Water Pollution and Digestive Cancers in China

Avraham Ebenstein*

August 2008

Following China's economic reforms of the late 1970's, rapid industrialization has led to a deterioration of water quality in the country's lakes and rivers. China's cancer rate has also increased in recent years, and digestive cancers (i.e. stomach, liver) now account for 8.1% of fatalities (WHO 2001). This paper examines a potential causal link between surface water quality and digestive cancers by exploiting large regional variation in water quality, which is driven in part by plausibly exogenous variation in rainfall patterns. I also exploit variation driven by the presence of manufacturing in an upstream river basin, which increases water pollution downstream and is plausibly exogenous to the digestive cancer rate downstream. Using a sample of 145 mortality registration points in China, I find using OLS that a deterioration of the water quality by a single grade is associated with a 14% increase in the death rate due to digestive cancer. Using 2SLS with rainfall and upstream manufacturing as instruments, I find that deterioration by a single grade is associated with a 30% increase in the digestive cancer death rate. The analysis rules out other potential explanations for the observed correlation, such as smoking rates, dietary patterns, and air pollution. The paper concludes with a brief discussion of policy options for China in light of the results.

---

* Robert Wood Johnson Scholar in Health Policy, Harvard University. I would like to thank Alison Flamm 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. Email: aebenste@rwj.harvard.edu.

1 Introduction

During the 1980's and 1990's, 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 side effects, including a severe deterioration in the quality of the country's rivers and lakes. Excessive use of fertilizers by farmers and irresponsible manufacturing practices has rendered the water in many lakes and rivers unfit for human consumption. China's water monitoring system indicates that roughly 70% 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 witnessed an increase in rural cancer rates, with digestive cancer death rates increasing 20% during the 1990s.[1] Following this increase, stomach cancer and liver cancer are respectively the 4th and 6th leading causes of death, and account for 8.1% of all fatalities (World Health Organization 2001). 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 2004), 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 Grant 2005) and diarrhea (Jalan and Ravalion 2003), and access to cleaner water may lower infant mortality as well (Galiani et al. 2005). The connection between water quality and cancer, however, has not been fully explored. A limited 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 Kantor (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 however represents an almost ideal context to investigate a causal association between contaminated water and digestive cancer. First, in most developing countries reliable data on pollution and mortality are unavailable. However, China's

[1] Author’s calculation.

efforts in the late 1980's 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.[2]

In the following 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.[3] Using 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 long-term averages of monthly precipitation.[4] I am also able to observe manufacturing output in each basin, including the basins flowing directly into the basin including the DSP site, which affects the water grade in the basin but should otherwise be exogenous to the digestive cancer rate at the site.

By comparing DSP sites in basins with better and worse water quality, I estimate using OLS that a deterioration of water quality by a single grade increases the incidence of digestive cancers by 14 percent, and this results is only changed slightly (13 percent)

---

[2] 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.
[3] 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.
[4] 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).

when control variables are added for air quality and other potential confounding factors also associated with industrialization (whether the site is urban, share employed in manufacturing, etc.). By exploiting plausibly exogenous variation in rainfall within each river basin, as well as the presence of manufacturing in the river basin upstream, I estimate 2SLS models of the relationship between digestive cancer rates and water quality, and the estimates indicate that a 1 grade increase in water contamination increases the digestive cancer rate by 30 percent, providing 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 appears no strong relationship in China between regional variation in smoking rates or dietary patterns and water quality.

The next section provides background information on China's waterways and regional variation in quality. Section 3 describes the data in more detail, and on production, emissions, water quality, cause-specific mortality, and the matched-pair design of the analysis. Section 4 reports the empirical results. Section 5 concludes.

2 Background

The pollution levels in China's water bodies are almost without historical precedent, and in spite of recent efforts to reduce water dumping by manufacturing firms, more than half of China’s surface water was found unfit for human use (World Bank 2007). 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.

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. Nonpoint 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 (non-point source), and manufacturing firms have dumped inorganic compounds into water as part of their production processes. When these chemicals drain into waterways, it stimulates a river's algal growth beyond its natural speed (known as eutrophication), and the water becomes populated by cyanobacteria (blue-green algae) such as microcystins (Davis and Masten 2004). These compounds in particular are thought to be carcinogenic, and have been linked directly to liver cancer (Codd 2000).

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 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 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 the north, is the most polluted basin in the country with 57% 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, since industrial areas of southern China have been largely spared the effects of water pollution due to heavier rainfall.

In China, the degradation of waterways has also led areas *without* industrial activity to experience a decline in water quality. Within a watershed, downstream river segments are contaminated by upstream sources of wastewater and this was the case in a famous episode 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 form, and dead fish." In this way, regions downstream of industrial firms 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.[5] In the next section, I describe how I will attempt to exploit both regional variation in water quality, as well as the flow dynamics of water, to estimate the causal link between water quality and cancer incidence.

3 Data

The analysis of mortality patterns in China are 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 for the 10 million residents of the points, and due to careful sample selection it yields an annual sample of deaths that mirror 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. A summary of cause-specific death rates during the sample period are shown in Table 1.

