.0 per thousand. As per capita GDP goes up, the infant mortality rate goes down. An increase of ¥1,000 (about US$150) in per capita GDP is associated with a 0.17 to 0.64 per thousand fall in the infant mortality rate, depending on which group is chosen.

Given that the sample selection terms (λ_i) are statistically significant in the type IV and type VI equations, if we did not explicitly take sample selection into account by using OLS, our estimates of the effects of water pollution on the infant mortality rate would be biased. We can predict the expected infant mortality rate $\widehat{\text{IMR}}_i = \hat{\beta}_j' \boldsymbol{x}_i$ for each category and calculate the averages and differences of the predicted infant mortality rate for each category. Using the estimates in table 4, we find the average predicted infant mortality rate $\bar{y}$ for the five types of water are 20.3, 24.0, 14.9, 12.1, and 5.8 per thousand, respectively. That is, the highest infant mortality rate is associated with type III water. In the cleanest areas (type I or II) and most polluted areas (types IV, V, and VI), the infant mortality rate is lower. The relationship between water quality and the infant mortality rate is nonmonotonic, and the most polluted areas (type VI) have the lowest infant mortality rate. On average, the infant mortality rate in the type I or II areas is 3.7 per thousand lower than that in the type III areas, and the infant mortality rates in the types IV, V, and VI areas are, respectively, 9.1, 11.9, and 18.2 per thousand lower than that in the type III areas.

Our goal is to estimate the effects of water quality on the infant mortality rate. An alternative, perhaps better, method is to calculate the counterfactual $\widetilde{\text{IMR}}_j$ for equation j, if all observations were to switch to category j. Specifically, we predict

$$\widetilde{\text{IMR}}_j = \hat{\beta}_j' \boldsymbol{x}_i + \hat{\rho}_j \hat{\sigma}_j \hat{\lambda}_i, \tag{9}$$

where $\hat{\lambda}_i$ is calculated as in (6), using the actual water$_i$.

The counterfactual $\widetilde{\mathrm{IMR}}_j$ tells us what the infant mortality rate would be if all observations were to switch to different levels of water pollution. We report the five predicted counterfactual average infant mortality rates in table 5. If we switch all observations to the type I or II area, the average infant mortality rate would be 18.7 per thousand. The corresponding mortality rates are 24.8 per thousand if we switch all observations to the type III area, 16.9 for the type IV area, 15.1 for the type V area, and 4.6 for the type VI area.

We still use the type III area as the reference group and compare its the infant mortality rate with that in other groups. The overall differences in table 5 show the estimated effects of water quality on the infant mortality rate in the selection model. If water quality changes from type I or II to type III, the infant mortality rate will increase by about 6.0 per thousand. Changing water quality from type III to type IV decreases the infant mortality rate by 7.9 per thousand. The most polluted areas (type VI) have the lowest infant mortality rate rates. If water quality deteriorates from type III to type VI, the infant mortality rate drops by 20.2 per thousand.

We also estimated the ordered probit selection model by gender. We calculated the counterfactual average infant mortality rate for each category and the differences between them, using type III areas as the reference group. The results for male infants are reported in table 5. Changing water quality from type I or II to type III would increase the male the infant mortality rate by roughly 9.7 per thousand. As water quality deteriorates from type III to types IV, V, and VI, the male infant mortality rate will drop by 7.1, 8.1, and 15.1 per thousand, respectively. The estimates for female infants are also reported in table 5. The results

TABLE 5

EFFECTS OF WATER QUALITY ON THE INFANT MORTALITY RATE

	Type I or II	Type III	Type IV	Type V	Type VI
Overall:					
Counterfactual averages	18.73	24.76	16.89	15.09	4.57
Differences	−6.03		−7.86	−9.67	−20.19
OLS estimates	−3.96		−5.68	−7.68	−7.96
Male:					
Counterfactual averages	12.79	22.46	15.36	14.40	7.38
Infant:					
Differences	−9.67		−7.10	−8.06	−15.08
OLS estimates	−4.55		−5.30	−6.70	−6.02
Female:					
Counterfactual average	26.43	27.55	18.67	15.96	1.28
Infant:					
Differences	−1.12		−8.87	−11.59	−26.26
OLS estimates	−3.10		−6.19	−8.90	−10.28

