The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China

Avraham Ebenstein[*]

November 2009

Abstract

China's rapid industrialization has led to a severe deterioration in water quality in the country's lakes and rivers. By exploiting variation in pollution across China's river basins, I estimate that a deterioration of water quality by a single grade (on a six-grade scale) increases the digestive cancer death rate by 9.7 percent. The analysis rules out other potential explanations such as smoking rates, dietary patterns, and air pollution. Estimates using 2SLS with rainfall and distance from the river's headwaters as instruments also indicate a strong relationship. I estimate that doubling China's levy rates for firm dumping of untreated wastewater would save roughly 17,000 lives per year but require an additional 500 million dollars in annual spending on wastewater treatment, implying a cost of roughly 30,000 dollars per averted death.

---

[*] Department of Economics, Hebrew University of Jerusalem. I would like to thank Joan Feng, Wallace Forman, Alison Flamm, Charlene Neo, and Dan Pam for excellent research assistance, and Scott Walker for invaluable help using the Hydro packages in ArcGIS. Special thanks to Jostein Nygard, Tamer Rabie, and Nicholas Bowden of the World Bank for helpful comments and generous access to environmental data. I would also like to thank Rodney Andrews, Jing Cao, David Card, Richard Crump, Janet Currie, Simon Galed, Alexander Gelber, Gopi Shah Goda, Jonathan Gruber, Ann Harrison, Larry Katz, Anna Levine, Ronald Lee, David Levine, Sanny Liao, Kevin Stange, and Ebonya Washington for helpful suggestions. Email: ebenstein@mscc.huji.ac.il.

1 Introduction

During the 1980s and 1990s, China's rapid economic growth transformed the country and lifted millions of its citizens out of poverty. The economic boom, however, has been accompanied by environmental degradation, including a severe deterioration in the water quality of the country's rivers and lakes. Extensive use of fertilizers by farmers and industrial wastewater dumping by manufacturing firms have rendered the water in many lakes and rivers unfit for human consumption. China's water monitoring system indicates that roughly 70 percent of the river water is unsafe for human consumption, although many farmers in rural areas still rely on these sources for drinking water (World Bank 2006).

Concurrent with the decline in water quality in China's lakes and rivers, the country has experienced an increase in rural cancer rates during the 1990s (see Figure 1). Stomach cancer and liver cancer now represent China's 4th and 6th leading causes of death and, in combination with other digestive tract cancers (e.g. esophageal), account for 11 percent of all fatalities and nearly one million deaths annually (World Health Organization 2002). Several media outlets have reported incidents of contaminated river water from industrial activity leading to outbreaks of cancer in rural villages in China (New York Times 2007, British Broadcasting Corporation 2007), but systematic analysis of these trends is lacking.

Researchers have found connections between water quality and acute water-borne diseases such as typhoid (Cutler and Miller 2005) and diarrhea (Jalan and Ravalion 2003), and access to cleaner water may lower infant mortality (Galiani et al. 2005). The connection between water quality and cancer, however, has not been fully explored. A limited but growing literature has linked water pollution to particular cancer types such as liver cancer (Lin et al. 2000, Davis and Masten 2004) or gastric cancer (Morales-Suarez-Varela et al. 1995). However, as described by Cantor (1997), the literature is incomplete regarding the causal link between water contaminants and cancer: "The epidemiologic data are not yet sufficient to draw a conclusion."

China represents an excellent context to investigate a causal association between contaminated water and digestive cancers. First, in most developing countries reliable data on pollution and mortality are unavailable. However, China's efforts in the late 1980s to begin carefully monitoring both mortality and water pollution provides reliable data on these patterns in areas where millions of inhabitants still rely on well water and lake water as their primary drinking sources. Second, since water quality is not randomly assigned to individuals, researchers must also pay attention to why a particular set of inhabitants live in an area of polluted water and the time-frame that survey respondents were exposed. In China, however, for most of the exposure window mobility was extremely limited by government regulations. Therefore, the location of residents at the time of observation in the data will likely reflect their true lifetime surface water pollution exposure. Third, China's high rates of cancer, high rates of pollution, and dramatic regional variation in water quality – driven in part by plausibly exogenous rainfall patterns – allow for more precise measurement of the causal effect of contaminated water on digestive cancer incidence.[1]

In this paper, I exploit rich data on water quality, air quality and cause-specific death rates to estimate the causal association between exposure to polluted water and cancer rates. Using a sample of 145 Disease Surveillance Points (DSP) in China and water quality measures from China's nationwide monitoring system, I examine the relationship between water quality and cancer incidence. At each DSP point I observe cause-specific death rates and the average water grade among monitoring stations in the same river basin.[2] Using Geographic Information System (GIS) software, I am able to examine several other environmental features of the river basins, such as the average air quality observed from satellite imagery and the long-term average of monthly precipitation.[3]

By comparing DSP points with better and worse water quality, I estimate using OLS that a deterioration of water quality by a single grade (on a six-grade scale[4]) increases the incidence of

---

[1] Northern China has a shorter rainy season than southern China, and as a consequence exhibits higher levels of pollutants in its surface water. This is discussed further in the next section.

[2] The river basins are identified by the United States Geological Survey project which uses satellite imagery to divide China into basins, or watersheds, which can be presumed to have similar water quality levels near the DSP point. This is described in greater detail in the data section.

[3] Air quality is proxied by average optical depth observed from NASA satellite imagery for 2002-2007. Precipitation is measured for 1961-1990 by the Global Precipitation Climatology Center (2008).

[4] The water grade is measured on a 6 point scale established by the Chinese government based on purpose of use: drinkable water (grade I or grade II), undrinkable but suitable for human contact (grade III), appropriate for general industrial water supply and recreational waters in which there is not direct human contact with the water (grade IV),

digestive cancers by 9.7 percent in my preferred specification, which includes control variables for air quality and other potential confounding factors also associated with industrialization, such as whether the site is urban, the share employed in farming, and region. Furthermore, the largest effect of water quality on digestive cancer rates is observed in areas where households are less likely to have access to tap water (which can be treated), consistent with the paper's hypothesis that the rural cancer epidemic is in part due to a lack of safe drinking water. By exploiting plausibly exogenous variation in rainfall and the DSP point's distance from the river's headwaters, I estimate 2SLS models which provide further support for a causal link between digestive cancer and surface water quality. I also rule out other factors that might confound the effect of water quality on cancer, such as smoking or diet, by demonstrating that there is no strong relationship in China between regional variation in smoking rates or dietary patterns and water quality.

In light of the potentially large health consequences of China's water pollution, I present an analysis of the benefits and costs of wastewater treatment in China. Industrial firms in China are subject to a system of levies for wastewater that fails to meet discharge standards, and I exploit regional variation in the policy's effective levy rate (yuan collected per ton discharged) to estimate the potential impact of revisions to China's current rates. Using provincial data from China's environmental yearbooks (1992-2002), I estimate that industrial cleanup (in tons) rises by 0.82 percent and spending on wastewater treatment (in yuan) rises by 0.14 percent with respect to a 1 percent increase in the effective levy rate. I also find that water quality is responsive to the total industrial dumping: a doubling of wastewater that fails to meet discharge standards increases the water grade within the basin by .22 units. In combination, these estimates imply that a doubling of China's levy rates would avert roughly 17,000 deaths per year, but require firms to spend roughly \$500 million[5] more per year on treatment, yielding a cost per averted death of roughly \$30,000. In addition, since these estimates do not include the potential benefits of cleaner water in reducing

---

appropriate only for agricultural water supply and general landscape requirements (grade V), and essentially useless (grade VI).

[5] I estimate that China's firms would need to increase spending on wastewater treatment by 14% from the level reported in 2001 of roughly \$3.7 billion, or an extra \$500 million in compliance costs.

the incidence of other causes of disease and death, they potentially understate the full benefits of tighter environmental regulations. Policymakers should recognize that cleanup efforts could yield large improvements in public health in a relatively cost-effective manner.

The next section provides background information on China's waterways and regional variation in industrial dumping and water quality. Section 3 describes the data in more detail and summarizes the patterns observed in the data in water quality, industrial dumping, and cause-specific mortality. Section 4 reports the empirical results of the analysis. Section 5 concludes.

2 Background

The pollution levels in China's water bodies have little historical precedent and are believed to be the highest in human history. Despite recent efforts to reduce water dumping by manufacturing firms, roughly 70 percent of China's surface water has been found unfit for human use, and 115 million rural inhabitants still rely on surface water as their main source of drinking water (World Bank 2006). In this section, I provide background information on environmental factors that affect water quality, geographic variation in these factors, and the variation in water quality that the analysis exploits to estimate its effect on digestive cancer rates. In this paper, I focus on digestive cancers due to their high relevance in China, where two-thirds of cancer cases are digestive, and because they are less likely to be affected by concurrent changes in smoking patterns that are occurring in China (Yang 1997). I also attempt to improve on existing cancer studies by examining data in a nationally representative sample, whereas the existing epidemiological literature is often focused on case studies of small sub-populations.

Water pollution is classified as either point source or non-point source pollution. Point source pollution is wastewater from domestic sewage and industrial wastes that is discharged from a single point. Non-point source pollution, such as urban and agricultural runoff, enters rivers and lakes at multiple points. China's experience following industrialization has led to the increase in both: farmers have attempted to increase yields through widespread fertilizer use (non-point source), and manufacturing firms have dumped inorganic compounds into water as part of their production processes. When these chemicals drain into waterways, they stimulate a river's algal growth beyond its natural speed – a process known as eutrophication – which, in turn, simulates the formation of carcinogenic compounds and chemicals. Through eutrophication, the water becomes populated by cyanobacteria (blue-green algae) leading to the formation of microcystins (Davis and Masten 2004) which have been linked directly to liver cancer (Codd 2000). Additionally, water pollution introduces nitrate into bodies of water, which, when digested, can undergo endogenous reduction to nitrite upon contact with bacteria in the gastrointestinal tract, forming highly carcinogenic N-nitroso compounds (Barrett et al. 1998, Gulis et al. 2002). Further, chlorine in water used to treat bacteria also poses a risk because it reacts with organic particles in river water to form halogenated hydrocarbons such as trihalomethane, a carcinogen.

In this paper, I build on a growing body of empirical evidence that cancer may be linked to polluted drinking water. The presence of nitrate in drinking water, a chemical observed in China's water monitoring stations, is linked to cancer increases in several studies (Armijo and Coulson 1975, Morales-Suarez-Varela 1994, Gulis et al. 2002). Tao et al. (1999) find that halogenated hydrocarbons, a by-product of chlorine, cause higher rates of esophageal cancer in men in Shanghai, China. Laboratory tests have also demonstrated that hexavalent chromium (Cr^{+6}) in drinking water for rats and mice leads to carcinogenic activity in the oral and intestinal cavities (Beaumont et al. 2008). In a recent study, Beaumont et al. (2008) found high levels of Cr^{+6} dumped by a ferrochromium factory in Liaoning province. In the exposed population, they find that stomach and lung cancer rates are significantly higher than in an unaffected areas.[6] As I will discuss further in Section 4, I also find large effects of polluted water on both digestive and lung cancer, consistent

---

[6] While the epidemiological literature focuses on a much more restricted sub-population, my estimates are broadly consistent with their findings. Gulis et al. (2002) estimate overall cancer incidence to increase by a factor of 1.14 and stomach cancer by 1.24 in high- versus low-nitrate areas of Slovakia. Beaumont et al. (2008) estimated a 1.82 relative risk ratio for stomach cancer in areas where the drinking water was contaminated by hexavalent chromium. My preferred OLS estimate for digestive cancers implies a relative risk ratio of 1.097 per water grade, and polluted water is roughly 2.5 grades worse than drinkable water, implying a relative risk ratio of 1.24. My preferred estimate is therefore somewhat higher than Gulis et al. (2002) and lower than Beaumont et al. (2008), but this back-of-the-envelope calculation suggests that my estimates are plausible relative to the existing estimates.

with these findings that a range of cancers are affected by polluted water.[7]

The deterioration of China's rivers and lakes over the past decades has been regionally bound, with water quality in northern regions declining more severely due to lower levels of precipitation. The rainy season may last as long as six to seven months in some southern areas and be as short as two or three months in more arid northern regions (World Bank 2006). As such, northern river systems have a lower capacity to absorb contaminants. In a thorough review of monitoring data for 1991-2005, the World Bank (2006) reported that 40 to 60 percent of the northern region's water is continuously in the non-functional water classification categories (grade V and VI), and therefore unfit even for agricultural use. The Hai river basin, located in northern China, is the most polluted basin in the country with 57 percent of monitored sections failing to meet Grade V, and therefore far below drinkable standards. The Yangtze river basin, however, has exhibited a far smaller deterioration in water quality, in spite of industrialization. Regional differences in water quality induced by rainfall patterns provide for observation of areas of China with similar levels of industrialization, but different levels of pollution.

In China, the degradation of waterways has also led areas without industrial activity to experience a decline in water quality. This occurred recently in Anhui, which has very low industrial activity of its own but is downstream of a major industrial zone located in the Huai river basin. According to Elizabeth Economy in her book *The River Runs Black* (2004), "Heavy rain flooded the [Huai] river's tributaries, flushing more than 38 billion gallons of highly polluted water into the Huai. Downstream, in Anhui Province, the river water was thick with garbage, yellow foam, and dead fish." In this way, regions without industrial firms can suffer from the same, or more serious, water pollution as those directly engaged in wastewater discharge, and in these rural areas the inhabitants have experienced the environmental costs of industrialization without realizing the economic benefits.[8]

The water quality available to villagers is also partly determined by whether a village is lo-

---

[7] Another chemical that has been tied to both digestive track cancers as well as lung cancer is arsenic. Arsenicosis has been linked to cancer of the lung (Yu et al. 2007) and bladder (Smith et al. 1998).