China's severe problems with water pollution in the 1980s following the country's reforms led to the creation of a national water monitoring system. The World Bank produced a comprehensive assessment of water quality patterns using this system, and described water patterns from 1991-2005 from the readings, which reflected persistent problems in the Northern river systems. The analysis presented here relies on the 2004 readings, which report water quality readings from China's nine river systems recorded for 484 geographic points. 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 center. The projects provides a suite of geo-referenced data sets that are created using a Digital Elevation Model (DEM) in which

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

China can be separated into a set of 989 basins, and a smaller set of larger basins. Satellite imagery is also 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 these cancer rates, and may be correlated with factors that induce water pollution (e.g. manufacturing). The measure is taken between zero and 1, with higher numbers representing high optical depth, implying the presence of more particulates and lower clarity. 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.[6] In order to examine how precipitation may affect water quality, I include measures of monthly rainfall collected by the Global Precipitation Climatology Center (2008). These measures are calculated by river basins in a manner similar to the air quality, where I use GIS software and average the rainfall measure across the area in the same basin as the DSP point. 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 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 IDs based on the topology of the land surface. Since it is hierarchical, it is possible to identify the watershed immediately downstream of each watershed by its numbering. This property is exploited to consider the impact of industrial activity upriver on cancer rates at DSP points in basins subordinate to the basin where the emissions are observed. The data on emissions are proxied by total value of manufacturing output, which is observed for each of China's counties (2,800+) at a particular latitude and longitude, and can therefore be placed in a river basin. The measure of upstream manufacturing is the total value of output in the level 4 basins that are upstream of the basin containing the DSP site.

[6] 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).

4 Empirical 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. 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 increases (e.g. IV to V) the digestive cancer rate by 14 percent. The coefficients are 35 percent, 14 percent, and 9 percent for the impact of water quality on esophageal, stomach, and liver cancer respectively, with the coefficients statistically significant for all but liver cancer.

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, and an imputed measure of ambient air quality, where a higher number reflects more particulates (and therefore worse air quality). The results are only changed slightly, with the estimates implying that water quality eroding by one grade induces a 13 percentage point increase in the digestive cancer rate. The estimates for the aforementioned types of digestive cancer are 34, 14, and 8 percentage points respectively. It may be unsurprising that the coefficients are largely unchanged by including controls, since Table 2 reflects that much of the water quality variation is regional, and the regions do not exhibit large differences in urbanization or air quality. Table 3 also indicates that air quality also has a statistically significant relationship with digestive cancer rates, with an increase in the particulate index variable (that varies from 0-1) by 0.10 induces a 7% increase in the digestive cancer rate. This may reflect a causal link between contaminants in the air and the likelihood of tumors forming in digestive organs (Jerret 2005 et al.), or 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 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 upstream manufacturing output. In the first column, I examine the first-stage relationship between monthly rainfall in milliliters, upstream manufacturing output, 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 quality is induced by variation in rainfall patterns. The impact of an additional million yuan of manufacturing output in the river basins directly upstream is associated with an increase in the water grade by 0.001 units, and the relationship is statistically significant at the 5% level. An F test of the joint significance of the two instruments is 9.48, which is highly significant as well (p-value=.0006).

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 30%. The estimates for esophageal cancer and stomach cancer imply that water quality declining by 1 grade increases the incidence of these diseases by 104% and 48% respectively, and both are statistically significant at the 5% 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.

In Table 5, I present an additional set of OLS regressions where I examine whether the relationship between water quality and digestive cancers is observed differently by gender or by particular pollutant. The table reflects a consistency between the estimated impact for both men and women, implying that declines in water quality operate similarly by gender. For example, an increase in the water grade by 1 unit is associated with a 33 percentage point increase in the esophageal cancer rate for men, and a 31 percentage point increase for women. The impact of overall water quality on stomach cancer is smaller (15 percentage points for men, 13 percentage points for women), and both coefficients are statistically significant. Among the other chemicals that appear related to these cancers is ammonium nitrate (from fertilizer run-off), biological oxygen demanding agents (from manufacturing waste), and oils (from firm waste). Note that the listed nitrogen compounds (ammonium nitrogen, nitrate) appear highly related to the digestive cancer, suggesting that efforts to control fertilizer use may improve China's waterways and lower the health risks that arise as a consequence of excessive algae growth. In summary, the results suggest that contaminated water results in an increase in digestive cancer for both men and women, and that both agricultural and non-agricultural waste appear to have a pronounced effect on digestive cancer rates.

In Table 6, I attempt to examine 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, and 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.[7] 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 factor in determining digestive cancer rates (Kono et al. 1996), it does not appear that this is correlated with water quality, and is less likely to be biasing the estimated effect of water quality on cancer rates.

In Table 7, 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 12 percentage point increase in the cancer rate (significant at 5%), but has a small and statistically insignificant

---

[7] National surveys reflect that smoking rates for men are in excess of 75%, but fewer than 8% of women smoke. 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. (Gonghuan, Yang. 1997. "1996 National Prevalence Survey on Smoking Patterns". China Science and Technology Press.).

relationship to the death rate from other leading causes of death such as heart disease or stroke. Interestingly, the fact that the overall death rate is uncorrelated with water quality in spite of water quality's impact on cancer rates suggests that other compensating effects of industrialization may mitigate the increase in cancer rates, such as greater wealth and better access to health care.