Note. Differences in infant mortality for each water quality type using estimates from both the ordered probit selection model and ordinary least squares (OLS). Type III water is the reference group.

are slightly different. Changing water quality from type I or II to type III would not significantly increase the female infant mortality rate: The difference is only about 1.1 per thousand. However, as water quality becomes more polluted, changing from type III to types IV, V, and VI, the female infant mortality rate would drop dramatically, with the respective estimated magnitudes of −8.8, −11.6, and −26.3 per thousand. The estimate of the effect of water pollution on the female infant mortality rate is greater than that for the male infant mortality rate. Female infants benefit more than males from water quality falling from fair to polluted.

We compared our sample-selection model results to an OLS regression model:

$$\mathrm{IMR}_i = \beta_0 + \beta_1 D1_i + \beta_2 D2_i + \beta_3 D3_i + \beta_4 D4_i + \beta_5' \boldsymbol{x}_i + \varepsilon_i, \quad (10)$$

where IMR_i is the infant mortality rate in county i, $\boldsymbol{x}_i$ is a vector of covariates, $D1$ is a dummy equal to 1 if a county has type I or II water, $D2 = 1$ if it has type IV water, $D3 = 1$ if it has type V water, and $D4 = 1$ if it has type VI water. Type III water is the reference group.

The regression results from OLS are reported in table 6. The relationship between water quality and the infant mortality rate is nonmonotonic in the OLS regressions. On the basis of the OLS estimates, changing water quality from type I or II to type III is associated with 4.0 per thousand drop in the infant mortality rate, and as water quality deteriorates from type III to types IV, V, and VI, the associated infant mortality rate would drop by 5.7, 7.7, and 8.0 per thousand, respectively.

The less flexible OLS model underestimates the effects of water pollution on the infant mortality rate, especially for the more polluted areas. For example, as water quality deteriorates from type III to type VI, the OLS estimates indicate the infant mortality rate would decrease by only 8 per thousand, whereas the ordered probit selection model predicts it would drop by 20 per thousand.

Assuming that people cannot perceive water quality changes between type I or II and type III, the associated 6.0 per thousand increase in the infant mortality rate can be interpreted as the pure health effect of water pollution. In other words, we find that a one-degree deterioration of water quality leads to about a 30% increase in the infant mortality rate. Is this result plausible? To answer this question, we conduct a crude calculation on the magnitude of water quality changes from type I or II to type III.

According to the Environmental Surface Water Quality Standard in China, a reduction in water quality from type I to type III approximately corresponds to (using type I as a reference) a 33% decrease in the concentration of dissolved

TABLE 6
OLS REGRESSION RESULTS ON THE INFANT MORTALITY RATE

	Overall	Male	Female
Type I or II	−3.964**	−4.552***	−3.102
	(1.74)	(1.56)	(2.26)
Type IV	−5.682***	−5.304***	−6.187**
	(2.19)	(1.99)	(2.73)
Type V	−7.683***	−6.699***	−8.895***
	(1.98)	(1.66)	(2.69)
Type VI	−7.956***	−6.018***	−10.28***
	(1.88)	(1.74)	(2.27)
Nonagricultural population (%)	−.128***	−.081***	−.187***
	(.03)	(.02)	(.04)
Illiteracy rate (%)	1.054***	1.026***	1.066***
	(.13)	(.11)	(.17)
Average rooms per household	−2.678**	−1.883**	−3.586***
	(1.02)	(.91)	(1.39)
Housing area per capita	−.356***	−.166**	−.587***
	(.08)	(.07)	(.12)
GDP per capita (¥1,000)	−.356***	−.352***	−.367***
	(.09)	(.08)	(.11)
Government expenditure per capita (¥1,000)	−.339	−.226	−.439
	(.56)	(.51)	(.71)
Hospital beds per 10,000 people	.007	.033	−.023
	(.03)	(.02)	(.04)
F-statistic	36.81	31.76	35.28
R^2	.49	.48	.43