[8] Lipscomb and Mobarak (2007) deal with a set of related political economy issues and find that pollution is higher near county boundary points, where neighboring counties will incur a larger share of the pollution's cost.

cated along a tributary versus a main river segment. Tributaries in China often have more pollution than main river system segments (World Bank 2006). Smaller streams have more eutrophication, partly due to slower stream flow which leads to greater algal bloom. If water flows slowly, pollutants are not transported away quickly in slow water, and the added time of exposure leads to greater algae growth and consequently worse water quality (Zhong et al. 2005). The World Bank (2006) also reports that cleanup efforts have generally focused on main stream segments, leaving many tributaries with very poor water quality. In 2001, while the main river segment of the Yangtze had no monitoring points graded at V or VI, these high levels of pollution were observed at 48 percent of the water in the tributaries. In the next section, I describe how I will attempt to exploit both regional variation in water quality and the placement of the DSP point along a river system to estimate the causal link between water quality and cancer incidence.

China's environmental conditions have continued to worsen in spite of long-running regulatory efforts to punish firms for dumping untreated wastewater. In 1982, China established a nationwide system of fine levies assessed on the tonnage of untreated wastewater emitted by factories. By 1998, Chinese regulators had collected about 40 billion RMB ($4.9 billion) in levies, with both private and state-owned enterprises being subject to the policy (Wang and Wheeler 2005). Though China's environmental regulatory agencies have gained increasing clout in administrative decisions nationally, incentive conflicts with local administrators who rely primarily on local industries for tax revenue have limited the effectiveness of the program (Ma and Ortolano 2000). However, when enforced, the levies have been found to induce reductions in chemical dumping by firms and higher spending on wastewater treatment facilities (Wang and Wheeler 1996, Wang 2002).[9] In my empirical analysis, using more recent data, I find that the levy system continues to be an effective policy measure at inducing firms to modify their behavior and limit the discharge of untreated wastewater.

---

[9] Wang and Wheeler (1996), in an analysis on provincial data from 1987-1989 and 1992-1993, estimate an elasticity of roughly minus 1 for the dischard of chemical oxygen demand (COD) pollution intensity (discharge/output) with respect to the effective levy rate. Wang (2002), using plant level level data, estimates an elasticity of .65 for firm spending on operating expenses and .27 for firm investment in waste-water treatment facilities.

3 Data

The analysis of mortality patterns in China is based on China's Disease Surveillance Point system (DSP). The DSP is a set of 145 sites chosen to form a nationally representative sample of China's population, and selects sites across different levels of wealth and urbanization. The coverage population was also chosen to reproduce geographic dispersion in China's population, relative to patterns in China's 1990 census. The DSP records all deaths among the coverage population of 10 million residents at the points and, due to careful sample selection of the DSP sites, yields an annual sample of deaths that mirrors patterns in the country nationwide (Yang 2005). This paper relies on the data taken from roughly 500,000 deaths recorded at DSP sites between 1991 and 2000, and population counts by age and sex that are used to convert the recorded deaths into death rates.[10] A summary of cause-specific death rates during the sample period is shown in Table 1.

China's severe problems with water pollution began in the 1980s, following economic reforms in the late 1970s that led to an industrial boom. The national water monitoring system was established during the late 1980s and collects annual readings of chemical content at a set of sites across China. The World Bank produced a comprehensive assessment of water quality patterns in China from 1991-2005 using data collected by the monitoring system. The analysis presented here relies on the 2004 readings, which report water quality readings for 484 geographic points across China's nine river systems.[11] The DSP and water quality data are geographically overlaid by using data on China's river basins created by the Hydro1k project, conducted by the United States Geological Survey (see Figure 2). The project provides a suite of geo-referenced data sets that are created using a Digital Elevation Model (DEM) in which China can be separated into a set

---

[10] While the DSP data provides information on cancer mortality rather than incidence, the difference is slim in the case of digestive cancers. Physical signs of gastric cancer are rarely presented in time for curative treatment, and the 5-year survival rate for patients presented with stage III cancer (likely the vast majority of cases in China) is only 10-15% (Cabebe and Mehta 2008).

[11] Data on water quality for China in GIS format is only available in 2004. Ideally, a longer series of water quality would be used to examine the relationship between water quality and digestive cancer. Cancer incidence is thought to peak 10 to 20 years following exposure to a carcinogen, as is observed between rises in smoking rates and lung cancer rates. As a robustness check, I examine the relationship between industrial output in each river basin between 1970 and 1990 and the digestive cancer rate at the DSP point. The results (shown in Appendix Table 6) reflect that industrial output during the 20 years before the DSP sampling frame is correlated with digestive cancer rates, and the results are statistically significant and similar to those observed using actual observed water quality in 2004.

of 989 basins, and a smaller set of larger basins. Satellite imagery is exploited to assess regional variations in air quality that might also affect cancer rates.

Using NASA estimates of optical depth from aerosol imagery, I proxy for the impact of air quality on digestive cancer rates. The measure is taken between zero and 1, with higher numbers representing higher optical depth and implying the presence of more particulates and worse air quality (see Figure 3). I assign to each river basin a measure of the average particulates over the basin's region between 2002 and 2007 to reduce annual fluctuations in the data.[12] In order to examine how precipitation may affect water quality, I include measures of monthly rainfall collected by the Global Precipitation Climatology Center for 1961-1990. These measures are calculated by river basin in a manner similar to how I calculate average air quality, where I use GIS software and average the rainfall measure across the area in the same basin as the DSP point (see Figure 4). Summary statistics are shown for the water quality measures assigned to each DSP point and other characteristics of the decedents at the points in Table 2. The characteristics are reported split between DSP points along polluted rivers versus cleaner rivers. The data indicate that water quality measures are statistically significantly different along these rivers but other observable features of the decedents (e.g. average education) are similar, suggesting no obvious correlation between water quality and observables. While it is difficult to rule out the possibility that water quality is correlated with unobservable features of areas that affects cancer rates, the similarity along observable dimensions is suggestive that the omitted variables bias associated with calculations using water quality may be small (Altonji, Elder, and Taber 2005).

The river basin data from the Hydro1k project are coded using a consistent numerical scheme that allows for inference regarding water flows within the network of basins (see Appendix Figure 1). The Pfafstetter coding system, designed in 1989 by Otto Pfafstetter, assigns watershed identification numbers based on the topology of the land surface. Since it is hierarchical, it is possible to identify the flow dynamics of rivers simply from the numerical scheme of the

---

[12] The NASA data on optical aerosol levels are only available beginning in 2002. However, China's industrialization exhibits a high degree of spatial concentration that suggests that the air quality during the available window is a reasonable proxy for air quality at the DSP points following China's large boom in manufacturing (Ebenstein and Hanink 2008).

basins (Appendix Figure 2). For each DSP point, I observe the nearest river and the flowlength to the river's headwaters. The flowlength is informative regarding whether the DSP point is along a tributary or main stem (lower values are found along tributaries), and since rivers are more polluted along tributaries than main river stems, this distance is predictive of the water quality at the DSP point. The river basin coding allows me to assign both current and historical industrial output in each basin.[13]

China's Environmental Yearbooks are produced by the State Environmental Protection Agency (SEPA) and provide the necessary data to examine the responsiveness of dumping to regulatory incentives. China's environmental regulations require manufacturing firms to register all emissions, and each Yearbook contains province-level totals for the tonnage of discharge of wastewater that fails to meet standards and the total levies collected as a result of these infractions, in a consistent format for 1992 to 2002. The data also contain information on the tonnage of dumping and treatment by chemical, allowing for more detailed analysis of the statistical relationships between firm behavior and water pollution graded by chemical. Lastly, the Yearbooks contain reported spending by firms in wastewater treatment in each year, both in terms of equipment investments and operating expenses. During the 1990s, many provinces began to ratchet up enforcement of water discharge standards, leading to an increase in the fine levy collections as well as a decline in industrial dumping of untreated wastewater relative to output (see Figure 5). Using variation across provinces in the timing of these increases, I am able to assess how firm spending on cleanup responds to the environmental regulations, which reflects the marginal cost to firms of compliance with respect to levy rate increases. The industrial dumping data is also used to assess the responsiveness of water quality to dumping within a river basin. Dumping within a basin is imputed using provincial dumping and the industrial mix of counties within a province.

---

[13] While water quality data is not available during the 1970s or 1980s, I am able to assign a comprehensive series of gross industrial output in each river basin for this period using provincial data. This is presumably correlated with industrial water pollution and provides variation in the total length of exposure to polluted water among inhabitants of any particular river basin. In the appendix materials, I demonstrate a correlation between historical industrial output (1970-1990) and cancer rates in the DSP (1991-2000).

4 Empirical Results

4.1 Main Results

In Table 3, I report the baseline results of the paper, where I examine OLS models of water quality and digestive cancer rates, measured in logs.

$$\ln(DeathRate_i) = \beta_0 + \beta_1 WaterQuality_i + \Gamma X_i + \varepsilon_i \tag{1}$$

I estimate models where the log of the death rate from digestive cancer at site i is a linear function of the water quality (i.e. grade) and demographic features of the residents X_i.[14] Note that water quality is graded on a 6 point scale, where I (1) is the best water and VI (6) indicates that the water is unfit even for agricultural use. In the first regression, I examine the partial correlation of digestive cancer with the overall water quality grade, and find that an increase in the water grade by 1 level (e.g. IV to V) increases the digestive cancer rate by 12.2 percent. The coefficients are 32 percent, 12 percent, and 8 percent for the impact of water quality on esophageal, stomach, and liver cancer respectively, with the coefficients statistically significant at the 5% level for all but liver cancer, which is significant at the 10% level.

In a second set of specifications, I assess the impact of water quality on the same set of dependent variables, but with a rich set of controls for factors that might also affect digestive cancer rates. Controls are included for whether the DSP point is urban, the average education of decedents at the site above the age of 20, the share who were employed in farming and manufacturing, an imputed measure of ambient air quality (where a higher number reflects more particulates), region fixed-effects, and county income controls. Region fixed-effects are based on the survey's classification of the 145 DSP points into three main regions: eastern, middle, and western.[15] The

---

[14] Alternative specifications for the functional form are discussed in the next section. Also see Appendix Table 3, which reflects that the results are similar when broader categories of water quality (drinkable water, somewhat polluted, very polluted) are used.

[15] A total of 51 DSP sites were classified into the eastern region, which consists of Beijing, Fujian, Guangdong, Hainan, Hebei, Jiangsu, Liaoning, Shandong, Shanghai, Tianjin, and Zhejiang provinces. The middle region, including 41 DSP sites, consists of Anhui, Heilongjiang, Henan, Hubei, Hunan, Jiangxi, Jilin, and Shanzi provinces. The

DSP sites are also classified by income: rural counties included in the sample were stratified into 4 classifications of wealth (1=poorest, 4=richest) and these categories are included as fixed effects to attempt to control for variation in cancer due to income differences. With these controls, results are somewhat lower, with the estimates implying that water quality eroding by one grade induces a 9.7 percentage point increase in the digestive cancer rate. The estimates for the aforementioned types of digestive cancer are 27, 11, and 5 percentage points respectively. It may be unsurprising that the coefficients are not dramatically changed by including controls, since Table 2 reflects that water quality variation is not highly correlated with observable features of a location, such as rates of urbanization or air quality. Table 3 also indicates that air quality also has a positive and statistically significant relationship with digestive cancer rates, with an increase in the particulate index variable (that varies from 0-1) by 1% inducing a .23% increase in the digestive cancer rate.[16] This may reflect a causal link between contaminants in the air and the likelihood of tumors forming in digestive organs (Jerret et al. 2005) or it may reflect a correlation between air quality and other carcinogenic environmental factors, such as water dumping or exposed carcinogenic chemicals.

In Table 4, I present an additional set of OLS regressions in which I examine whether the relationship between water quality and digestive cancers is observed differently by the share of households within the county that have access to tap water. Note that tap water is much more likely to be treated and consequently cleaner than water taken directly from rivers or wells.[17] The results reported by gender reflect that water quality and digestive cancer rates are highly correlated in areas without tap water, but much less strongly linked in areas with tap water, suggesting that living near polluted rivers is less dangerous when households have access to tap water. The results in Table 4 also reflect a consistency between the estimated impact for men and women. For example, an

---

remaining 53 sites were attributed to the western region, which consists of Chongqing, Gansu, Guangxi, Guizhou, Neimenggu, Ningxia, Qinghai, Shaanxi, Sichuan, Xicang, Xinjiang, and Yunnan provinces. This classification generally accounts for region-specific patterns in cancer rates. While fixed effects by province would be preferable, they do not provide sufficient degrees of freedom. Since N=145 and there are 30 provinces, in many cases there are only 2 or 3 DSP sites per province.

[16] The air quality measure has a mean of .48 and a standard deviation of .19 in the sample of DSP sites.