5 Conclusion

Despite an increase in clean-up efforts in recent years, overall degradation of China's waterways continues. Since the early 1990s, there has been some increase in both the number and capacity of industrial wastewater treatment facilities. Efforts have begun to build modern wastewater treatment plants which will lower the pollutants discharged into the rivers (World Bank 2007). Regulations aimed at controlling industrial water dumping have been more strictly enforced in recent years, and have been shown to affect the pollution intensity of firm production methods (Wang and Wheeler 2000). In combination with evidence presented here regarding the health consequences of water pollution, policies that increase the costs to pollution for firms may better reflect the societal costs of the dumping. Likewise, the results indicate that chemicals from agricultural runoff have grave consequences on digestive cancer rates. Efforts to clean waterways or discourage excessive fertilizer should also be considered, due to the strong relationship observed between nitrogen compounds and cancer rates.

China's historical record provides a useful natural experiment to discern water pollution's relationship with digestive cancer, but it is also a pressing issue in its own right. While China's economy has grown rapidly following economic reform, the adverse health effects of pollution threaten to mitigate the health benefits of the country's newfound wealth. In the context of water pollution, many of China's residents who do not have access to tap water are the least likely to enjoy the benefits of growth and are now the most exposed to the consequences. In light of the evidence of water pollution's costs to both acute illnesses and to cancer rates, policymakers in China should continue to pursue policies that aggressively promote better waste-water management and widen the availability of treated water to rural areas. Recent reports indicate no improvement in water quality in key river basins, which may continue to undermine China's efforts to improve public health. As China considers how to spend its newfound wealth, the results presented here indicate that improving the country's waterways may be an effective way to improve quality of life and lower the incidence of digestive cancer.

6 Bibliography

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

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

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.

Davis, Mackenzie L. and Susan J. Masten. 2004. *Principles of environmental engineering and science*. McGraw-Hill.

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.”

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

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.

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

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.

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

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.

Lipscomb, Molly and Ahmed Mushfiq Mubarak. 2007. “Decentralization and Water Pollution Spillovers: Evidence from the Re-drawing of County Boundaries in Brazil”. Draft.

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

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

National Aeronautics Space Administration. 2008. See http://disc.sci.gsfc.nasa.gov/. Optical air readings from MODIS sensor on Terra satellite.

Pfafstetter, O., 1989, "Classification of hydrographic basins: coding methodology", unpublished manuscript, DNOS.

Schneider, U., T. Fuchs, A. Meyer-Christoffer, B. Rudolf. 2008. “Global Precipitation Analysis Products of the GPCC.” Global Precipitation Climatology Centre (GPCC) Deutscher Wetterdienst, Offenbach a. M., Germany.

Wang, Hua and David Wheeler. 2000. “Endogenous Enforcement and Effectiveness of China’s Pollution Levy System.” The World Bank: Development Research Group, Infrastructure and Environment. Policy Research Working Paper 2336.

World Bank. 2006. *Water Quality Management Policy and Institutional Considerations.* Washington, DC: World Bank.

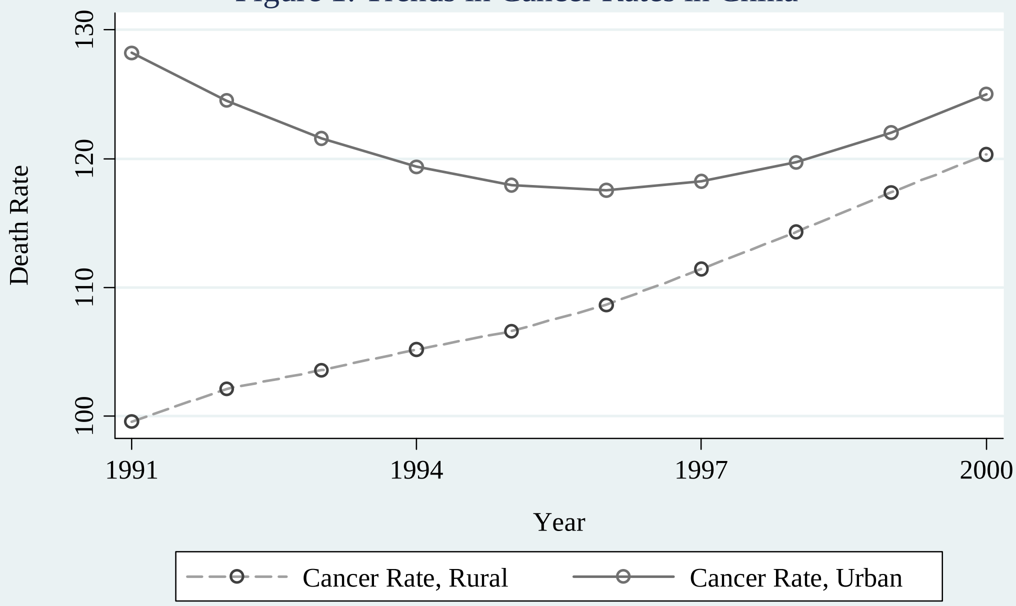
World Bank and SEPA. 2007. *Cost of Pollution in China: Economic Estimates of Physical Damages*. Washington, DC: World Bank.

