

NBER WORKING PAPER SERIES

(RE)SCHEDULING POLLUTION EXPOSURE:  
THE CASE OF SURGERY SCHEDULES AND PATIENT MORTALITY

Jialin Huang  
Jianwei Xing  
Eric Zou

Working Paper 28708  
<http://www.nber.org/papers/w28708>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
April 2021, Revised July 2021

We are extremely grateful for Professor Ling Li at National School of Development, Peking University for generous help with data access. We thank Karen Brandon, Shengmao Cao, Tatyana Deryugina, Qing Gong, Elaine Hill, Wei Huang, Anupam Jena, Shanjun Li, Christopher Worsham, anonymous referees, and seminar participants at the ASHEcon conference and Xiamen University for helpful comments. Junting Chen provided excellent research assistance. All errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Jialin Huang, Jianwei Xing, and Eric Zou. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

(Re)scheduling Pollution Exposure: The Case of Surgery Schedules and Patient Mortality  
Jialin Huang, Jianwei Xing, and Eric Zou  
NBER Working Paper No. 28708  
April 2021, Revised July 2021  
JEL No. C44,I18,O13,Q53

### **ABSTRACT**

Many human activities can be strategically timed around forecastable natural hazards to mute their impacts. We study air pollution shock mitigation in a high-stakes healthcare setting: hospital surgery scheduling. Using newly available inpatient surgery records from a major city in China, we track post-surgery survival for over 1 million patients, and document a significant increase of hospital mortality among those who underwent surgeries on days with high particulate matter pollution. This effect has two special features. First, pollution on the surgery day, rather than exposure prior to hospitalization, before or after the surgery, is primarily explanatory of the excess mortality. Second, a small but high-risk group – elderly patients undergoing respiratory or cancer operations – bears a majority of pollution’s damages. Based on these empirical findings, we build and analyze a model of hospital surgery scheduling. For over a third of the high-risk surgeries, there exists an alternative, lower-pollution day within three days such that moving the surgery may lead to a Pareto improvement in survival.

Jialin Huang  
National School of Development  
Peking University  
No.5 Yiheyuan Road  
Beijing  
100871  
China  
huangjl@pku.edu.cn

Eric Zou  
Department of Economics  
University of Oregon  
1415 Kincaid Street  
Eugene, OR 97403  
and NBER  
ericzou@uoregon.edu

Jianwei Xing  
National School of Development  
Peking University  
No.5 Yiheyuan Road  
Beijing 100871  
China  
jerryxing@nsd.pku.edu.cn

# 1. Introduction

Scheduling activities around looming natural hazards is an important human adaptation strategy in the face of the changing environment. Today, we adjust daily schedules to changes in weather, thanks to the development of weather forecasting technologies. This paper eyes on the potential of adaptation with respect to air pollution, another environmental hazard that a growing number of technologies can provide “nowcasts” and forecasts about. Recent economic analysis reveals people are already using such information to arrange activities, such as outdoor recreation, to avoid exposure ([Cutter and Neidell, 2009](#); [Neidell, 2009](#); [Graff Zivin and Neidell, 2009](#)). In this paper, we further explore the potential value of such adaptation and adjustment of scheduled activities in a high-stakes healthcare context. We examine how high levels of pollution that coincide with the date of surgery affect patients’ post-surgery mortality risk, and whether adjustments in hospital surgery schedules could improve survival.

We highlight two broad motivations for this study. First, while air pollution is often seen as an *outdoor* problem, indoor exposure may be surprisingly high in many contexts. Modern surgery rooms use high-efficiency filtration technologies to reduce contamination of air to extremely low levels ([Mangram et al., 1999](#); [American Institute of Architects, 2001](#); [Chinn and Schulster, 2003](#)). However, the general hospital indoor environment – unlike a dedicated surgery room – does not have elaborate filtration systems that prevent the penetration of air pollution from outdoor. In fact, hospitals sometimes *prefer* to have wards open to the outside air to help control infection. Our review of available studies on hospital wards air quality suggests a high indoor-outdoor fine particulate matter (PM<sub>2.5</sub>) concentration ratio of up to 0.8 (Section 2.1). Concerns about such exposure may be particularly high for patients who are ill, and those who are recovering from a recent procedure. On the other hand, many surgical procedures are arranged days ahead, making patients’ exposure to pollution *pre-scheduled* in nature. This raises the possibility of using strategic surgery scheduling to mitigate pollution exposure at such a sensitive time, and for any particularly vulnerable patients.