[17] Other research in China has found a link between the source from which people draw water and local cancer rates. In a prospective cohort study, Chen et al. (2005) found high rates of colon and rectal cancer among those who relied on well water relative to those relying on municipal water in Jiashan County, Zhejiang Province, China.

increase in the water grade by 1 unit is associated with a 12.8 percentage point increase in the esophageal cancer rate for men in areas without tap water and a 14.1 percentage point increase for women. The impact of overall water quality on esophageal cancer is also positive and similar by gender in areas without tap water (27 percentage points for men, 24 percentage points for women) and this holds for stomach cancer as well (13 percentage points for men, 11 percentage points for women).

These findings are compelling evidence that environmental factors are responsible for the correlation between water quality and digestive cancer rates. In particular, if water quality did not directly affect digestive cancer rates but was instead reflecting an unobserved correlation between water quality and omitted factors, such as occupational exposure to carcinogens, one would expect to find larger elasticities for men, who are more likely to work in mines and other hazardous occupations. However, the similarity by gender is suggestive instead that factors shared by men and women are responsible for the correlation, such as water quality. Likewise, the lack of an effect in areas with tap water is compelling that drinking water is responsible for the correlation between surface water quality and digestive cancer rates.

In Table 5, I consider whether the OLS results could be explained by unobserved correlation between water quality and other potential risk factors for digestive cancer, such as smoking rates and dietary patterns. Using province-level information on smoking rates and dietary practices from household survey data (China Household Income Survey 1995, China Health and Nutrition Survey 1989-2006), I examine whether either smoking or diet patterns covary with water quality. The results indicate that smoking rates are similar across the water quality readings, suggesting that the estimated impact of water quality is not being confounded by smoking patterns.[18] Likewise, no large difference in diet is observed across sites with better and worse quality, suggesting that regional differences in diet are not responsible for the correlation between water quality and digestive cancer. So, although diet is a known risk factor for digestive cancers, it is uncorrelated

---

[18] National surveys reflect that smoking rates for men are in excess of 75%, but fewer than 8% of women smoke (Yang 1997). The age profile of smoking rates was very similar in both the national smoking survey of 1984 and in a follow-up survey in 1996, suggesting that smoking patterns are unlikely to be responsible for the recent increase in China's digestive cancer rate.

with water quality and is therefore unlikely to be biasing the estimated effect of water quality on cancer.

Although dietary patterns in China are known to vary by region, it is unlikely to explain the patterns in cancer mortality I observe in the data, which reflect high digestive cancer rates among northern areas with lower rainfall (and consequently worse surface water quality). First, while salty and pickled foods are thought to be associated with higher digestive cancer rates (Kono and Hirohata 1996), southern China is not very different than northern China in this dietary dimension. In fact, the principal difference between northern and southern China in terms of diet is the South's "rice culture" versus the northern "wheat culture". Carbohydrates are thought to be a risk factor for Asian men with high rates of this disease (Ji et al. 1998) but inhabitants of both regions consume large amounts of carbohydrates. Since regional differences in diet are not thought to be risk factors for digestive cancer, it is unlikely that unobserved differences in diet are confounding the regression analysis.

In Table 6, I perform a falsification exercise where I attempt to assess whether water quality's correlation with cancer is an artifact of a correlation between water quality and higher death rates in general. As shown in the table, water quality appears largely unrelated to other causes of death, but is strongly correlated with cancer rates. A deterioration of water quality by a single grade induces an 9.0 percentage point increase in the cancer rate (significant at 5%), but has a small and statistically insignificant relationship to the death rate from causes other than cancer. The results also indicate that the correlation between water quality and cancers of all types is similar to what is found between digestive cancers alone (9.7 percent). Since digestive cancers represent nearly two-thirds of all cancers, this is perhaps unsurprising, but it reflects that non-digestive cancers, such as lung cancer, are also positively correlated with water pollution and may be causally linked to water pollution as well. Table 6 reflects that the correlation between water grade and lung cancer is 11.5 percent, potentially implicating water pollution in rising lung cancer rates observed in China. Water pollution has been blamed by local residents for the outbreak of throat and lung cancer in some of China's "cancer villages" (Voss 2008) and has been linked to the incidence of certain respiratory tract cancers in China (Yu 2007).[19] While the analysis here focuses on digestive cancers, the link between water quality and cancer incidence may exist across a broader class of cancer types, lung in particular, and represents an area for further research.[20] As a consequence of the large impact on cancer rates, which represent roughly a sixth of all deaths, the 9.0 percentage point increase in cancer deaths results in a roughly 1.7 percent increase in the total death rate, while other causes are largely unaffected.

4.2 Robustness Checks

In Table 7, I present a set of 2SLS estimates of water quality's relationship with digestive cancer rates, exploiting plausible exogenous variation in water quality due to differences in precipitation across the DSP sites and variation in the distance from the DSP point to the nearest river's headwaters. Regional differences in water quality are induced by rainfall patterns, since additional rainfall dilutes the pollutants and contributes to flow speed, which in turn makes the river less prone to eutrophication.[21] The second instrument, distance from a river's headwaters, is correlated with water quality variation within a region and in fact within a river basin, since the tributaries of a river system are generally more polluted than main river segments. In the first column, I examine the first-stage relationship between monthly rainfall in milliliters and distance in kilometers (in 000s) from the river's headwaters and the observed water grade within the river basin. The coefficient implies that an increase by 100 milliliters lowers the water grade by 1.2 levels, significant at the 1% level, which suggests that large variation in surface water quality is induced by variation in rainfall patterns. The water grade is declining with respect to distance from the river's headwaters, which is consistent with an expectation that tributaries will have worse water quality than main stem river

---

[19] Voss (2008) documents high rates of cancer and poor water quality in Shenqiu County (Henan Province). Accessible online at http://www.stephenvoss.com/stories/ChinaWaterPollution/

[20] A comparison of cancer rates in China relative to the United States reveals that in spite of China's high male smoking rate, which is roughly 3 times the American, lung cancer is less common in China and represents a smaller share of total cancer deaths (see Appendix Table 4). The table suggests that the causal links between behavior, environment, and cancer incidence may operate differently in China and the United States.

[21] In the appendix materials, I present further discussion of the validity of using rainfall as an instrument for water quality. See Appendix Table 7, where I examine the reduced form relationship between rainfall and other causes of death. The results indicate that there is only a weak correlation between rainfall and non-cancer death rates.

segments. Each additional thousand kilometers of river flow reduces the water grade by .45 units. An F test of the joint significance of the two instruments is 14.61, which is highly significant as well.

In column 2, I exploit this variation and regress the log of the death rate from digestive cancer on the predicted water quality reading from the first-stage and the covariates included from Table 3 (e.g. urban, years of education, etc.). The 2SLS estimates are larger than the OLS estimates and imply that increasing the water quality grade by 1 level increases the digestive cancer rate by 27 percent. The estimates for esophageal cancer and stomach cancer imply that increasing the water quality grade by 1 level increases the incidence of these diseases by 87 percent and 41 percent respectively, and both are statistically significant at the 1% level. The 2SLS estimate for liver cancer is 2% and not statistically significant. Overall, the 2SLS results support the claim that there is a causal link between water quality and digestive cancers, though the point estimates are somewhat larger than what I find using OLS in Table 3.

The larger coefficient estimates using 2SLS may reflect that measurement error in water grade is attenuating the OLS estimates and the true effect is understated by OLS. Alternatively, water quality and digestive cancer may be related non-linearly with respect to grade. When I use broader categorical measures of water quality, I find that digestive cancer rates are 19 percent higher in areas with medium water quality (grade III) and 33 percent higher in areas with very poor water (grade IV+) relative to areas with potable surface drinking water (grade I and grade II). If the instruments are exploiting variation in water quality where the water quality-digestive gradient that is steeper than the average slope, the 2SLS estimates will be larger than OLS. While it is unclear whether the 2SLS estimates are closer to the average causal effect of water grade on digestive cancer incidence than my baseline OLS estimates, the results are consistent with a direct impact of water quality on digestive cancer rates.

4.3 Estimating the Cost of Averting a Death through Cleanup

Digestive cancers are responsible for nearly one million deaths annually (WHO 2002) and policy efforts to lower the incidence of these diseases can have large benefits in terms of population health and life expectancy. Digestive cancers represent 20 percent of deaths among those age 40 to 60 and are more common at these ages than other leading causes of death, such as stroke (see Figure 6). The conservative estimate of the impact of improving China's water grade is that almost 93,000 deaths could be averted annually, since nearly 1 million people (980,000) die each year of these diseases, and each water grade improvement is associated with 9.7 percent fewer digestive cancer deaths. As such, it is of great policy interest to know the cost of improving China's waterways by a single grade. In combination with my estimates of the potential benefit in averted cases of digestive cancer, it provides information regarding the trade-offs associated with tighter wastewater regulations in China.

In order to assess the cost of improving China's water, in Table 8, I examine three relationships: the relationship between China's surface water quality and industrial dumping, the relationship between industrial dumping and the levy rates, and the relationship between levy rates and firm spending on wastewater treatment facilities.[22] In column 1 of Table 8, I report the relationship between the overall water grade and the total dumping of untreated wastewater within a river basin, which indicates that an increase in dumping by 10% would induce a .022 unit increase in water grade, and the result is statistically significant at the 1% level.[23] In column 2 and 3, I examine how China's levy rates affected firm dumping behavior for 1992-2002, the window for which China's environmental yearbooks contain the necessary data on industrial wastewater treatment (in tons) and total spending by firms in wastewater treatment. Raising fines by 1 percent increases the tonnage of cleanup by 0.82 percent (significant at the 1% level) and spending on cleanup by 0.14 percent (significant at the 10% level). This is estimated with province and year

---

[22] Summary statistics of the industries with the largest share of industrial pollution are presented in Appendix Table 5. Firms classified as producing chemicals or chemical products were responsible for 19% of the dumping of untreated wastewater, the largest share among the 21 industrial categories.

[23] Dumping is assigned to each river basin using data on the industrial mix within the river basin, and county-level data of industrial production. Details regarding this calculation are in the appendix.

fixed effects that absorb province- or year-specific variation in levies, and the standard errors are clustered at the province level. Since China's levy rates have been generally rising, this strategy essentially exploits the timing of levy increases across China and is robust to either time-invariant or province-invariant factors driving levy rates and dumping behavior. These coefficients indicate that the marginal cost of abatement in China is much lower than the average cost, since anticipated wastewater treatment is anticipated to increase by almost 6 times as much as the total spending on cleanup, implying that during the 1990s many provinces could have induced large increases in cleanup by raising levy rates.[24]

In Table 9, I synthesize the preceding analysis to calculate the anticipated savings (in lives) of raising China's levy rate and the compliance costs required of firms in wastewater treatment spending. A full 100 percent increase in China's levy rate is predicted to reduce untreated dumping by 82 percent, which in turn improves the water grade by 22 percent (from Table 8) of 82 percent, yielding a predicted improvement in water quality of .18 units (.82*.22). In the preferred OLS specification in Table 3, each unit decrease in water grade is associated with roughly 9.7 percent fewer deaths due to digestive cancer, or roughly 95,000 deaths due to digestive cancer. Since water quality is expected to improve by .18 units, the proposed levy increase would avert roughly 17,000 deaths. In terms of the anticipated compliance costs, I estimate that China's firms would need to increase spending on wastewater treatment by 14 percent from the level reported in 2001 of 29 billion yuan, or roughly $3.7 billion on wastewater treatment, which implies an anticipated extra $500 million in compliance costs.[25] This implies a cost per death averted of roughly $30,000 ($500 million/17,000 deaths averted). Since each digestive cancer death imposes a cost of slightly more than 20 years in life expectancy (20.12), this amounts to a cost of roughly $1,500 per year.[26]

---

[24] An alternate interpretation is that the province and year fixed-effects are over-controlling for the relevant incentives. The simple correlation between the levy rate and spending on cleanup is roughly 0.43, which would imply marginal costs roughly three times larger than the preferred estimate of 0.14, but much of this variation is absorbed by the province and year fixed-effects. In terms of the cost to avert a death by increasing the levy rates, this would yield an estimate three times larger than what I present in Table 9.

[25] The environmental yearbook estimate for 2000 (in the 2001 yearbook) is the most recent year in which China's environmental yearbook reported both operating expenses and equipment value. This calculation also assumes an exchange rate of 8 yuan per dollar.

[26] This is calculated as the weighted average of remaining life expectancy, where the weights are defined by the share of digestive cancer deaths that occur at that age in the DSP. Alternatively, I have calculated that life expectancy at birth

This estimate is low relative to conventional valuation placed on a human life, even in low-income countries. According to surveys conducted in China by the World Bank in 2005, estimates based on the contingent valuation method indicate a mean value of a statistical life among the participants of 1.4 million yuan, or $175,000 (World Bank 2007).[27] While it is difficult to measure the full cost in quality and length of life of contracting digestive cancer, the simple back-of-the-envelope calculation here suggests that the cost of compliance with higher pollution levies is justified by their benefit. My estimates suggest that even if the cost per averted death was much higher than the estimated $30,000, the cost to saving a life through cleanup would still be justified by the benefit in improved health outcomes.

In addition, my estimate of the potential health benefit of raising levies may be very conservative. First, the preferred OLS estimate of 9.7 percent is smaller than point estimates without regional control variables (12 percent) or estimates from 2SLS (27 percent), which serves to understate the impact of improving water quality. Second, because I am focusing on a narrow measure of the health benefits of cleanup, the estimate presented here can be thought a lower bound of the full impact on mortality. Third, these calculations only count the cost of a death, when in fact digestive cancer is also associated with years of poor health and distress preceding death. Lastly, China's rapid income increases have led to large reductions in infant mortality and the incidence of infectious diseases. As the population ages, reducing the prevalence of digestive cancer will avert an increasing number of deaths, since the disease's share of deaths is higher among those in middle and old age (see Figure 6).