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

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.

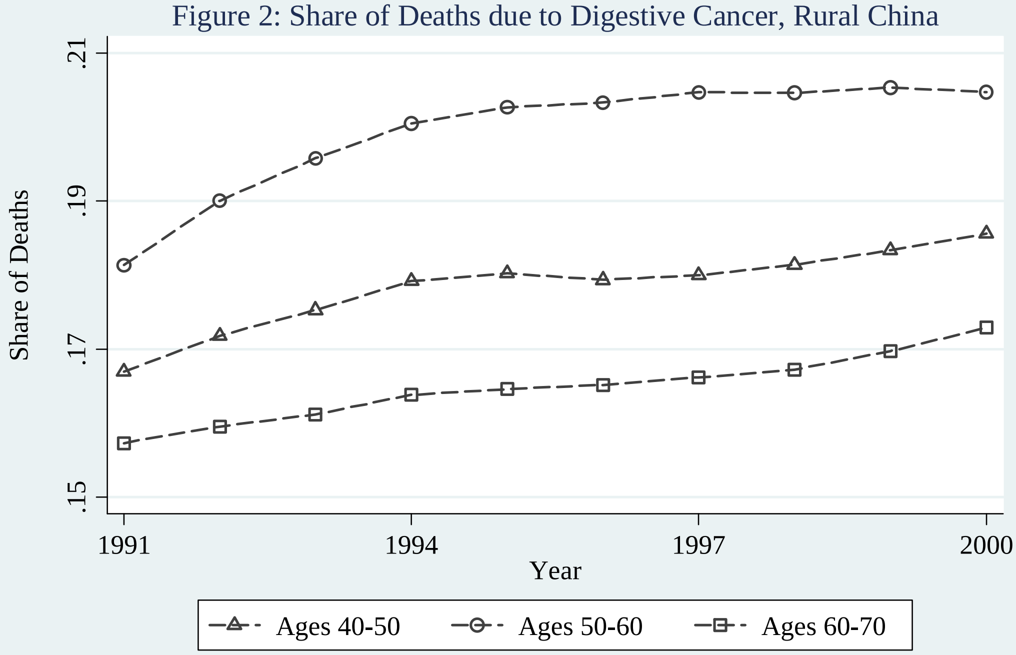
7 Figures and Tables

Figure 1: Trends in Cancer Rates in China



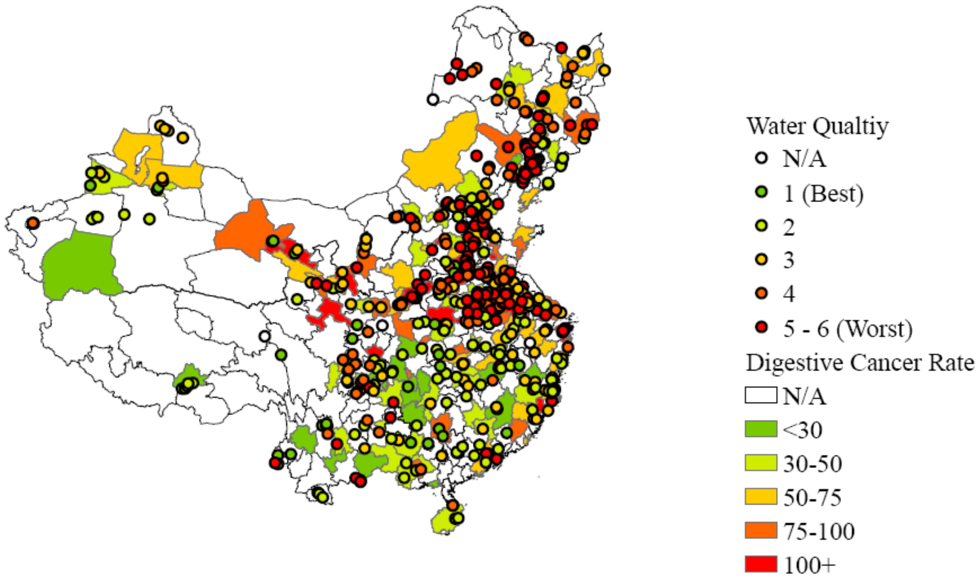
Source: China DSP. Graphs displayed are created using STATA lowess smoother.
Death rates are per 100,000 persons.