Second, knowledge about the adverse determinants of post-surgery death (known as *perioperative mortality*) – and ways to mitigate them – is of academic and clinical importance in its own right. Death is the most feared consequence of surgery care. An estimated 30 percent of the global burden of disease requires surgery care ([Shrime et al., 2015](#)). U.S. and European data show that 5 to 10 percent of patients die following an inpatient surgery ([Ghaferi, Birkmeyer, and Dimick, 2009](#); [Pearse et al., 2012](#)). Despite being a high-stakes health outcome and a crucial marker of healthcare delivery quality, current understanding on risk factors of perioperative mortality is limited ([Bainbridge et al., 2012](#)). Existing studies only examine basic patient demographics and health conditions such as the presence of chronic diseases (e.g., [Alves et al., 2005](#); [Bozic et al., 2012](#)) and healthcare provider characteristics (e.g., [Aiken et al., 2002](#);

Sahni et al., 2016). A key message of our work is that environmental conditions vary considerably from day to day, which may provide leeway to adjust scheduling to take advantage of better air quality conditions. As we show in this paper, incorporating air pollution levels as one consideration in setting surgery schedules may lead to better survival among certain vulnerable patients.

Our study examines Guangzhou, a major city in China with a population of 15 million. Several features of the city pose an opportunity for our research questions. First, air quality in Guangzhou exhibits substantial day-to-day fluctuations, with a mean  $PM_{2.5}$  concentration of  $36.5 \text{ ug/m}^3$  (standard deviation =  $19.8 \text{ ug/m}^3$ ).<sup>1</sup> Patients with observably similar characteristics undergo similar surgeries on days with substantially different air quality. This plausibly exogenous variation in pollution allows us to study its impact on patient outcomes. Second, Guangzhou is a city with abundant, advanced healthcare resources. Indeed, it is recognized as the premier healthcare destination in southern China. This setting helps us better tease out the impact of pollution shocks from impact of healthcare resource constraints that could also have an independent impact on surgical outcomes (Meara et al., 2015; Guidetti, Pereda, and Severnini, 2020). Finally, Guangzhou is one of the first cities in China to implement a healthcare informatics reform. Our analysis is based on newly available, rich microdata about patients and the surgical procedures they were provided in operations conducted in all 23 major (“3-A”) hospitals in the city. These microdata incorporate information from more than 1.3 million surgical records spanning 2014 to 2017.

Our analysis has two parts. **The first part** is a reduced-form econometric exercise that leads to our headline finding: High ambient air pollution levels on the day of a patient’s surgery lead to significantly worse post-surgery survival outcomes. We show this with both a classic Kaplan-Meier survival framework, and a fixed effects regression approach that exploits plausibly exogenous day-to-day fluctuations in air pollution levels by controlling flexibly for patient, hospital, and surgery characteristics, and for atmospheric weather patterns and seasonality. We find that a log unit increase in surgery-day  $PM_{2.5}$  results in an increase of mortality rate by 0.428 per 1,000 patients over the month following the surgery. This effect size converts to about 1 percent increase in surgery patient mortality per  $10 \text{ ug/m}^3$  increase in  $PM_{2.5}$ .

A main threat to our research design is whether the pollution effect is a result of patient selection. For example, high pollution events may lead sicker people to seek treatment at a hospital; surgery case mix or the severity of patients’ conditions may differ across days with high and low levels of pollution. Our study population, however, consists of patients who had *already* been admitted to the hospital and, on average, had spent four days in hospital between the admission and the surgery. Thus, the extent of patient

---

<sup>1</sup> To put these statistics in perspective, the Los Angeles County of California – one of the U.S. counties with the highest  $PM_{2.5}$  concentration – had a mean  $PM_{2.5}$  level of  $11.1 \text{ ug/m}^3$  (standard deviation =  $6.6 \text{ ug/m}^3$ ) during the same time period.

selection with respect to pollution *on the day of the surgery* is low. To verify this claim, we use our primary regression specification to conduct “balance tests,” using patients’ *pre-surgery* characteristics as outcome variables. In contrast to selection, we find no evidence of any significant differences in patients’ observable characteristics across surgeries on high versus low pollution days, which include 14 indicators for demographics, health, and surgery characteristics. To further alleviate concerns about endogeneity and measurement errors in pollution, we implement an instrumental variable (IV) approach that exploits variation in local  $PM_{2.5}$  attributable to pollution transported from upwind cities that locate over 100 kilometers away from Guangzhou. Our IV estimates are similar in magnitude with the ordinary least square (OLS) estimates.