---

would be increased by 1.5 years through the elimination of this cause from a standard life table. The life expectancy at birth in the DSP sample (1991–2000) is 73.9 years, and is 75.4 years when the death rate from digestive cancer is set to zero and the death rates from other causes are assumed to equal their distribution in the DSP. Results available upon request.

[27] The World Bank (2007) reports that the survey was administered in Chongqing and Shanghai (twice) and the survey questionnaire, with minor changes, was identical to those administered in the U.S., Canada, U.K., France, Italy, and Japan. See Krupnick et al. (2006) for more information regarding the surveys in China.

5 Conclusion

Despite an increase in clean-up efforts in recent years, the overall degradation of China's waterways continues. While the capacity of wastewater treatment facilities has grown, it has not kept pace with the growth of industrial output. The pollution intensity of China's industrial firms has declined (discharge per yuan of output), but the tonnage of water dumping has continued to increase (World Bank 2007).

Although China's economy has grown rapidly and brought with it many benefits, the adverse health effects of pollution threaten to mitigate the health benefits of the country's new-found wealth. The results presented here highlight one channel by which China's industrial growth has led to deterioration in health outcomes. The dumping of untreated wastewater in densely populated areas has contributed to China's increasing cancer rate, and cancer is now the country's leading cause of death (Chinese Ministry of Health 2008). The cost of industrial pollution is also disproportionately borne by the millions of Chinese farmers who are unable to access safe drinking water and are least able to share in the benefits of China's urban manufacturing boom. Recent estimates by the World Bank (2006) indicate that as many as half of China's inhabitants still lack access to safe drinking water. In 2005, China's Ministry of Water Resources announced ambitious plans to reduce the number of residents without access to clean drinking water by a third by 2010 and to provide safe access to drinking water to all rural residents by 2030. Even if these goals are met, however, in the near future the need to curb industrial dumping of untreated wastewater is clear and pressing.

The analysis reveals a relatively low cost to averting deaths via water cleanup of roughly \$30,000, suggesting that dumping regulations need to be more aggressively enforced. The gaps in enforcement of China's regulations reveal inappropriately inexpensive opportunities to avert deaths relative to the value of life that Chinese citizens report in contingent valuation surveys. These surveys indicate average valuations of roughly \$175,000 for the value of a statistical life (Krupnick et al. 2006). Protests by villagers who are justifiably angered by the contamination of the water supply also suggest that the current Chinese policy may represent an ongoing threat to political stability in China. The government reported 50,000 environmental protests in 2005 alone (Los Angeles Times 2006), providing further motivation for tightening environmental standards on China's industrial firms. Wastewater dumping is in part responsible for China's emerging cancer epidemic, and addressing this problem through stricter levy enforcement may yield large improvements in public health and life expectancy at a reasonable cost. Failure to act could prove costly for the millions of rural Chinese farmers who continue to rely on surface water for their drinking supply.

6 Data Appendix

6.1 China Disease Surveillance Points

The analysis of mortality patterns in China is based on China's Disease Surveillance Point system (DSP). The DSP is a set of 145 sites chosen to form a nationally representative sample of China's population and selects sites across different levels of wealth and urbanization (see Appendix Table 1). The coverage population was also chosen to reproduce geographic dispersion in China's population, relative to patterns in China's 1990 census. The DSP records all deaths among the coverage population of 10 million residents at the points and, due to careful sample selection of the DSP sites, yields an annual sample of deaths that mirror patterns in the country nationwide (Yang 2005). In each DSP site, there is at least one township hospital and the 'Disease Prevention Unit' in these hospitals is responsible for vital registration. While deaths in hospitals are registered using standard protocols, in rural areas roughly 80% of deaths occur in the home. In these instances, a village health worker reports the incident to the responsible Prevention Unit and a Chinese CDC staff member visits the household and completes a death certificate based on a description of symptoms from family members. The data is then processed by the Chinese CDC in Beijing, and the data are validated using an internal procedural check system and then evaluated for consistency using statistical measures described in greater detail by Yang et al. (2005). This paper relies on the data taken from roughly 500,000 deaths recorded at DSP sites between 1991 and 2000 and population counts by age and sex that are used to convert the recorded deaths into death rates. The age- and sex-specific death rates by causes are aggregated at each site by the age and sex distribution observed in China's 2000 census, allowing for comparison of mortality rates across different DSP points that have different age distributions.

6.2 China's River Basins and Stream Network

China's river basins are identified using GIS files provided by the United States Geological Survey (USGS) Hydro1k project, which provides a digital elevation model for China and identifies river basin divides, or watersheds. The Hydro1k file is overlaid with a GIS file from China's 2000 census taken from the Harvard Geospatial Library, which provides the geographic location of each of the DSP points and identifies the river basin in which the point is located.[28] The river basin data from the Hydro1k project are coded using a consistent numerical scheme that allows for inference regarding water flows within the network of basins. The Pfafstetter coding system, designed in 1989 by Otto Pfafstetter, assigns watershed identification numbers based on the topology of the land surface. The finest breakdown of river basins is at a level 6 river basin, which divides China into 989 different principal basins. Lower level basins are composed of several principal basins, and there are 387 level 4 river basins. At the highest aggregation level, China is broken into 4 very large level 1 river basins.

---

[28] Documentation for the Hydro1k project is available online at http://edc.usgs.gov/products/elevation/. The China census coded by river basin is available for download at the author's website at http://www.demog.berkeley.edu/~ebenstei/pollution/data_appendix/Basin_Data/.

6.3 Water Quality, Air Quality, and Rainfall Patterns

Data on water quality is taken from China's national monitoring system. The analysis presented here relies on the 2004 readings, which report water quality readings for 484 geographic points across China's nine river systems (see Appendix Table 2). The water grade is measured on a 6 point scale: drinkable water (grade I or grade II), undrinkable but suitable for human contact (grade III), appropriate for general industrial water supply and recreational waters in which there is not direct human contact with the water (grade IV), appropriate only for agricultural water supply and general landscape requirements (grade V), and water that is essentially useless (grade VI). Each reading is observed for a set of monitored chemicals and the overall water grade is the grade recorded for the highest pollutant. These data are available for download from the author's website.[29]

The water grades reported at the monitoring sites are averaged across the river basins to form the critical variable in the analysis. In order to assign water grades to the DSP point, I merge the water quality point data with the river basin data. The water quality for each river basin is identified for the local river basin (level 6) and for larger basins composed of local basins. I assign to each DSP point the average water grade reading for the smallest basin in which a monitoring station is observed. Among the 145 DSP points, 74 are observed within a level 6 basin in which I also observe at least one water quality monitoring station. The remaining DSP points are assigned the average of water quality readings within the smallest river basin, with the vast majority being linked to water monitoring stations within the same level 4 basin.

Air quality readings are taken from NASA estimates of optical depth from aerosol imagery, generated from the MODIS sensor on NASA's Terra satellite.[30] The measure is taken between zero and 1, with higher numbers representing higher optical depth and implying the presence of more particulates and worse air quality (see Figure 3). The extinction or total aerosol optical depth is a measure of radiation extinction due to aerosol scattering and absorption. I assign to each level 6 river basin a measure of the average aerosol optical depth over the basin's region between 2002 and 2007 to reduce annual fluctuations in the data. The rainfall measures for each level 6 basin are taken from the Global Precipitation Climate Center for 1961-1990. The GPCC rainfall measure is gauge-based gridded monthly precipitation data sets for the global land surface based on the complete GPCC monthly rainfall station data-base which contains readings from more than 70,000 different stations. The data used in this analysis is from the GPCC normals taken at 0.5 degree resolution and are available online.[31]

6.4 Manufacturing Output and Industrial Wastewater Dumping by River Basin

Manufacturing output in each river basin is taken from a survey of manufacturers that contains information for all firms with greater than 5 million RMB in sales per year in 2003. These manufacturing data are used in conjunction with province-level wastewater discharge (or dumping)

---

[29] http://www.demog.berkeley.edu/~ebenstei/pollution/data_appendix/Water_Quality_Data/

[30] A detailed description is available online at http://disc.sci.gsfc.nasa.gov/giovanni/secondary/users-manual/G3_manual_parameter_appendix.shtml. The data are available for download at the author's website at http://www.demog.berkeley.edu/~ebenstei/pollution/data_appendix/Rainfall_Data/.

[31] See description at ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf.

data in 2003 to create basin-level measures of wastewater discharge exploiting different patterns in wastewater discharge by industry (see Appendix Table 5). Each county is assigned a share of the province's wastewater dumping by assuming that each industry's share of wastewater discharge is in proportion to the county's share of production by industry and assuming that each industry's share of wastewater discharge is the same as the proportion observed in 2000, which is the most recent year in which China's Environmental Yearbook reports wastewater dumping by industry.[32] The county's total wastewater dumping is then calculated as the county's share of provincial dumping times the province's total wastewater dumping in 2003. The county-level dumping measures are aggregated at the level 4 river basin level and are used in the basin-level regression of the impact of industrial dumping on water quality reported in Table 8 (column 1). Historical estimates for industrial production within each river basin are taken from provincial data on gross industrial output value used in combination with the 2000 China census distribution of manufacturing workers within each province. The county-level production data for 1970-1990 is then aggregated to the river basin level and assigned to the DSP points by merging the data by level 6 river basin. These are used in Appendix Table 6.

6.5 Chinese Smoking and Dietary Patterns

The China Household Income Survey (1995) asked detailed questions on tobacco spending and use, and these provide smoking rates by province and gender.[33] A province- and sex-specific smoking rate is assigned to each DSP point using the province's fraction of men and women 30 and older who report being a smoker in the survey. Data are available for only 19 provinces, and the 102 DSP points within those provinces. The China Health and Nutrition Survey (1989) is used for the assessment of the relationship between regional variation in diet and water quality. The CHNS website provides a constructed nutrient intake file which contains data for roughly 83,000 observations of individuals who report 3-day average intake of total calories and the share of the diet consumed on carbohydrates, fat, and protein.[34] These data are available for only 9 provinces, or 56 of the 145 DSP points.

7 Is Rainfall a Valid Instrument for Water Quality?

In Appendix Table 7, I examine whether rainfall can be used as an instrument for water quality. Rainfall has a large impact on surface water quality through three primary channels. First, areas with more rainfall have pollution diluted by the relatively clean water from the atmosphere. Second, areas with more rainfall have faster river currents. If water flows slowly, pollutants are not transported away quickly in slow water, and the added time of exposure leads to greater algae growth and consequently worse water quality (Zhong et al. 2005). Third, areas with insufficient rainfall may attempt to compensate by the overuse of fertilizer. This leads to excessive runoff and degrades the surface water quality further. Although rainfall has a large impact on water quality, its

---

[32] The concordance used to match the industry definitions used in the two surveys, county-level production by industry for 2003, and the province-level wastewater dumping for the period are available online at http://www.demog.berkeley.edu/~ebenstei/pollution/data_appendix/Dumping_Data/.

[33] These data are available for download at http://www.icpsr.umich.edu/.

[34] http://www.cpc.unc.edu/projects/china/data/data.html

suitability as an instrument is debatable. Rainfall affects many other factors, including income and crop types, and these may directly affect cancer rates. In order to assess whether rainfall satisfies the exclusion restriction of only affecting cancer rates through its effect on surface water quality, I present the reduced form relationship between cause-specific death rates and monthly rainfall within the river basin, while including the complete set of control variables (e.g. air quality, county income, etc.). I also stratify the 145 DSP sites by whether a majority of residents of the county have access to tap water, since areas without tap water are more likely to experience negative health consequences of surface water pollution. The results indicate that areas with more rainfall have significantly lower cancer rates, but rainfall has only a very weak relationship with mortality rates for other causes. Furthermore, rainfall's impact on the cancer rate of the DSP site is only significant in counties with low rates of tap water use. I find that an increase in monthly rainfall by 1 milliliters decreases cancer mortality by 0.60 percent among those without access to tap water, significant at the 1% level. The effect on digestive (0.80) and lung cancer (0.60) are particularly large and both estimates are significant at the 1% level. In contrast, I find almost no relationship between rainfall and cancer rates in areas with high rates of tap water use, and almost no relationship with rainfall and other causes of death. While rainfall affects many facets of life, the results suggest that the reduced form relationship between rainfall and cancer rates is consistent with an interpretation that drinking water quality in rivers and wells is affected by the amount of rainfall, and the dilution of chemicals that are in the waterways. Also note that the scientific literature is extremely limited regarding a identifying a causal link between cancer rates and rainfall patterns, suggesting there is no obvious mechanism by which rainfall would be affecting cancer rates other than through affecting drinking water quality.[35]

8 Estimating the Cost of Averting a Death through Cleanup

In Section 4.3, I examine the potential policy impact of raising China's fine levies as a mechanism for inducing improvement in China's water quality and consequently reducing mortality. These calculations assume that the only benefit of water clean-up on health is through a decline in digestive cancer rates. The death rate from digestive cancer at site i is given by $DeathRate_i$, the water quality and dumping at site i is given by $WaterQuality_i$ and $Dumping_i$, and the effective tax applied to dumping is given by $TaxRate_i$. By definition, the total deaths from digestive cancer is related to the death rate (measured as deaths per 100,000) by the following equation, where N is the total population.

$$TotalDeaths = DeathRate \cdot (\frac{N}{100,000}) \tag{2}$$

---

[35] As mentioned in the previous section, one plausible mechanism by which rainfall could affect cancer rates (other than through surface water quality) is through its impact on diet. Rainfall is associated with eating more rice and less wheat in China. Wong et al. (1998) examine data on rice and wheat intake across eight major cities in China and fail to find any relationship between gastric cancer and dietary patterns.