Figure 2: Share of Deaths due to Digestive Cancer, Rural China



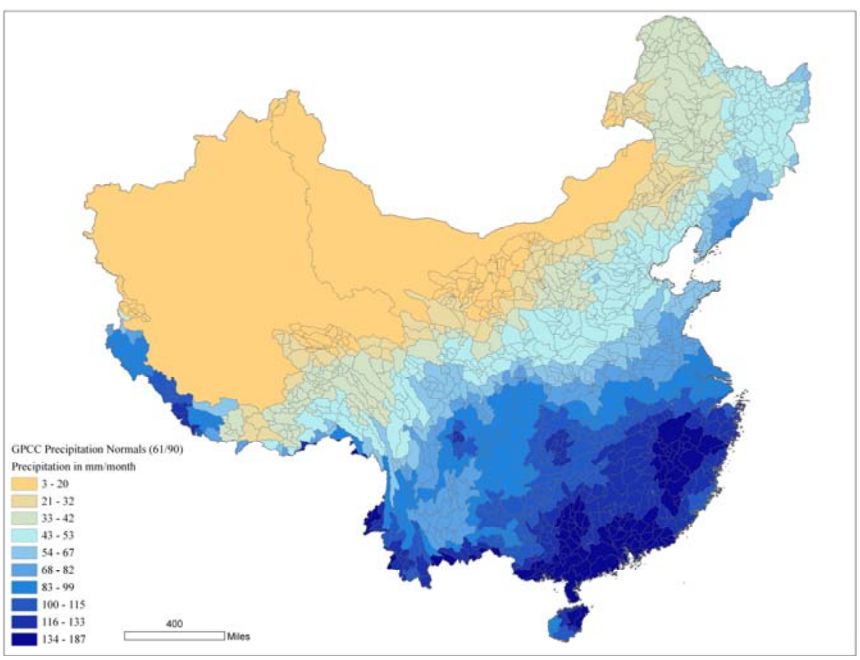
Source: China DSP. Graphs displayed are created using STATA lowess smoother.

Figure 3: Water Quality and Digestive Cancer Rates in Select Locations, 1991-2000



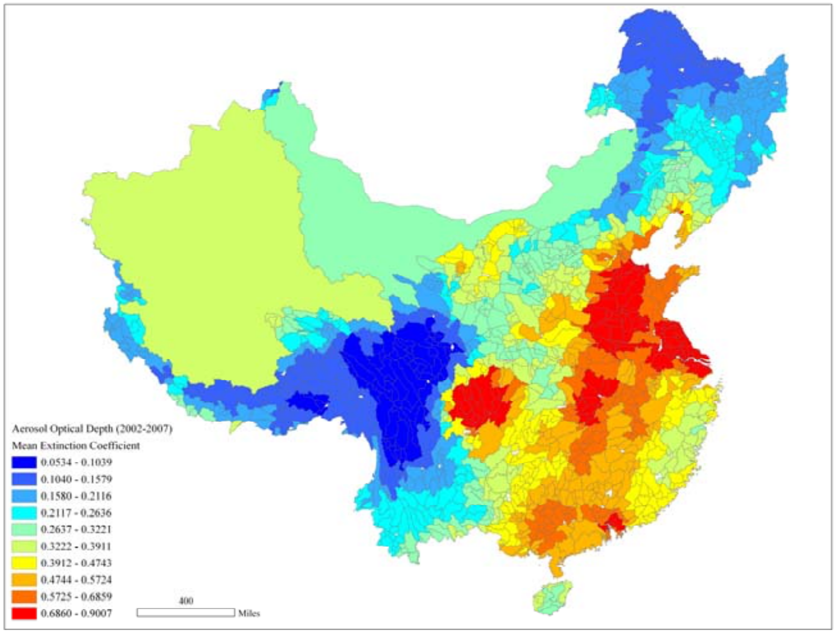
Source: China Center for Disease Control. Disease Surveillance Points (DSP) comprise a nationally representative sample of mortality for China. Rates reported per 100,000 and are age and sex adjusted using the 2000 census population structure.

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



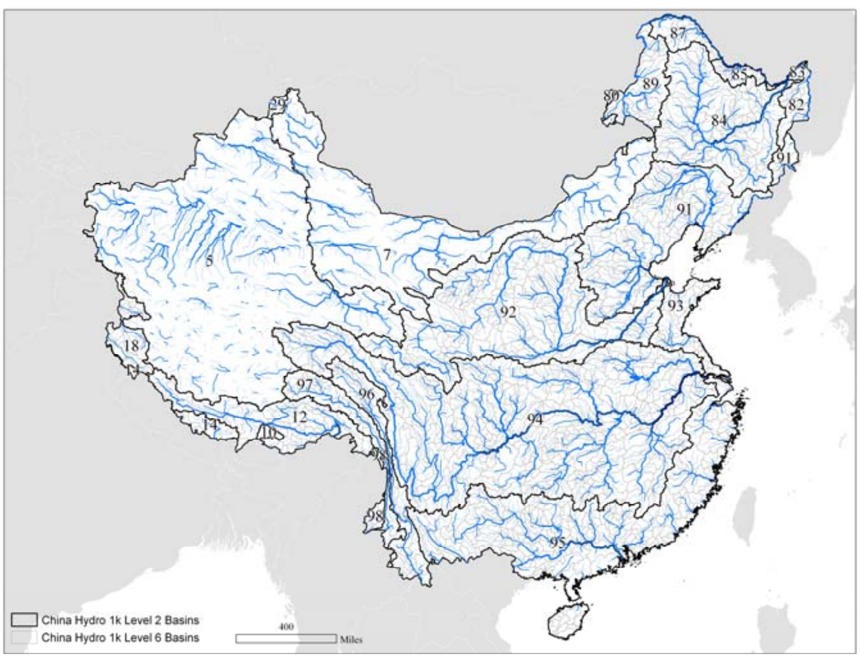
Source: Global Precipitation Climatology Center.