Another potential concern on the interpretation of our results is whether the estimated effect reflects a “special” effect of surgery-day pollution, or just the fact that air pollution is in general harmful for survival. We find that higher post-surgery mortality is particularly associated with high pollution that coincides with the day of the surgery itself – as opposed to pollution exposure that occurs prior to hospital admission, between the admission and the surgery, or after the date of the procedure. In a more generalized exercise, we compare observed pollution effects with the distribution of “placebo” effect estimates that would be obtained by randomly assigning surgery dates to patients. This randomized inference again provides strong evidence that high pollution levels on the day of surgery are particularly bad for post-surgery survival. This is, in fact, a key result of our paper: days with high outdoor  $PM_{2.5}$  concentrations are linked to a substantial reduction of survival among patients who undergo surgeries on that day, while similar  $PM_{2.5}$  changes on other days have no similar effect on the same patients. This result is the basis of the second part of this paper, which considers alternative surgery scheduling to reduce surgery-day pollution’s adverse impact.

How does surgery-day pollution affect patient survival? We examine channels that fall into two groups of explanations: the effects of air pollution on physicians and on patients. We find limited evidence to support a physician effect. We begin with institutional evidence. Our review of both architectural codes of surgery operating room construction and available air quality inspection studies suggests most operating rooms in Guangzhou are equipped with high-efficiency air-filtering technologies that can remove vast majority of particulate pollution. The immediate cognitive stress on a surgeon or other medical staff due to pollution exposure while performing the operation is likely very low. We also look for changes in other performance indicators that may reveal treatment differences, such as antimicrobial agents use, the incidence of non-healing surgical wounds, or medical error-related post-admission injuries. We find no evidence of systematic differences in these indicators of treatment decisions or performance on days with different levels of pollution. Moreover, we find that the mortality risk of high pollution levels on the day

of is concentrated among a key group of elderly patients with respiratory and cancer diagnoses, which plausibly contrasts with a physician channel whereby one would expect to observe risks emerging across a broader patient population. Due to privacy restrictions, we do not have access to identifying information on surgeons and other medical personnel, and therefore unable to estimate models with physician fixed effects, or to assess the role of the physician's treatment style or experience ([Gong, 2018](#); [Molitor, 2018](#)). We leave these directions for future work to analyze.

On the other hand, we find evidence to support the view that the impacts stem from the effects of pollution on the patient. As we noted, hospital's indoor environment provides limited protection against air pollutants, such as fine particulate matter, which can easily penetrate the air inside buildings. In such situations, air pollution may pose a significant risk to the patients, even if the surgery per se is associated with very low exposure. Inspired by the prior medical research on the effect of operating room air quality and surgical complications ([Gryska and O'Dea, 1970](#); [Charnley, 1972](#); [Lidwell et al., 1982](#)) – a literature that underlies the use of cleanroom technologies in surgery theaters – we examine the effect of surgery-day air quality on markers of surgical site infections, finding an increase of infection incidents as outdoor pollution level rises. This result may arise as the typical surgical wound takes 24 to 48 hours to close after the operation ([Mangram et al., 1999](#)), leaving the patient susceptible to exogenous infections from in-hospital exposure to pollution.