The anticipated change in total deaths from a change in the tax rate can be re-written in terms of elasticities as follows.

$$\begin{aligned} \frac{\partial TotalDeaths_i}{\partial TaxRate_i} &= \frac{\partial \ln TotalDeaths_i}{\partial \ln TaxRate_i} \cdot \frac{TotalDeaths_i}{TaxRate_i} \\ &= \frac{\partial \ln DeathRate_i}{\partial \ln TaxRate_i} \cdot \frac{TotalDeaths_i}{TaxRate_i} \end{aligned} \tag{3}$$

where the second line follows from the definition of $TotalDeaths$ in (2). By the chain rule, we can express the relationship between the death rate and the tax rate as the product of several partial derivatives that are observed in the data.[36]

$$\frac{\partial TotalDeaths_i}{\partial TaxRate_i} = \frac{\partial \ln DeathRate_i}{\partial WaterQuality_i} \cdot \frac{\partial WaterQuality_i}{\partial \ln Dumping_i} \cdot \frac{\partial \ln Dumping_i}{\partial \ln TaxRate_i} \cdot \frac{TotalDeaths_i}{TaxRate_i}$$

The first term can be estimated by regressing the log of the death rate from digestive cancer on the water quality (i.e. grade) and demographic features of the site X_i.

$$\ln(DeathRate_i) = \beta_0 + \beta_1 WaterQuality_i + \beta_2 X_i \tag{4}$$

The second term can be estimated by regressing the water quality on the log of dumping of untreated waste $Dumping_i$ and millimeters of monthly rainfall R_i.

$$WaterQuality_i = \gamma_0 + \gamma_1 \ln(Dumping_i) + R_i \tag{5}$$

Firms will optimize by adjusting three dimensions of behavior: they can change the amount of pollution per unit of production at existing plants, they can alter output, or they can choose to relocate to a location with less regulation. These three factors will yield a reduced form pattern in the data in which water dumping and the tax on dumping are negatively correlated. The elasticity of dumping to the tax rate is estimated as follows.

$$\ln(Dumping_i) = \lambda_0 + \lambda_1 \ln(TaxRate_i) \tag{6}$$

The increase in the tax rate also requires firms to spend more on wastewater treatment. Suppose the total cost of spending by firms at site i is given by $TotalCost_i$ and there are T firms.

$$TotalCosts = \sum_{i}^{T} TotalCost_i \tag{7}$$

---

[36] Since no data set has reliable information on the direct relationship between the death rate and the tax rate, I estimate the parameters in separate data sets with different sample sizes. The relationship between death rates and water quality is observed at the 145 DSP sites (See Table 3). The relationship between water quality and firm dumping is observed at the river basin level in 2003/2004 (See Table 8). The relationship between water dumping, cleanup spending, and the tax rate is observed by province and year for 1992-2002 (See Table 9). For expository purposes, in this appendix I refer to the data as being observed at site i.

The anticipated change in total costs from a change in the tax rate can be re-written in terms of elasticities as follows.

$$\frac{\partial TotalCost_i}{\partial TaxRate_i} = \frac{\partial \ln TotalCost_i}{\partial \ln TaxRate_i} \cdot \frac{TotalCost_i}{TaxRate_i} \tag{8}$$

The first term can be estimated by regressing the log of the total cost of cleanup on the log of the tax rate on dumping.

$$\ln(TotalCost_i) = \delta_0 + \delta_1 \ln(TaxRate_i) \tag{9}$$

The statistic of interest is the cost of saving a life through water cleanup. Using (3), we can predict the number of averted deaths due to an increase in the tax rate by applying the elasticity to the base-period level of deaths from digestive cancer, roughly 980,000 (World Health Organization 2002). Using (8), we can predict the amount of increased costs to firms in cleanup spending induced by an increase in the tax rate by applying the elasticity to the base-period level of spending on cleanup, roughly $3.7 billion (China Environmental Yearbook 2001). The estimated cost to avert a death from wastewater cleanup implied by an increase in the tax rate can be expressed in terms of the reduced-form coefficients that define the elasticities, the base-period level of cleanup costs, and the base-period level of deaths from digestive cancer.

$$\frac{\partial TotalCosts/\partial TaxRate}{\partial TotalDeaths/\partial TaxRate} = \frac{\delta_1 \cdot TotalCosts}{\beta_1 \cdot \gamma_1 \cdot \lambda_1 \cdot TotalDeaths} \tag{10}$$

References

[1] Altonji, Joseph, Todd Elder, and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113(1):151-184.

[2] Armijo, Rolando and Anne H. Coulson. 1975. "Epidemiology of Stomach Cancer in Chile—The Role of Nitrogen Fertilizers." *International Journal of Epidemiology* 4:301-209.

[3] Barrett, Jennifer H., Roger C. Parslow, Patricia A. McKinney, Graham R. Law, David Forman. 1998. "Nitrate in drinking water and the incidence of gastric, esophageal, and brain cancer in Yorkshire, England." *Cancer Causes & Control* 9(2):153-9.

[4] Beaumont, James J., Richard M. Sedman, Stephen D. Reynolds, Claire D. Sherman, Ling-Hong Li, Robert A. Howd, Martha S. Sandy, Lauren Zeise, and George V. Alexeeff. 2008. "Cancer mortality in a Chinese population exposed to hexavalent chromium in drinking water." *Epidemiology* 19(1):12-23.

[5] Cabebe, Elwyn C. and Vivek K. Mehta. 2008. "Gastric Cancer." Available for download at http://emedicine.medscape.com/article/278744-overview.

[6] Chen, Kun, Weiping Yu, Xinyan Ma, Kaiyan Yao, Qinting Jiang. 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(6):652-656.

[7] China Ministry of Health. 2008. "Cancer, stroke top killers for Chinese." *China Daily*. Available for download at: http://www.china.org.cn/health/2008-04/30/content_15038853.htm.

[8] Cantor, Kenneth P. 1997. "Drinking Water and Cancer." *Cancer Causes & Control* 8(3):292-308.

[9] Codd, Geoffrey A. 2000. "Cyanobacterial toxins, the perception of water quality, and the prioritisation of eutrophication control." *Ecological Engineering* 16:51–60.

[10] Cutler, David and Grant Miller. 2005. "The Role of Public Health Improvements in Health Advances: The 20th Century United States." *Demography* 42(1):1-22.

[11] Davis, Mackenzie L. and Susan J. Masten. 2004. *Principles of Environmental Engineering and Science*. McGraw-Hill.

[12] Ebenstein, Avraham Y. and Dean M. Hanink. 2008. "A Spatial Analysis of Selected Manufacturing and Service Sectors in China's Economy using County Employment Data for 1990 and 2000." Unpublished manuscript.

[13] Economy, Elizabeth C. 2004. *The River Runs Black: The Environmental Challenge to China's Future*. Cornell University Press.

[14] Galiani, Sebastian, Paul Gertler and Ernesto Schargrodsky. 2005. "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113(1):83-120.

[15] Griffiths, Dan. 2007. “China’s ‘cancer villages’ pay price.” *British Broadcasting Corporation News*. Available for download at: http://news.bbc.co.uk/2/hi/asia-pacific/6271103.stm.

[16] Gulis, Gabriel, Monika Czompolyova and James R. Cerhan. 2002. “An ecologic study of nitrate in municipal drinking water and cancer incidence in Trnava district, Slovakia.” *Environmental Research* 88(3):182-7.

[17] Jalan, Jyotsna and Martin Ravallion. 2003. “Does piped water reduce diarrhea for children in rural India?” *Journal of Econometrics* 112:153-173.

[18] Jerrett, Michael, Richard Burnett, Renjun Ma, C. Arden Pope III, Daniel Krewski, K. Bruce Newbold, George Thurston, Yuanli Shi, Norm Finkelstein, Eugenia E. Calle and Michael J. Thun. 2005. “Spatial Analysis of Air Pollution and Mortality in Los Angeles.” *Epidemiology* 16:727-736.

[19] Ji, Bu-Tian, Wong-Ho Chow, Gong Yang, Joseph K. Mclaughlin, Wei Zheng, Xiao-Ou Shu, Fan Jin, Ru-Nie Gao, Yu-Tang Gao and Joseph F. Fraumeni, Jr. 1998. “Dietary Habits and Stomach Cancer in China.” *International Journal of Cancer* (76):659-664.

[20] Kahn, Joseph and Jim Yardley. 2007. “As China Roars, Pollution Reaches Deadly Extremes.” *New York Times*. Available for download at:

http://www.nytimes.com/interactive/2007/08/26/world/asia/choking_on_growth.html.

[21] Kono, Suminori and Tomimo Hirohata. 1996. “Nutrition and Stomach Cancer.” *Cancer Causes & Control* 7(1):41-55.

[22] Krupnick, Alan, Sandra Hoffmann, Bjorn Larsen, Xizhe Peng, Ran Tao and Chen Yan. 2006. “The willingness to pay for mortality risk reductions in Shanghai and Chongqing, China.” *Resources for the Future*, The World Bank: Washington, D.C.

[23] Lin, Nian Feng, Jie Tang and Hoteyi S. Mohamed Ismael. 2000. “Study on environmental etiology of high incidence areas of liver cancer in China.” *World Journal of Gastroenterology* 6(4):572-576.

[24] Lipscomb, Molly and Ahmed Mushfiq Mobarak. 2007. “Decentralization and Water Pollution Spillovers: Evidence from the Re-drawing of Boundaries in Brazil”. Unpublished manuscript.

[25] Ma, Xiaoying and Leonard Ortolano. 2000. *Environmental Regulation in China: Institutions, Enforcement, and Compliance*. Rowman & Littlefield Publishers, Inc.

[26] Magnier, Mark. 2006. “As China Spews Pollution, Villagers Rise Up.” *Los Angeles Times*. Available for download at: http://articles.latimes.com/2006/sep/03/world/fg-enviro3.

[27] Morales-Suarez-Varela, Maria M., Augustin Llopis-Gonzalez, and Maria L. Tejerizo-Perez. 1995. “Impact of Nitrates in Drinking Water on Cancer Mortality in Valencia, Spain.” *European Journal of Epidemiology* 11(1):15-21.

[28] National Aeronautics Space Administration. 2008. Optical air readings from MODIS sensor on Terra satellite. Available for download at: http://disc.sci.gsfc.nasa.gov/.

[29] Pfafstetter, Otto. 1989. “Classification of hydrographic basins: coding methodology.” Unpublished manuscript.

[30] Schneider, Udo, Tobias Fuchs, Anja Meyer-Christoffer and Bruno Rudolf. 2008. “Global Precipitation Analysis Products of the GPCC.” Global Precipitation Climatology Centre (GPCC) Deutscher Wetterdienst, Offenbach, Germany.

[31] Smith, Allan H., Mario Goycolea, Reina Haque and Mary Lou Biggs. 1998. “Marked Increase in Bladder and Lung Cancer Mortality in a Region of Northern Chile Due to Arsenic in Drinking Water.” *American Journal of Epidemiology* 147(7):660-9.

[32] Tao, Xuguang, Huigang Zhu and Genevieve M. Matanoski. 1999. “Mutagenic Drinking Water and Risk of Male Esophageal Cancer: A Population-based Case-Control Study.” *American Journal of Epidemiology* 150(5):443-52.

[33] Voss, Stephen. 2008. “Water Pollution in China.” Text and online photo album available for download at: http://www.stephenvoss.com/stories/ChinaWaterPollution/.

[34] Wang, Hua. 2002. “Pollution regulation and abatement efforts: evidence from China.” *Ecological Economics* 41:85-94.

[35] Wang, Hua and David Wheeler. 1996. “Pricing Industrial Pollution in China: An Econometric Analysis of the Levy System.” Policy Research Working Paper 1644, The World Bank, Washington, D.C.

[36] Wang, Hua and David Wheeler. 2005. “Financial incentives and endogenous enforcement in China’s pollution levy system.” *Journal of Environmental Economics and Management* 49:174-96.

[37] Weyer, Peter, James Cerhan, Burton Kross, George Hallberg, Jiji Kantamneni, George Breuer, Michael Jones, Wei Zheng, Charles Lynch. 2001. “Municipal Drinking Water Nitrate Level and Cancer Risk in Older Women: The Iowa Women’s Health Study.” *Epidemiology* 12(3):327-338.

[38] Wong, BCY, CK Ching, SK Lam, ZL Li, BW Chen, YN Li, HJ Liu, JB Liu, BE Wang, SZ Yuan, CP Xu, Xh Hou, AT Zhang, and ZT Zheng. 1998. "Differential north to south gastric cancer-duodenal ulcer gradient in China." *Journal of Gastroenterology and Hepatology* 13:1050-7.

[39] World Bank. 2006. “Water Quality Management Policy and Institutional Considerations.” *Discussion Paper,* The World Bank, Washington, D.C.

[40] World Bank. 2007. “Cost of Pollution in China: Economic Estimates of Physical Damages.” *Conference Edition*, The World Bank, Washington, D.C.