Figure 5: Air quality patterns in China



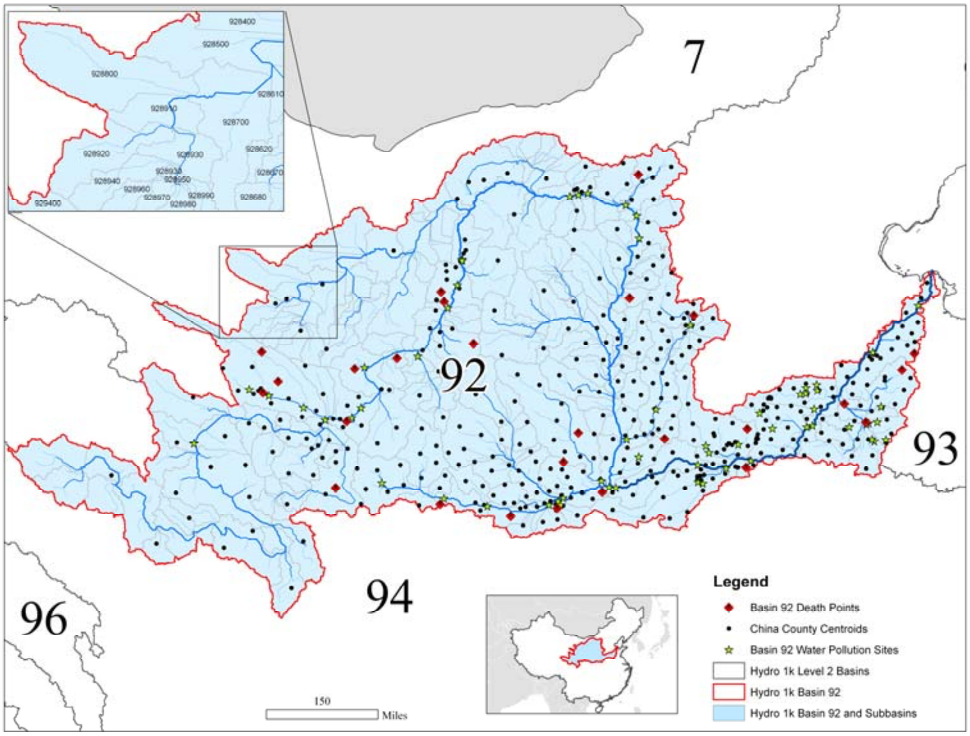
Source: NASA satellite imagery.

Figure 6: Main river basins in China



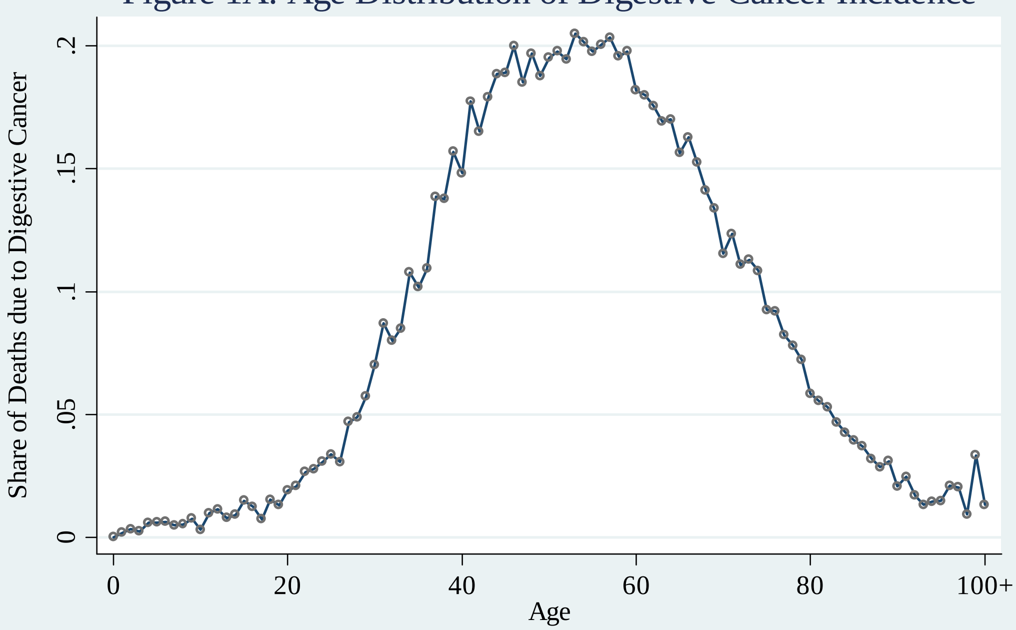
Source: Hydro 1k Project. The above figure reflects the principal 2-digit basins (or watersheds) that comprise the hydrological surface of China. The dark lines reflect the breakdown of the 2-digit basins, and the lighter outline is the breakdown of China into 989 lower-level basins.