Note. Relationship between infant mortality and surface water quality estimated by a set of ordinary least square (OLS) regressions. Huber-White robust standard errors in parentheses. $N = 460$.
** Significant at 5% level.
*** Significant at 1% level.

oxygen, a 200% increase in the concentration of potassium permanganate (KMnO4), a 33% increase in the concentration of chemical oxygen demand, a 560% increase in the concentration of ammonia nitrogen (NH3-N), a 900% increase in the concentration of total phosphorus, and a 400% increase in the concentration of total nitrogen. The changes in the concentrations for other pollutants are even more pronounced. For example, the maximum allowed number of fecal coliform for type I water is 200 per liter, it is 2,000 per liter for type II water, and this maximum number increases to 10,000 per liter for type III water. Thus, the actual change in water quality from type I to type II to type III, even though it cannot be visually detected, is huge.

The large effect of water pollution on the Chinese infant mortality rate is consistent with the findings of studies covering other countries. Cutler and Miller (2005) argued that the adoption of clean water technologies such as filtration and chlorination was responsible for up to a 75% reduction in the infant mortality rate in early twentieth-century America. Galiani et al. (2005) found that privatization of the water supply in Argentina reduced the mortality of children under age 5 by 26%. Brainerd and Menon (2012) found that a 10% increase in the average of fertilizer chemicals in water in the month of conception increased the infant mortality rate by 4% and neonatal mortality by 7%.

Our study has two data limitations. We lack data on nonwater pollution by county and on alternative sources of drinking water. However, we believe that our key qualitative result of a nonmonotonic relationship between surface water and the infant mortality would hold even if such data were available.

First, if other types of pollution are correlated with water pollution, our estimated effects for surface water may reflect the combined effects of exposure to a range of environmental pollutants and risks. However, other types of pollution, such as air pollution and toxin accumulation, usually have a monotonic effect on health. While people can protect themselves from being harmed by polluted water at relatively low cost (such as by boiling the water or using ground water), avoiding the harms of polluted air and many other types of pollution can be costly and less effective. Given that water pollution is positively correlated with other types of pollution, the nonmonotonic relationship we find between polluted water and infant mortality is unlikely to be a result from other types of pollution.

Second, due to a lack of detailed data on alternative water sources such as ground, tap, and bottled water at the county level, we are unable to explore the exact channels through which people can mitigate the harms of polluted surface water. If detailed information about alternative water sources was available, we could estimate how people substitute between surface water and the various more expensive alternative sources of water. However, lacking such data, our study shows that people substitute away from extremely dangerous surface water—we just do not know to which alternative. If they did not substitute, the death rates would be very high.

VII. Conclusion

We show that the relationship between surface water quality and the infant mortality rate is nonmonotonic in China. As surface water quality deteriorates, the infant mortality rate first increases and then decreases. The infant mortality rate is the highest in areas where surface water quality is fair. This finding is robust to a variety of specifications and models.

Our explanation is that, as surface water deteriorates from a good level, people do not detect a quality change and continue to consume the water, so more infants die. As the water pollution increase more, the low quality becomes obvious, so people reduce their usage of polluted water, and more infants survive.

We find strong evidence supporting this argument. Both the OLS and the ordered-probit selection models show that the infant mortality rate is highest when water quality is fair. In regions with cleaner or more polluted water, the infant mortality rate is lower. The infant mortality rate is the lowest in regions with the most polluted surface water (type V and type VI), suggesting that avoidance behavior significantly mitigates the health risks from water pollution.