[41] World Health Organization. 2002. "Mortality and Burden of Disease Estimates for WHO Member States in 2002." Available for download at: http://www.who.int/evidence/en/.

[42] Yang, Gonghuan. 1997. “1996 National Prevalence Survey on Smoking Patterns.” *China Science and Technology Press*.

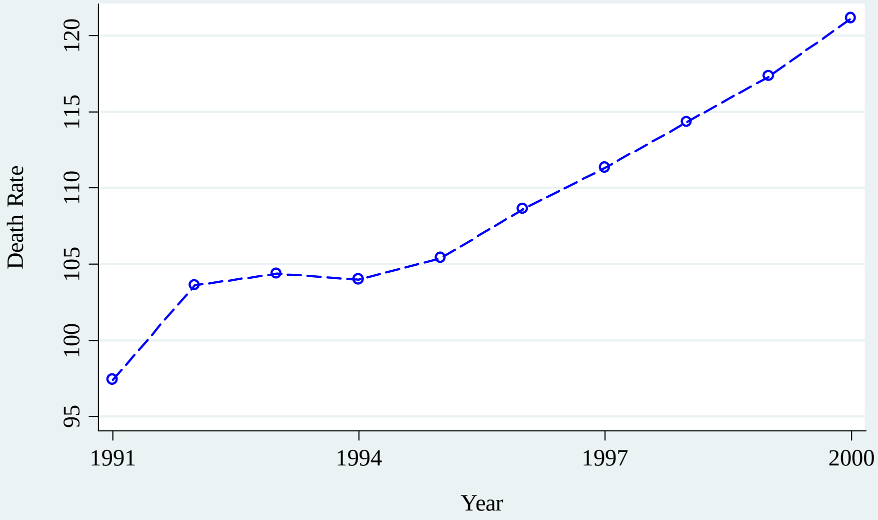
[43] Yang, Gonghuan, Jianping Hu, Ke Quin Rao, Jeimin Ma, Chalapati Rao and Alan D Lopez. 2005. “Mortality registration and surveillance in China: History, current situation and challenges.” *Population Health Metrics* 3(3).

[44] Yu, Guangqian, Dianjun Sun and Yan Zhang. 2007. “Health Effects of Exposure to Natural Arsenic in Groundwater and Coal in China: An Overview of Occurrence.” *Environment Health Perspectives* 115(4):636-642.

[45] Zhong, Chenghua, Zhiguo Xing, Wenqian Zhao, Derui Wang, Chunguang Deng, Yongjian Li, and Mei Xing. 2005. “Eutrophication investigation and assessment of the Daning River after water storage of the Three Gorges River.” *Chinese Journal of Geochemistry* 24(2):149-154.

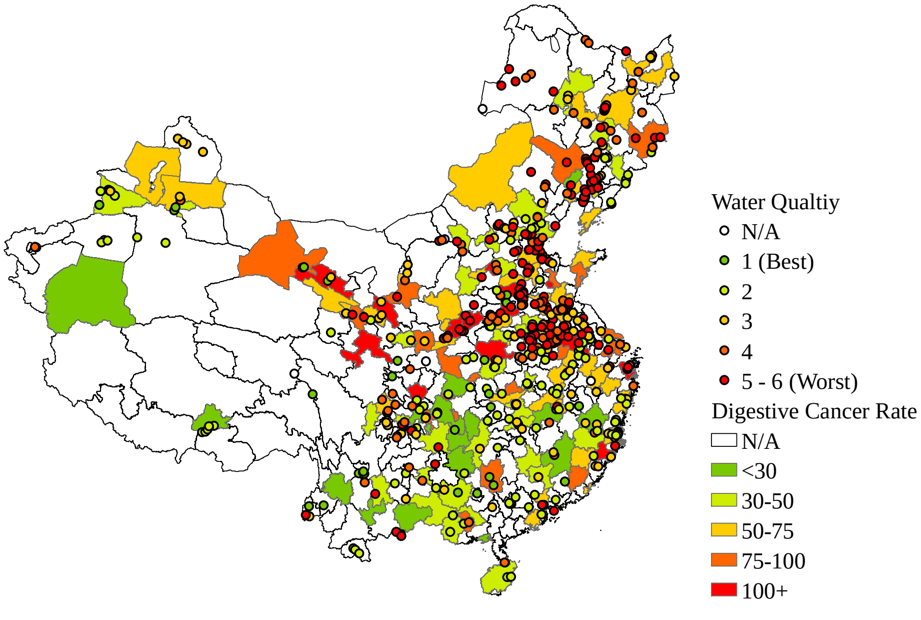
7 Figures and Tables`

Figure 1: Trends in Cancer Death Rates in Rural China



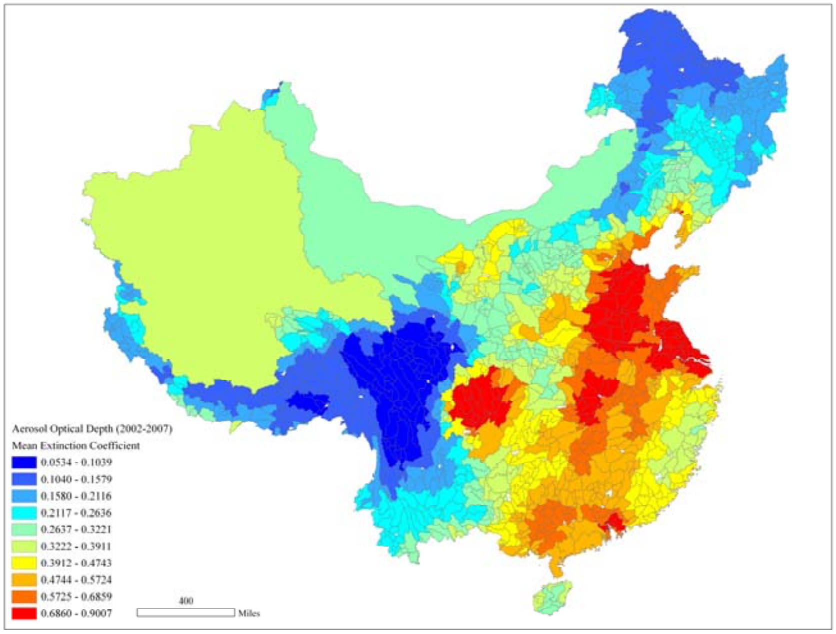
Source: China DSP. Graph displayed is created using STATA lowess smoother.
Death rates are per 100,000 persons.

Figure 2: Water Quality and Digestive Cancer Rates in Select Locations, 1991-2000



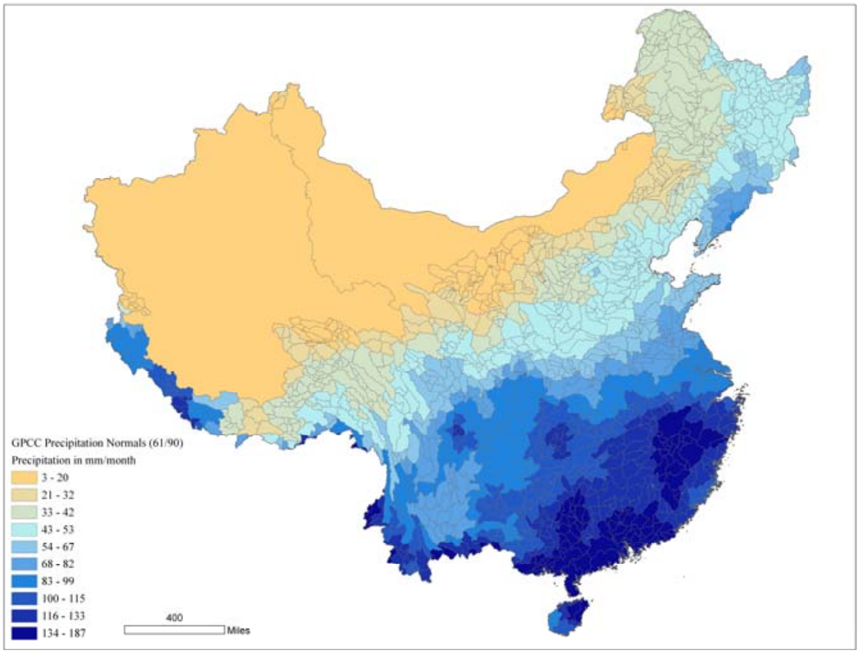
Source: China Center for Disease Control. Disease Surveillance Points (DSP) comprise a nationally representative sample of mortality for China. Rates reported per 100,000 and are age and sex adjusted using the 2000 census population structure.

Figure 3: Air Quality Patterns in China



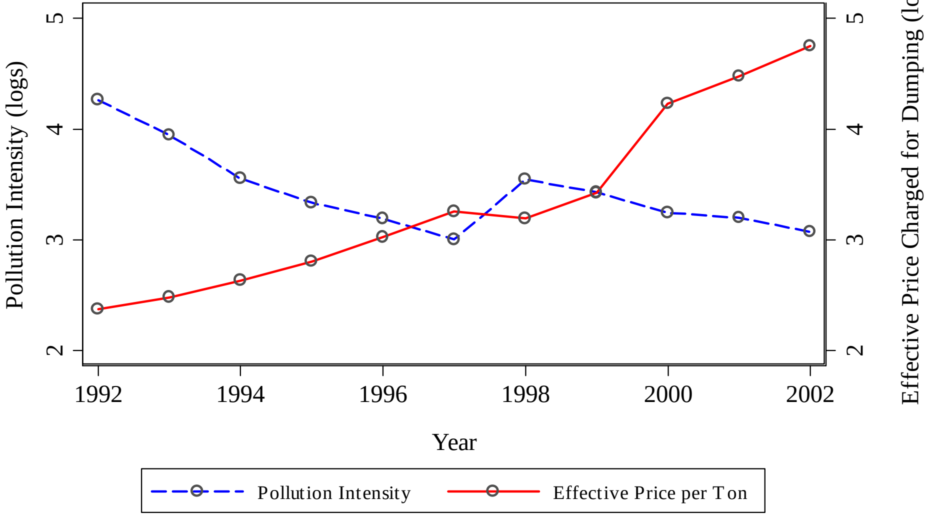
Source: NASA satellite imagery.

Figure 4: Monthly Precipitation Patterns in China, 1961-1990



Source: Global Precipitation Climatology Center.

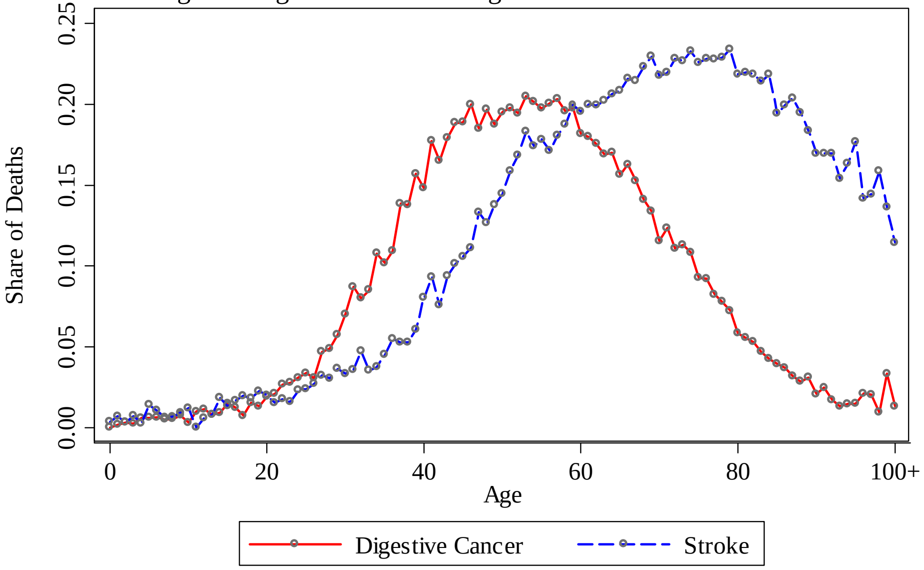
Figure 5: Pollution Intensity and the Effective Price Charged for Dumping



Soure: China Environmental Yearbook (1993-2003).