We also recover a series of empirical patterns that have commonly emerged in prior studies on air pollution's effects on general population health. First, patients' vulnerability to pollution on the day of surgery varies substantially across subgroups. In our case, over 60 percent of the observed effects are explained by a small, high-risk group (6 percent of all patients) composed patients who are over age 60 and are being treated for lung diseases or cancers. Second, the relationship between post-surgery mortality and pollution features a concentration response, with damage rising sharply for initial increases in pollution levels, and then flattening out at very high levels (above 40  $\mu\text{g}/\text{m}^3$ ). Finally,  $\text{PM}_{2.5}$  concentrations tend to be the most robust predictor for patient survival, as indicated by a multiple-pollutant model that also considers the simultaneous impacts of other harmful air pollutants including ozone, nitrogen dioxides, sulfur dioxides, and carbon monoxides on the surgery day. Together, these familiar characteristics suggest that the pollution effects we observe in our study likely share common toxicological and physiological underpinnings that have been evidenced in prior studies ([Jerrett et al., 2009](#); [Brook et al., 2010](#); [Pope III et al., 2011](#); [Crouse et al., 2012](#); [Pope III, Cropper, Coggins, and Cohen, 2015](#); [Deryugina et al., 2019](#)) – although the day-of-surgery effect we document might stem from the particular vulnerability of patients at that time, leaving them exposed to greater risk of pollution's effects than otherwise would have been the case.

**The second part** of our paper presents a stylized structural analysis that builds on the observed link between pollution and post-surgery survival. Our goal is to examine whether hospitals can better internalize the adverse effect of pollution to improve patient survival through relatively targeted, small changes in the status quo scheduling practices.

Surgery scheduling is a complex undertaking. Large shifts in scheduling practices may induce unintended consequences that are not captured by our stylized structural model. For example, rescheduling many surgeries to days with good air quality may add to capacity stress on those days, strain healthcare resources, and cause adverse outcomes. We exploit two empirical features that allow us to consider relatively modest changes in surgery schedules. First, our model considers rescheduling only for the aforementioned high-risk patients (respiratory and cancer patients aged over 60) – that is, the 6 percent of all patients who account for 60 percent of the negative health effects. We verify that this restriction ensures that the impact of rescheduling on surgical capacity is very small compared to the hospital’s overall capacity constraints. Second, we limit our counterfactual scenarios to include only those patients who were originally intended to undergo surgery within three days of their hospital admission (41 percent of all surgical cases), and consider alternative days for surgery that are within this three-day window. In practice, this short-term approach means pollution expectation is more accurate; it also circumvents practical problems such as patient consent and any health consequences that might surface with longer delays. For example, we find supportive evidence that patient’s admission-to-surgery delay within the three-day time window has statistically insignificant effect on post-surgery mortality, suggesting the health damage associated with changes in surgery delays within the narrow time window might be limited. As we show in the paper, there is abundant day-to-day variations in air pollution levels to allow for a meaningful rescheduling exercise even within such a narrow time window.

We first build and parameterize a stylized model of hospital surgery scheduling among the small group of high-risk patients, and estimate key parameters that represent hospitals’ implicit trade-offs between patient mortality hazard and other non-health-related considerations. We then consider counterfactual scenarios in which the hospital would take into account the adverse consequences of surgery-day  $PM_{2.5}$  exposure, and observe how hospitals would re-optimize surgery schedules as a result. The structural exercise reveals important opportunities for patient survival improvements. For about 40 percent of the surgery cases among our targeted patient group, there exists a better air quality day within the three-day window such that switching the surgery to that day would likely improve post-surgery survival. The average “switcher” patient experiences a 4.2 percent improvement in mortality risk relative to the mean (a risk reduction of 1.6 deaths per 1,000 patients). This is likely a significant mortality improvement, as the estimated effect size is of same order of magnitude with those from well-known

surgical quality improvement programs such as the WHO’s Safe Surgery Saves Lives Program and the Veteran Affairs’ National Surgical Quality Improvement Program. We expect little change in health hazards among the rest of the patients whose surgeries are not rescheduled. This is because the counterfactual surgery schedule differs little from the original surgery schedule in terms of *overall* surgery capacity utilization, and is well below capacity constraints of the facility. Besides the main structural exercise, we also provide an example surgery postponing rule in which the hospital preserves the status quo scheduling practice while pushing high-risk surgeries back if pollution on the surgery day turns out to be exceptionally high. This practice does not reschedule surgeries to *earlier* days (a practice that is allowed in the full structural exercise but could be costly to implement in practice) while still able to achieve important survival improvements.