An important policy implication of this study is that the Chinese government should intervene in regions where the surface water is moderately polluted. The government could provide health information or, even better, provide safe tap water to those regions in which public awareness of water pollution is low but health risks are high.

References

Basu, A. M., and R. Stephenson. 2005. "Low Levels of Maternal Education and the Proximate Determinants of Childhood Mortality: A Little Learning Is Not a Dangerous Thing." *Social Science and Medicine* 60, no. 9:2011–23.

Behrman, J. R., J. Hoddinott, J. A. Maluccio, A. Quisumbing, R. Martorell, and A. D. Stein. 2003. "The Impact of Experimental Nutritional Interventions on Education into Adulthood in Rural Guatemala: Preliminary Longitudinal Analysis." Presented at the Meeting of the Social Policy Monitoring Network, "Health and Nutrition," November 6–7.

Bhalotra, S. 2010. "Fatal Fluctuations? Cyclicality in Infant Mortality in India." *Journal of Development Economics* 93, no. 1:7–19.

Boyle, M. H. 2006. "The Influence of Economic Development Level, Household Wealth and Maternal Education on Child Health in the Developing World." *Social Science and Medicine* 63, no. 8:2242–54.

Brainerd, E., and N. Menon. 2012. "Seasonal Effects of Water Quality on Infant and Child Health in India." IZA Discussion Paper no. 6559, Institute for the Study of Labor, Bonn.

Brennan, M. E., and R. Lancashire. 1978. "Association of Childhood Mortality with Housing Status and Unemployment." *Journal of Epidemiology and Community Health* 32, no. 1:28–33.

Cantor, K. P. 1997. "Drinking Water and Cancer." *Cancer Causes and Control* 8, no. 1: 292–308.

Chen, K., W. Yu, X. Ma, et al. 2005. "The Association between Drinking Water Source and Colorectal Cancer Incidence in Jiashan County of China: A Prospective Cohort Study." *European Journal of Public Health* 15, no. 6:652–56.

Chiburis, R., and M. Lokshin. 2007. "Maximum Likelihood and Two-Step Estimation of an Ordered-Probit Selection Model." *Stata* 7, no. 2:167–82.

Cowell, A. J. O. 2006. "The Relationship between Education and Health Behavior: Some Empirical Evidence." *Health Economics* 15, no. 2:125–46.

Cutler, D., and A. Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." NBER Working Paper no. 12352, National Bureau of Economic Research, Cambridge, MA.

Cutler, D., and G. Miller. 2005. "The Role of Public Health Improvements in Health Advances: The 20th Century United States." *Demography* 42, no. 1:1–22.

Currie, J., J. Graff Zivin, K. Meckel, M. Neidell, and W. Schlenker. 2013. "Something in the Water: Contaminated Drinking Water and Infant Health." *Canadian Journal of Economics/Revue Canadienne D'économique* 46, no. 3:791–810.

Davis, M. L., and S. J. Masten. 2004. *Principles of Environmental Engineering and Science*. New York: McGraw-Hill.

Delgado, J., M. E. Ramirez-Cardich, R. H. Gilman, R. Lavarello, N. Dahodwala, A. Bazán, V. Rodríguez, R. I. Cama, M. Tovar, and A. Lescano. 2002. "Risk Factors for Burns in Children: Crowding, Poverty, and Poor Maternal Education." *Injury Prevention* 8, no. 1:38–41.

Duflo, E., and C. Udry. 2004. "Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts, and Consumption Choices." NBER Working Paper no. 10498, National Bureau of Economic Research, Cambridge, MA.

Ebenstein, A. Y. 2012. "The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China." *Review of Economics and Statistics* 94, no. 1:186–201.

Freedman, D. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21:291–313.

Frost, M. B., R. Forste, and D. W. Haas. 2005. "Maternal Education and Child Nutritional Status in Bolivia: Finding the Links." *Social Science and Medicine* 60, no. 2: 395–407.