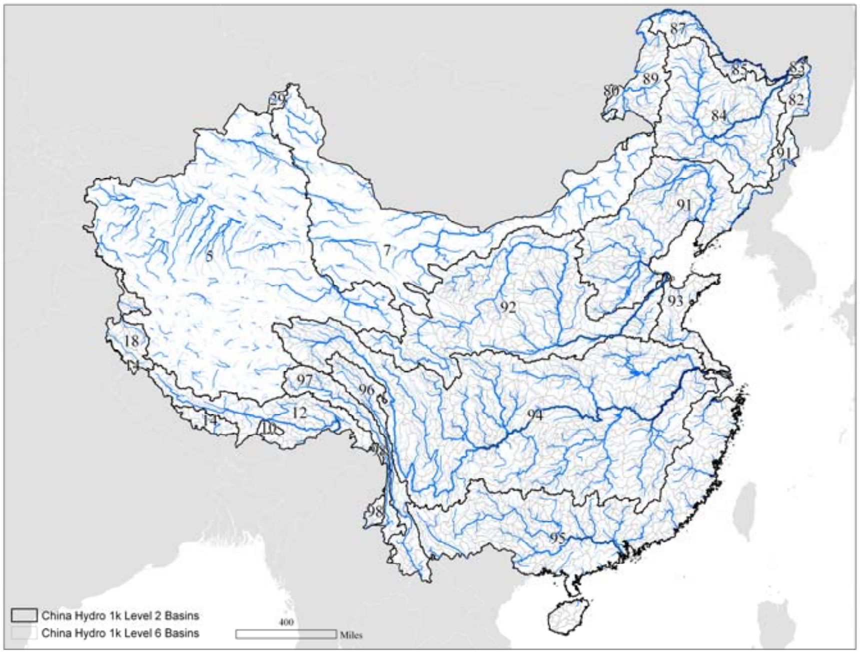
Note: Pollution intensity is the ratio of dumping in tons to 10,000 yuan of industrial output. The effective price is the value of collected yuan per hundred tons not meeting wastewater discharge standards

Figure 6: Age Distribution of Digestive Cancer and Stroke Deaths



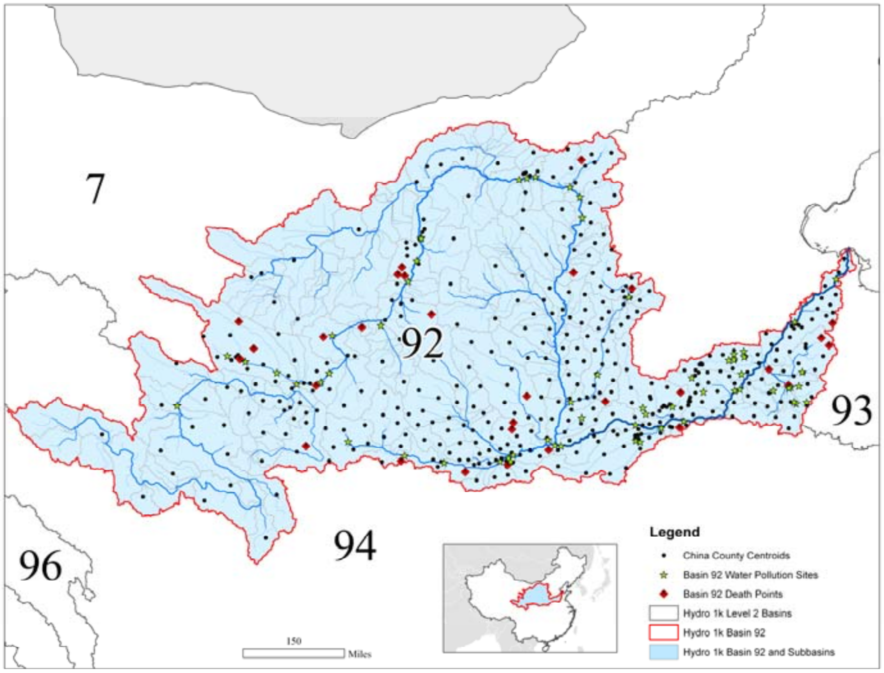
Source: China Disease Surveillance Points (1991-2000).

Appendix Figure 1: Main river basins in China



Source: Hydro 1k Project. The above figure reflects the principal 2-digit basins (or watersheds) that comprise the hydrological surface of China. The dark lines reflect the breakdown of the 2-digit basins, and the lighter outline is the breakdown of China into 989 lower-level basins.

Appendix Figure 2: The Yellow River Basin



Source: China Hydro 1k Project.

Table 1

Age-adjusted Death Rates (per 100,000) by Cause in China, 1991-2000

	Males		Females	
	Rural	Urban	Rural	Urban
	(1)	(2)	(3)	(4)
Panel 1: Death Rates by General Cause				
All Causes	726	599	600	463
Cancer	133	150	78	85
Digestive Cancers	94	79	49	38
Lung Cancers	26	50	12	21
Other Cancer	14	21	18	25
Heart	133	100	134	91
Stroke	125	125	107	100
Respiratory Illnesses	126	72	120	58
Accidents / Violence	91	49	59	31
Other	118	102	100	97
Panel 2: Death Rates for Types of Digestive Cancer				
Esophageal Cancer	21	11	11	4
Stomach Cancer	32	22	17	11
Liver Cancer	33	31	14	11
Other Digestive Cancers	9	15	7	12

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

Note : N=145. Age adjustment is performed by calculating age-specific death rates and creating weighted averages using the population structure in China's 2000 census. Other digestive cancers include colon cancer, intestinal cancer, and pancreatic cancer. The reported death rates are the average rates for the 145 sites, weighted by the population at each site. These calculations exclude roughly three thousand deaths (of the 500,000 deaths) in the sample where I am missing information on the age or sex of the decedent.

Table 2

Sample Means for Disease Surveillance Points by River System

Statistic	Points Along Polluted Rivers	Points Along Cleaner Rivers	Differences
	(1)	(2)	(3)
Digestive Cancer Rate	79.6	63.3	16.34**
Overall Water Grade	4.58	3.13	1.45***
Ammonia Nitrogen	3.88	2.52	1.36***
Biological Oxygen Demand	3.74	1.58	2.16***
Lead	1.16	1.08	0.08*
Oils	3.14	1.69	1.45***
Permanganate	4.12	2.27	1.85***
Volatile Phenol	2.33	1.29	1.04***
Average Years of Education	3.96	4.30	-0.34
Share in Farming	0.70	0.63	0.07
Urban site (1=yes)	0.24	0.23	0.01
Share in Manufacturing	0.07	0.10	-0.03
Air Pollution Reading	0.54	0.49	0.05
Monthly Rainfall (mm)	52.6	97.6	-45.1***
Share of Households with Tap Water	0.51	0.48	0.026
County Income Level (1=poor, 4=rich)	3.02	2.85	0.170
# of Sites	59	86	27

Source : Chinese Disease Surveillance Points Mortality Registration System (DSP), China National Monitoring Center (2004), Global Precipitation Climatology Center (2008)

Note : All sites along rivers which have average water grade higher than 4.0 are in column 1, which includes the Huai, Hai, Yellow, and Songhua rivers. Sites along all other rivers, including the Yangtze and Pearl, are in column 2. Higher grades reflect lower water quality (1=best, 6=worst) and a greater concentration of the listed pollutants.

The water grade measure at each DSP site reflects the average water grade among monitoring sites in the same river basin. The air pollution reading is taken from satellite imagery and takes on values from 0-1, with higher values reflecting more particulates in the air, and is reported as the average reading in the river basin containing the DSP site. The rainfall measure is the average monthly rainfall in millileters in the river basin containing the DSP site from 1961-1990. The DSP survey defines each site by the income level of the coverage population of rural counties into 4 categores (1=poor, 4=rich). The sample means are the average values (e.g. average education) among decedents at each site restricted to deaths among persons age 20 and older. Sample means are reported weighted by the population at each site.

Table 3

Ordinary Least Squares (OLS) Regressions of Log of Digestive Cancer Rates on Water Grade

	No Controls				With Controls			
Statistic	Digestive (all)	Esophageal	Stomach	Liver	Digestive (all)	Esophageal	Stomach	Liver
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Water Grade (1=best, 6=worst)	0.122**	0.315***	0.120*	0.080*	0.097**	0.274***	0.105	0.050
	(0.052)	(0.108)	(0.065)	(0.045)	(0.044)	(0.093)	(0.062)	(0.032)
Average Education					-0.016	-0.014	-0.080	0.012
					(0.044)	(0.109)	(0.066)	(0.031)
Share in Farming					0.029	0.719	0.510	-0.484***
					(0.293)	(0.991)	(0.415)	(0.175)
Urban (1=yes)					-0.260	0.109	0.060	-0.588***
					(0.194)	(0.667)	(0.254)	(0.128)
ln(Air Pollution)					0.232*	0.647**	0.070	0.253**
					(0.127)	(0.263)	(0.181)	(0.120)
County Income Controls					Yes	Yes	Yes	Yes
Region Controls	No	No	No	No	Yes	Yes	Yes	Yes
R Squared	0.083	0.121	0.048	0.045	0.267	0.227	0.213	0.315

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

Source : China Disease Surveillance Points Mortality Registration (DSP), China National Monitoring Center (2004)

Note : N=145. The first four columns represent OLS regressions of the logarithm of the death rate of a cause on the average water grade of the river basin in which the DSP site is located. I add covariates for columns (5)-(8), which are the average values (e.g. education) among decedents at each site restricted to deaths among persons age 20 and older. Standard errors are robust and clustered at the province level. The water grade measure at each DSP site reflects the average water quality among monitoring sites in the same river basin. County income controls are fixed effects for the 4 income classifications of rural DSP sites. Region controls are fixed effects for whether the site is located in an eastern, middle, or western province (see the text for details). Regressions are weighted by the population at each DSP site.

Table 4

OLS Regressions of Log of Digestive Cancer Rates on Water Grade by Sex and Access to Tap Water

	Men		Women	
	Low Tap Water Share	High Tap Water Share	Low Tap Water Share	High Tap Water Share
	(1)	(2)	(3)	(4)
Digestive (all)	0.128**	0.022	0.141**	0.054
	(0.056)	(0.054)	(0.062)	(0.065)
Esophageal	0.265*	0.216**	0.242	0.206*
	(0.136)	(0.085)	(0.165)	(0.117)
Stomach	0.131*	0.034	0.105	0.046
	(0.077)	(0.072)	(0.083)	(0.076)
Liver	0.087**	-0.014	0.140***	0.042
	(0.040)	(0.045)	(0.045)	(0.052)
Observations	74	71	74	71

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

Source : China Disease Surveillance Points Mortality Registration (DSP), China National Monitoring Center (2004), China 2000 Census

Note : Each reported coefficient represents a separate regression, with water grade as the independent variable and the log of the death rate of the listed cause as the dependent variable. Columns 1 and 3 represent regressions for the 74 DSP points located in counties where a majority of households do not have access to tap water, and columns 2 and 4 represent the 71 DSP points where a majority have tap water according to the 2000 census. All regressions include the controls shown in Table 3.

Table 5

Smoking and Dietary Habits by Water Grade in China

	Smoking Rates		Dietary Patterns				
Water Grade	Men	Women	Caloric Intake	Carbo-hydrates	% Fat	% Protein	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Level I (Best)	0.732	0.034	2,172	15.21	2.89	2.86	79.04
Level II	0.685	0.065	2,376	15.12	2.96	2.84	79.08
Level III	0.697	0.025	2,303	14.75	3.06	2.92	79.28
Level IV	0.704	0.035	2,238	15.38	2.75	2.97	78.90
Level V	0.704	0.059	2,311	16.13	2.41	2.99	78.47
Level VI (Worst)	0.710	0.046	2,316	15.19	2.82	2.92	79.06

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

Note: The smoking rates are shown for the DSP sites which were in the 19 provinces included in the CHIS (1995), which includes 102 of the 145 DSP sites. Information on diet is shown for DSP sites located in the 9 provinces included in the CHNS, which includes 56 of the 145 sites.

Table 6

OLS Regressions of Log of Death Rates by Cause in China on Water Grade

	All	Men	Women
	(1)	(2)	(3)
All Causes	0.019	0.016	0.023
	(0.02)	(0.02)	(0.02)
Cancer (all types)	0.090***	0.087**	0.099***
	(0.03)	(0.03)	(0.03)
Digestive	0.097**	0.091**	0.110**
	(0.04)	(0.04)	(0.05)
Lung	0.114***	0.086***	0.167***
	(0.03)	(0.03)	(0.03)
Other	0.034	0.030	0.044
	(0.02)	(0.03)	(0.03)
Non-Cancer (all types)	0.009	0.005	0.014
	(0.03)	(0.02)	(0.03)

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

Source : Chinese Disease Surveillance Points Mortality Registration System, China National Monitoring Center (2004)

Note : N=145. Each reported coefficient represents a separate regression, with water grade as the independent variable and the log of the death rate of the listed cause as the dependent variable. All regressions include the controls shown in Table 3.

Table 7

Two-stage Least Squares (2SLS) Regressions of Log of Digestive Cancer Rates on Water Grade using Annual Rainfall and Distance from Headwaters

	First-Stage	Two-stage Least Squares			
Statistic	Water Grade	Digestive (all)	Esophageal	Stomach	Liver
	(1)	(2)	(3)	(4)	(5)
Monthly Rainfall (mm)	-0.016***				
	(0.01)				
Distance from Headwaters (km x 10^3)	-0.450***				
	(0.10)				
Water Grade (1=best, 6=worst)		0.265**	0.873***	0.411***	0.021
		(0.105)	(0.252)	(0.147)	(0.101)
Average Education	0.127	-0.035	-0.046	-0.112	0.009
	(0.110)	(0.050)	(0.108)	(0.080)	(0.034)
Share in Farming	-1.067	-0.335	0.581	0.069	-0.839***
	(0.99)	(0.30)	(0.82)	(0.57)	(0.17)
Urban (1=yes)	-1.147	-0.264	0.377	0.155	-0.749***
	(0.785)	(0.183)	(0.524)	(0.355)	(0.156)
ln(Air Pollution)	0.889**	0.292*	0.465	0.092	0.415***
	(0.323)	(0.150)	(0.301)	(0.206)	(0.147)
F Test of Instruments	14.61***				

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

Source : China Disease Suveillance Points Mortality Registration (DSP), China National Monitoring Center (2004), Global Precipitation Climatology Center (2008)

Note : N=145. The first column is the first-stage relationship between water grade at the DSP site, the covariates (e.g. urban), and two instrumental variables: the average monthly rainfall in millileters in the basin and the distance from the river's headwaters in thousands of kilometers. This is calculated as the longest path from the DSP point's nearest stream node and the drainage basin divide. The regressions in columns (2) through (5) represent 2SLS regressions where the dependent variable is the logarithm of the death rate of a cause on the predicted average water grade from colum (1) and the other covariates. Standard errors are robust and clustered at the province level. Regressions are weighted by the population at each DSP site.