Figure 7: Example of a River Basins



Source: China Hydro 1k.

Figure 1A: Age Distribution of Digestive Cancer Incidence



Source: China DSP.

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	601	460
Cancer	133	150	78	84
Digestive Cancers	82	69	42	33
Lung Cancers	23	43	10	18
Other Cancer	29	38	26	33
Heart	133	100	134	91
Stroke	125	125	107	100
Respiratory Illnesses	126	72	122	58
Accidents / Violence	91	49	59	31
Other	118	102	101	96
Panel 2: Death Rates for Types of Digestive Cancer				
Esophageal Cancer	18	9	10	4
Stomach Cancer	27	19	15	9
Liver Cancer	29	27	12	10
Other Digestive Cancers	8	13	6	10

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 includes colon cancer, intestinal cancer, and pancreatic cancer. The reported death rates are the average rates for the 145 sites, weighted by the total population at each site.

Table 2

Sample Means for Disease Surveillance Points by Region

Statistic	North	South	Overall
	(1)	(2)	(3)
Digestive Cancer Rate	68.6	55.0	59.7
Overall Water Grade	4.46	3.25	3.67
Ammonium Nitrate	3.78	2.59	3.01
Biological Oxygen Demand	3.61	1.64	2.32
Nitrite	4.46	3.25	3.67
Oils	3.16	1.70	2.21
Permanganate	4.01	2.35	2.93
Urban site (1=yes)	0.24	0.23	0.23
Average Years of Education	3.92	4.33	4.19
Share in Manufacturing	0.06	0.10	0.09
Share in Farming	0.71	0.63	0.65
Air Pollution Reading	0.53	0.49	0.51
Monthly Rainfall (mm)	50.7	99.6	82.5
Upstream Manufacturing (millions)	63.8	8.8	28.0
# of Sites	66	79	145

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

Note : 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. Upstream manufacturing is based on the total value of output in yuan of firms with greater than 500,000 yuan in sales. 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

Statistic	No Controls: Digestive (all)	Esophageal	Stomach	Liver	With Controls: Digestive (all)	Esophageal	Stomach	Liver
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Water Grade (1=best, 6=worst)	0.135**	0.348***	0.141**	0.090*	0.127**	0.338***	0.139**	0.078**
	(0.053)	(0.102)	(0.065)	(0.047)	(0.050)	(0.101)	(0.067)	(0.038)
Urban (1=yes)					-0.023	-0.003	-0.092	0.007
					(0.045)	(0.099)	(0.074)	(0.028)
Average Education					-0.230	0.734	0.166	-0.715***
					(0.29)	(0.93)	(0.46)	(0.21)
Share in Farming					0.655*	1.360	0.791	0.503
					(0.338)	(1.142)	(0.576)	(0.444)
Share in Manufacturing					-0.356**	0.101	0.002	-0.760***
					(0.167)	(0.611)	(0.247)	(0.135)
Air Pollution					0.739**	1.463**	0.432	0.834**
					(0.358)	(0.663)	(0.486)	(0.343)
R Squared	0.096	0.140	0.064	0.054	0.177	0.214	0.122	0.219

* 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 point 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 point reflects the average water quality among monitoring sites in the same river basin.

Table 4

Two-stage Least Squares (2SLS) Regressions of Log of Digestive Cancer Rates on Water Grade using Annual Rainfall and Upstream Output as Instruments

	First-Stage	Two-stage Least Squares			
Statistic	Water Grade	Digestive (all)	Esophageal	Stomach	Liver
	(1)	(2)	(3)	(4)	(5)
Monthly Rainfall	-0.012**				
	(0.005)				
Upstream Output	0.001*				
	(0.0005)				
Water Grade (1=best, 6=worst)		0.296**	1.035***	0.480**	0.024
		(0.144)	(0.316)	(0.196)	(0.136)
Urban (1=yes)	-0.343	-0.349*	0.129	0.016	-0.763***
	(0.870)	(0.194)	(0.523)	(0.415)	(0.143)
Average Education	0.013	-0.018	0.015	-0.083	0.006
	(0.11)	(0.05)	(0.10)	(0.08)	(0.03)
Share in Farming	-1.585	0.012	1.728	0.651	-0.792**
	(1.032)	(0.354)	(1.061)	(0.687)	(0.299)
Share in Manufacturing	-4.474***	1.473*	4.716**	2.431*	0.244
	(1.249)	(0.801)	(2.212)	(1.268)	(0.813)
Air Pollution	1.412	0.554	0.702	0.060	0.893**
	(0.836)	(0.397)	(0.835)	(0.549)	(0.395)
F Test of Instruments	9.48***				

* 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 output upstream of the basin containing the DSP site. The F statistic for the joint significance of the two instruments is reported (9.48), and the instruments also pass a Sargan-Hansen over-identification test, failing to reject the null hypothesis of their validity. 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.