Galiani, S., P. Gertler, and E. Schargrodsky. 2005. "Water for Life: The Impact of Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113, no. 1:83–120.

Greene, W. H. 2002. *LIMDEP Version 8.0: Econometric Modeling Guide*. Plainview, NY: Econometric Software.

Greenstone, M., and R. Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review* 104, no. 10:3038–72.

Heckman, J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153–62.

Jalan, J., and J. Ravallion. 2003. "Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching." *Journal of Business and Economic Statistics* 21, no. 1:19–30.

Lavy, V., J. Strauss, D. Thomas, and P. de Vreyer. 1996. "Quality of Health Care, Survival and Health Outcomes in Ghana." *Journal of Health Economics* 15:33–57.

Lee, L., M. R. Rosenzweig, and M. M. Pitt. 1997. "The Effects of Improved Nutrition, Sanitation, and Water Quality on Child Health in High-Mortality Populations." *Journal of Econometrics* 77:209–35.

Levine, D., and D. Yang. 2006. "A Note on the Impact of Local Rainfall on Rice Output in Indonesian Districts." Photocopy, University of California, Berkeley, and University of Michigan.

Mansfield, C., F. R. Johnson, and G. Van Houtven. 2006. "The Missing Piece: Valuing Averting Behavior for Children's Ozone Exposures." *Resource and Energy Economics* 28, no. 3:215–28.

Martin, A. E. 1967. "Environment, Housing and Health." *Urban Studies* 4, no. 1:1–21.

Merrick, T. W. 1985. "The Effects of Piped Water on Early Childhood Mortality in Urban Brazil, 1970 to 1976." *Demography* 22, no. 1:1–24.

Miguel, E. 2005. "Health, Education, and Economic Development." In *Health and Economic Growth*, ed. G. Lopez-Casasnovas, B. Rivera, and L. Currais. Cambridge, MA: MIT Press.

Ministry of Water Resources of the People's Republic of China. 2010. "China Water Resource Bulletin." China Water and Power.

Moretti, E., and M. Neidell. 2011. "Pollution, Health, and Avoidance Behavior Evidence from the Ports of Los Angeles." *Journal of Human Resources* 46, no. 1:154–75.

National Bureau of Statistics of China. 2000. "Provincial Statistical Yearbooks." China Statistics.

Neidell, M. 2009. "Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations." *Journal of Human Resources* 44, no. 2:450–78.

Puhani, P. A. 2000. "The Heckman Correction for Sample Selection and Its Critique." *Journal of Economic Surveys* 14:53–68.

Rocha, R., and R. Soares. 2011. "Early Life Health and Educational Performance: Evidence from Rainfall Fluctuations in the Brazilian Northeast." FGV Conferences, 33rd Meeting of the Brazilian Econometric Society.

Shimshack, J., M. Ward, and T. Beatty. 2007. "Mercury Advisories: Information, Education, and Fish Consumption." *Journal of Environmental Economics and Management* 53, no. 2:158–79.

Van de Poel, E., O. O'Donnell, and E. Van Doorslaer. 2009. "Urbanization and the Spread of Diseases of Affluence in China." *Economics and Human Biology* 7, no. 2: 200–216.

Victora, C. G., et al. 1988. "Water Supply, Sanitation and Housing in Relation to the Risk of Infant Mortality From Diarrhea." *International Journal of Epidemiology* 17, no. 3:651–54.

World Bank. 2006. "Water Quality Management Policy and Institutional Considerations." Discussion paper, World Bank, Washington, DC.

Zhong, C., X. Zhiguo, Z. Wenqian, W. Derui, D. Chunguang, L. Yongjian, and X. Mei. 2005. "Eutrophication Investigation and Assessment of the Daning River after Water Storage of the Three Gorges Reservoir." *Chinese Journal of Geochemistry* 24, no. 2:149–54.

Zivin, J. G., M. Neidell, and W. Schlenker. 2011. "Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption." *American Economic Review: Papers and Proceedings* 101, no. 3:448–53.