Table 8

OLS Regressions examining Water Quality, Industrial Dumping, and Firm Response to Pollution Levies

	Water Grade	Log of Total Industrial Cleanup (tons)	Log of Spending on Cleanup (100 million yuan)
	(1)	(2)	(3)
Log of Total Untreated Waste Water (tons)	0.220*** (0.08)		
Log of Effective Fine Levy		0.815*** (0.21)	0.137* (0.08)
Period Available	2003-2004	1992-2002	1992-2002
R Squared	0.241	0.706	0.902
N	125	319	319

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

Source : China National Monitoring Center (2004), China Environmental Yearbooks (1993-2003), China Manufacturing Census (2003)

Note : In the first column, the dependent variable is the average water grade of monitoring stations and the independent variable is the log of the number of tons of industrial wastewater not meeting discharge standards within the level 4 river basin. The number of observations corresponds to the total number of level 4 river basins in which water quality readings are available. Controls (not shown) include monthly rainfall and the total area of the river basin. In the second and third columns, the dependent variable is the log of the total tons of reported cleanup and the log of spending on cleanup respectively, and the independent variable is the effective fine levy. The effective fine levy is yuan collected per ton of wastewater discharge failing to meet regulatory standards. Cleanup and spending on cleanup is reported for all registered manufacturing firms by province and year. The regressions in columns 2 and 3 include province- and year- fixed effects, and the standard errors are robust and clustered at the province level. Regressions are weighted by the population of each province.

Table 9

Estimated Benefits and Costs of Raising Fine Levies

				Benefits and Costs of Doubling Fine Rates			
	$\frac{\partial \ln DeathRate_i}{\partial WaterQuality_i}$	$\frac{\partial WaterQuality_i}{\partial \ln Dumping_i}$	$\frac{\partial \ln Dumping_i}{\partial \ln TaxRate_i}$	Deaths Averted Per Year	Extra Compliance Cost Per Year ($)	Cost per Averted Death	Cost Per Year of Life
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				(1)*(2)*(3)*(980,000[a])	(Costs[b])*(0.137)	(5)/(4)	(6)/20.12[c]
Overall Grade	0.097**	0.220***	0.815***	17,044	$ 504,888,032	$ 29,622	$ 1,472
	(0.044)	(0.075)	(0.21)				

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

Source : China Disease Surveillance Points Mortality Registration (DSP), China National Monitoring Center (2004), China Environmental Yearbooks (1993-2003), World Health Organization (2002)

Note : In column 1, I report the relationship between the log of the digestive cancer rate on the overall water grade (See Table 3). In column 2, I report the relationship between the water grade and the log of total dumping (See Table 8). In column 3, I report the relationship between the log of total dumping and the log of the effective fine levy on water dumping (See Table 8).[a]In column 4, I predict the number of lives saved by raising the levy rate, using the fact that the total annual deaths due to digestive cancers is roughly 980,000 (World Health Organization 2002). [b]In column 5, I predict the cost of compliance of raising the levy rate, using the fact that total reported compliance costs in 2000 was roughly $3.7 billion dollars (China Environmental Yearbook 2001), and the estimated elasticity of spending on cleanup of .137 (see Table 8). In column 6, I present the cost to firms in addtional spending required to avert one additional death from digestive cancer. [c]In column 7, I report this cost in terms of extra years of life expectancy, using the projection that decedents have on average 20.12 years of remaining life expectancy.

Appendix Table 1

Sample Means (and Standard Deviatons) for China Disease Surveilance Points by Urbanization, 1991-2000

Statistic	Poor Rural	Medium Rural	Rich Rural	Urban
	(1)	(2)	(3)	(4)
Share of Deaths Occuring in the Home	0.826	0.794	0.801	0.365
	(0.08)	(0.18)	(0.09)	(0.16)
Share of Deaths Occuring in the Hospital	0.073	0.109	0.104	0.490
	(0.05)	(0.17)	(0.09)	(0.15)
Share of Decedents Employed in Farming	0.919	0.837	0.788	0.013
	(0.09)	(0.20)	(0.23)	(0.04)
Average Education among Decedents	3.38	3.64	3.63	6.27
	(1.11)	(1.33)	(1.48)	(1.27)
Total Deaths Recorded	124,492	153,388	124,115	110,642
Total Person Years Covered	25,016,184	30,227,522	21,918,116	23,584,446
Crude Death Rate (Deaths/Persons)	0.0050	0.0051	0.0057	0.0047
Number of Sites	32	31	32	50

Source : China Disease Surveillance Points Mortality Registration (DSP)

Note : The table above summarizes differences across the 145 sites covered by the DSP. The sites form a nationally representative sample of deaths for China (see Yang et al. 2005). Employment and education for decedents is restricted to deaths among persons age 20 and older. Total deaths recorded is for the entire sample frame from 1991-2000. The total person years covered refers to the total number of individuals covered by each DSP site summed over the entire sample frame from 1991-2000. The DSP sites are classified into 4 income categories, and the table is composed of columns of poor rural (income category 1), medium rural (income category 2), and rich rural (income categories 3 and 4).

Appendix Table 2

Summary Statistics of Water Quality in China's Main River Systems

River System	Region	Overall	Ammonia Nitrogen	Biological Oxygen Demand	Lead	Oils	Perman-ganate	Volatile Phenol
Liao River	Northeast	3.10	2.65	1.45	1.03	1.68	2.36	1.33
Hai River	North	4.85	4.24	4.17	1.27	3.29	4.42	2.58
Huai River	North	4.63	3.94	3.54	1.13	2.71	4.00	1.75
Yellow River	North	2.08	1.88	1.08	1.00	3.46	1.54	1.00
Songhua River	Northeast	4.94	3.90	4.07	1.14	2.93	4.31	3.11
Fujian/Zhejiang	South	2.71	1.62	1.40	2.20	1.86	2.10	1.00
Yangtze River	South	2.58	1.94	1.39	1.09	1.38	1.67	1.12
Nu/Yarlung Zangbo	Southwest	3.90	2.71	2.30	1.00	1.87	3.68	1.85
Pearl River	South	2.85	2.43	1.39	1.13	1.68	1.97	1.20
Inward Flowing Systems	West	4.23	3.65	2.74	1.40	3.05	3.42	2.26

Source : China National Monitoring Center (2004)

Note : Higher numbers reflect lower water quality (1=best, 6=worst) and a greater concentration of the listed pollutants. These readings are taken from 484 water quality monitoring systems across China. The systems are displayed in descending order of overall water quality.

Appendix Table 3

Ordinary Least Squares (OLS) Regressions of Log of Digestive Cancer Rates on Water Grade by Category

	No Controls				With Controls			
Statistic	Digestive (all)	Esophageal	Stomach	Liver	Digestive (all)	Esophageal	Stomach	Liver
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Medium Quality Water (1=yes)	0.275	0.374	0.340	0.224	0.189	0.254	0.247	0.130
	(0.190)	(0.399)	(0.244)	(0.177)	(0.121)	(0.288)	(0.219)	(0.096)
Lowest Quality Water (1=yes)	0.388**	0.873**	0.472*	0.257*	0.329*	0.757**	0.413*	0.192*
	(0.170)	(0.353)	(0.235)	(0.147)	(0.162)	(0.342)	(0.243)	(0.102)
Average Education					-0.007	0.014	-0.070	0.016
					(0.045)	(0.114)	(0.066)	(0.030)
Share in Farming					-0.026	0.634	0.441	-0.520***
					(0.311)	(1.055)	(0.421)	(0.182)
Urban Site (1=yes)					-0.332	-0.056	-0.025	-0.630***
					(0.213)	(0.728)	(0.268)	(0.134)
Air Pollution					0.264**	0.729**	0.103	0.270**
					(0.127)	(0.292)	(0.184)	(0.113)
County Income Controls					Yes	Yes	Yes	Yes
Region Controls	No	No	No	No	Yes	Yes	Yes	Yes
R Squared	0.077	0.085	0.069	0.046	0.268	0.204	0.226	0.320

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

Source : China Disease Surveillance Points Mortality Registration (DSP), China National Monitoring Center (2004)

Note : N=145. See Table 3 for a description of the control variables included in columns 4-6. Standard errors are robust and clustered at the province level. Regressions are weighted by the population at each DSP site. The water grade is measured on a 6 point scale: the base category is drinkable water (grade I or grade II), the medium quality water category is water that is not drinkable but suitable for human contact (grade III), and the lowest quality water category is water that is not suitable for human contact (grade IV+). The coefficient for medium quality and lowest quality water are reported relative to drinkable water.

Appendix Table 4

Comparison of Age-adjusted Death Rates (per 100,000) Due to Cancer, China and the United States

	China		United States (Overall)		Asian Americans	
	Men	Women	Men	Women	Men	Women
All Cancers	137	80	228	157	137	92
Digestive	90	46	57	35	46	31
Lung	32	14	71	41	37	18
Other	15	20	101	82	54	43
% Digestive	66%	58%	25%	22%	34%	34%
% Lung	23%	17%	31%	26%	27%	20%
% Other	11%	25%	44%	52%	39%	47%

Source : China DSP 1991-2000, United States Cancer Statistics: 2004 Incidence and Mortality, Center for Disease Control (CDC)

Note : Statistics for the United States reported relative to US standard population in 2000, which is not directly comparable to the rates reported for China, which are adjusted using the age distribution of the 2000 China census, and would be slightly altered if evaluated using the US standard population. The US population is older than the Chinese population: the census for both countries in 2000 reflects that the US median age (41) is 4 years older than the Chinese median age (37).

Appendix Table 5

Industrial Wastewater Dumping and Manufacturing Output by Industry, China (1998-2000)

Industry	Dumping (10,000 tons) (1)	Percent of Dumping (2)	Output (million yuan) (3)	Percent of Output (4)	Pollution Intensity (2)/(4)
Chemicals and chemical products	1,081,982	18.9%	1,236,764	7.3%	2.6
Ferrous metal smelting and pressing	967,518	16.9%	318,901	1.8%	9.2
Paper and paper products	707,785	12.4%	814,550	4.7%	2.6
Electricity, gas, and water production	610,458	10.7%	471,520	2.7%	3.9
Food, beverages and tobacco	512,592	9.0%	1,512,839	8.7%	1.0
Mining	377,351	6.6%	523,205	3.0%	2.2
Textiles	357,017	6.2%	1,687,521	9.7%	0.6
Mechanical and electronic equipment	296,958	5.2%	4,645,496	26.8%	0.2
Construction materials	215,395	3.8%	757,880	4.4%	0.9
Other	597,352	10.4%	4,973,618	28.7%	0.4
Total	5,724,408	100.0%	16,942,295	100.0%	1.0

Source : Chinese Environmental Yearbooks (1998-2000), China Manufacturing Firm Database (1998-2000)

Note : Manufacturing output is based on the total value of output in yuan of firms with greater than 5 million yuan in sales. Dumping by industry is reported for firms registered with China's State Environmental Protection Agency (SEPA).

Appendix Table 6

OLS Regressions of Digestive Cancer Rate on Historical Industrial Output in the River Basin (1970-1990)

	Digestive (all)	Esoph-ageal	Stomach	Liver
	(1)	(2)	(3)	(4)
Log of Industrial Output in River Basin (1970-1990)	0.071** (0.03)	0.237*** (0.08)	0.103*** (0.04)	-0.0150 (0.03)
Controls	Yes	Yes	Yes	Yes
R Squared	0.238	0.191	0.204	0.299

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

Source : China Disease Suveillance Points Mortality Registration (DSP), China Statistical Yearbooks (1970-1990), China 2000 Census

Note: N=145. The dependent variable in each regression is the listed cause of death. The independent variable is the log of total gross industrial output value in the river basin containing the DSP from 1970 to 1990. This is calculated using annual province-level data and the distribution of manufacturing employment within the province's counties in the 2000 census. The included controls are the same as in Table 3.

Appendix Table 7

OLS Regressions of Log of Cause-Specific Death Rates on Monthly Rainfall (mm) by Access to Tap Water

	Low Tap Water Share	High Tap Water Share
	(1)	(2)
All Causes	0.001	-0.001
	(0.0011)	(0.0015)
Cancer	-0.006***	0.000
	(0.0018)	(0.0018)
Digestive (all)	-0.008***	0.000
	(0.0026)	(0.0022)
Lung	-0.007***	-0.001
	(0.0020)	(0.0017)
Other Cancer	0.001	0.002
	(0.0026)	(0.0019)
Non-Cancer (all types)	0.002	-0.002
	(0.0012)	(0.0015)
Observations	74	71

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

Source : China Disease Surveillance Points Mortality Registration (DSP), China National Monitoring Center (2004), China 2000 Census

Note : Each reported coefficient represents a separate regression, with monthly rainfall as the independent variable and the log of the death rate of the listed cause as the dependent variable. Columns 1 represents regressions for the 74 DSP points located in counties where a majority of households do not have access to tap water, and column 2 represents the 71 DSP points where a majority have tap water according to the 2000 census. All regressions include the controls shown in Table 3.