Table 5

OLS Regressions of Log of Digestive Cancer Rates on Water Grade by Chemical

	Men			Women		
	Esophageal	Stomach	Liver	Esophageal	Stomach	Liver
Chemical	(1)	(2)	(3)	(4)	(5)	(6)
Overall Grade	0.326***	0.146**	0.067*	0.306**	0.130*	0.109***
	(0.10)	(0.07)	(0.04)	(0.14)	(0.07)	(0.04)
Ammonium Nitrate	0.305***	0.127**	0.051	0.313***	0.117**	0.089***
	(0.10)	(0.06)	(0.03)	(0.12)	(0.06)	(0.03)
Biological Oxygen Demand	0.303***	0.143***	0.046	0.270**	0.145***	0.086**
	(0.08)	(0.05)	(0.03)	(0.11)	(0.05)	(0.03)
Mercury	0.063	0.038	-0.020	-0.027	0.038	0.029
	(0.22)	(0.11)	(0.06)	(0.29)	(0.13)	(0.10)
Nitrite	0.326***	0.146**	0.067*	0.306**	0.130*	0.109***
	(0.10)	(0.07)	(0.04)	(0.14)	(0.07)	(0.04)
Oils	0.312***	0.164**	0.014	0.264**	0.150**	0.050
	(0.10)	(0.07)	(0.04)	(0.13)	(0.07)	(0.04)
Permanganate	0.323***	0.133**	0.059	0.299**	0.125**	0.110***
	(0.09)	(0.06)	(0.04)	(0.12)	(0.06)	(0.04)

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

Note : N=145. Each reported coefficient represents a separate regression. The regressions are estimated using the column variable as the dependent variable and the water grade measure in the row as the independent variable. Overall grade and chemical pollution are graded on a 6-point scale, with 1 being the highest quality and 6 being the most polluted. The dependent variable is the logarithm of the age and sex adjusted death rate by cause. The independent variable is a measure of water pollution, which is the amount of the chemical found in readings among water surveillance points in the river basin. All specifications include the control variables shown in Table 2. Standard errors are robust and clustered at the province level.

Table 6

Smoking and Dietary Habits by Water Grade in China

	Smoking Rates		Dietary Patterns				
Water Grade	Males	Females	Caloric Intake	Carbo-hydrates	% Fat	% Protein	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Level 1 (Best)	0.732	0.034	2,172	15.21	2.89	2.86	79.04
Level 2	0.705	0.066	2,376	15.12	2.96	2.84	79.08
Level 3	0.697	0.025	2,303	14.75	3.06	2.92	79.28
Level 4	0.705	0.034	2,238	15.38	2.75	2.97	78.90
Level 5	0.704	0.059	2,311	16.13	2.41	2.99	78.47
Level 6 (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 Household 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 7

Impact of Water Pollution on Log of Death Rates by Cause in China, 1991-2000

	All Causes	Cancer	Heart	Stroke	Resp-iratory	Violent/ Accidents	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Overall Grade (1=best, 6=worst)	0.008	0.121***	0.007	0.054	-0.044	-0.051**	-0.034
	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.02)	(0.03)
Ammonium Nitrate	0.004	0.106***	-0.008	0.041	-0.024	-0.034	-0.037
	(0.02)	(0.03)	(0.04)	(0.04)	(0.05)	(0.02)	(0.03)
Biological Oxygen Demand	0.004	0.094***	0.018	0.038	-0.059	-0.041**	-0.040*
	(0.02)	(0.03)	(0.04)	(0.03)	(0.05)	(0.02)	(0.02)
Mercury	(0.05)	(0.08)	(0.07)	(0.06)	(0.09)	(0.05)	(0.05)
Nitrite	0.008	0.121***	0.007	0.054	-0.044	-0.051**	-0.034
	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.02)	(0.03)
Oils	0.010	0.088**	0.082*	0.036	-0.069	-0.054**	-0.043
	(0.02)	(0.04)	(0.05)	(0.03)	(0.05)	(0.02)	(0.03)
Permanganate	-0.002	0.106***	0.014	0.052	-0.066	-0.068***	-0.060**
	(0.02)	(0.03)	(0.05)	(0.03)	(0.06)	(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. The regressions are estimated using the column variable as the dependent variable and the water quality measure in the row as the independent variable. Each water reading is graded on a 6-point scale, with 1 being the highest quality and 6 being the most polluted. The dependent variable is the arithemetic average of the death rate by cause for each DSP site (see Table 2). The independent variable is a measure of water pollution, which is calculated as the average chemical reading among the water surveillance points in 2004 for the watershed containing the DSP. All specifications include a control for whether the site is classified as urban or rural. Standard errors are robust and clustered at the province level.

Appendix Table 1

Sample Means (and standard deviaions) 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 cites are chosen to form a nationally representative sample of deaths for China, with sites being chosen to match China's census in 1990 (see Yang et al. 2005 and the data section). Employment and education for decedents restricted to deaths among persons age 20-100. 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 point summed over the entire sample frame from 1991-2000.