The disease burden of ambient air pollution has been widely studied (e.g., [Dominici, Greenstone, and Sunstein, 2014](#); [Cohen et al., 2017](#); [Landrigan et al., 2018](#)), with a particular focus on the effect of outdoor air pollution on health of the general population.<sup>2</sup> However, few studies have considered the relevance of pollution in the context of healthcare delivery. In this paper, we show that air pollution is a direct determinant of post-surgery survival, one of the most important indicators of the quality of surgical care. Beyond documenting the health effect, we show that the pre-scheduled nature of pollution exposure in the surgical contexts allows for avoidance of unnecessary exposure among the most vulnerable. This is a natural extension of the idea that adaptation can occur through avoidance ([Cutter and Neidell, 2009](#); [Neidell, 2009](#); [Graff Zivin and Neidell, 2009](#)), and here we provide a concrete illustration on how better scheduling practices may provide such adaptation in an institutional context. The analysis framework of our paper may be applied to other economic scheduling problems where the outcomes depend on a predictable future hazard. Finally, our study adds to a nascent literature on the importance of indoor air pollution exposure ([Duflo, Hanna, and Greenstone, 2008](#); [Jeuland, Pattanayak, and Bluffstone, 2015](#); [Stafford, 2015](#); [Barron and Torero, 2017](#); [Greenstone, Lee, and Sahai, 2021](#)).

This paper also connects to several strands of literature beyond economics. A key contribution of our analysis is the operation-level relevance of the findings: the fact that it is feasible to mitigate the adverse effect of pollution exposure through better scheduling of the surgeries. In this respect, we add to the epidemiology literature on the incidents and risk factors of perioperative mortality ([Ghaferi, Birkmeyer, and Dimick, 2009](#); [Bainbridge et al., 2012](#); [Pearse et al., 2012](#)), and to the operation research literature on

---

<sup>2</sup> In particular, this paper is related to the economics literature on the mortality burden of PM<sub>2.5</sub> exposure ([Chay, Dobkin, and Greenstone, 2003](#); [Currie and Neidell, 2005](#); [Moretti and Neidell, 2011](#); [Knittel, Miller, and Sanders, 2016](#); [Ebenstein et al., 2017](#); [Deschenes, Greenstone, and Shapiro, 2017](#); [Deryugina et al., 2019](#); [Anderson, 2020](#)).



surgery theater management ([Cardoen, Demeulemeester, and Beliën, 2010](#)). To the best of our knowledge, neither field has previously considered the relevance of air quality.

Section 2 continues with a brief survey of institutional background on Guangzhou’s healthcare system and in-hospital air pollution exposure; it also introduces the primary data sources. Section 3 presents the reduced-form analysis of the effect of pollution on patients’ post-operative survival. Section 4 describes the structural model of surgery scheduling, and discusses possible counterfactuals. Section 5 concludes.

## 2. Background and Data

### 2.1 Background

Guangzhou, the focal city of our study, is a prefecture-level city located near the southern coast of China (Appendix Figure A.1, panel A). Guangzhou is one of the four most-developed, “Tier-1” cities in China (along with Beijing, Shanghai, and Shenzhen). In 2017, the city has a population of 14.5 million, a total GDP of 318.5 billion USD (ranked fourth among all cities), and a per capita GDP of 22,317 USD (ranked seventh). Widely recognized as the healthcare center in southern China, Guangzhou has some of the best healthcare resources in the country. For every 1,000 residents, Guangzhou has 2.8 physicians (national average: 2.0), 4.6 nurses (national average: 2.7), and 4.6 hospital beds (national average: 4.3) according to data from 2017. Annual per capita healthcare spending in 2018 was 1,541 PPP USD (national average: 838 PPP USD).<sup>3</sup>

**Air Quality in the Operating Room.** Are patients exposed to particulate pollution while undergoing surgery? An examination of the potential for pollution exposure inside the operating room itself warrants an understanding of China’s operating room standards and compliance in the city of Guangzhou. The construction of hospital operation theaters follows a set of protocols in the Architectural Technical Code for Hospital Clean Operating Department (2002).<sup>4</sup> Based on their ability to maintain a nearly particulate-free environment, operating rooms are classified into one of four main categories. For example, Class I (cleanest) operating rooms are those that meet International Organization for Standardization (ISO) Class 5 cleanliness standards,<sup>5</sup> which correspond to a maximum particle concentration of 3,500 particles

---

<sup>3</sup> For reference, in the United States in 2017, the corresponding national average figures were 2.6 physicians, 0.6 nurses, and 2.9 beds for every 1,000 residents. Annual per capita healthcare spending in that year was 10,210 USD. The statistics from China are taken from the National Health Yearbook and the Guangdong Province Health Yearbook. The raw statistics are scaled to reflect accounting definitions used by the World Health Organization’s Global Health Observatory and the Global Health Expenditure Database.

<sup>4</sup> GB 50333-2002 and GB 50333-2013.

<sup>5</sup> ISO 14644 Part 1: Classification of Air Cleanliness.

per cubic meter of air for particles over 0.5 micrometers; Class II (second cleanest) operating rooms are defined by ISO Class 6 cleanliness (concentration  $\leq 35,000$  particles/m<sup>3</sup> for particles over 0.5 micrometers), and so forth. These high levels of cleanliness are often achieved through the use of High Efficiency Particulate Air (HEPA) or Ultra Low Particulate Air (ULPA) filters combined with a laminar (unidirectional) air flow system. In practice, this means that Class I, II, III and IV operating rooms filter out, respectively, over 99.99%, 99%, 95%, and 70% of particles over 0.5 micrometers. Filtration rates for larger particles, such as PM<sub>2.5</sub> (particles over 2.5 micrometers) are expected to be even higher.

While the filtering capacity of operating rooms must comply with these standards, actual performance also depends on appropriate practices, such as timely cleaning and replacement of air filters. The best data we are aware of come from the Guangzhou Center for Disease Control and Prevention, which inspected 53 operating rooms across 18 hospitals in 2014 ([Li et al., 2014](#)). The study found that over 83% of inspected operating rooms satisfy the particulate cleanliness standards.<sup>6</sup> Overall, we believe particulate pollution exposure inside the operating room is likely very low.

**Air Quality in Other Hospital Areas.** Outside of the operating rooms and other cleanrooms with special purposes (such as the Intensive Care Units), air pollution control within the hospitals is limited. For small particulate pollution such as PM<sub>2.5</sub>, outdoor-to-indoor penetration is likely a major source of indoor exposure (e.g., [Hanley et al., 1994](#); [Riley et al., 2002](#); [Chen and Zhao, 2011](#)). Data on indoor air pollution in hospital settings are scarce. One useful case study monitored indoor and outdoor air pollution in two hospital rooms in Erfurt, Germany, finding an average indoor/outdoor PM<sub>2.5</sub> ratio of 0.83 (open-window) and 0.63 (closed-window); outdoor PM<sub>2.5</sub> also strongly predicts indoor PM<sub>2.5</sub> *variation*, with a linear regression R-squared of 0.84 ([Cyrus et al., 2004](#)). In examining seven peer-reviewed studies conducted in Chinese hospitals between 2007 and 2014, [Zheng \(2014\)](#) finds that indoor/outdoor PM<sub>2.5</sub> ratios are generally near one. The health benefits of ventilation and air quality control in hospital wards and patient rooms remain an important but underexplored area in both research and practice.

## 2.2 Data and Summary Statistics

**Surgery Records.** Our surgery data are based on the universe of hospitalization records in Guangzhou from 2014 to 2017. These data are sourced from administrative medical records submitted by individual hospitals to the Guangzhou Health Information Center. Access to the data was granted by the

---

<sup>6</sup> A critical finding of [Li et al. \(2014\)](#), however, is that some hospitals tend to achieve high particle-filtration rates by overusing air changes, causing less-than-optimal humidity conditions in the operating room.

Guangzhou Municipal Health Commission. Because we are among the first researchers to use this type of inpatient database in China, below we discuss some relevant institutional details of the data source.

The availability of the healthcare data is linked with the ongoing health informatics reform in China in the wake of the establishment of the Health Level Seven (HL7) China committee in 2006.<sup>7</sup> As part of this agenda, legislation in 2014 created a national electronic health records system, the Basic Dataset of Electronic Medical Records. The Basic Dataset covers various aspects of healthcare. Our data are extracted from Part 10 of the Basic Dataset, known as the Home Page of Inpatient Medical Records.<sup>8</sup> These data provide an “abstract” of each inpatient record, containing the most critical information associated with the hospital stay including patient demographics, hospital admission, inpatient care (including surgery care), and payment. The Home Page information is primarily filled out by the attending physician(s), and then verified by a clinical coder to ensure consistency in the diagnosis and procedure codes. Participation in this digital reporting system is mandatory. Each hospital’s compliance is reviewed and graded annually by the National Health Commission. Digital medical records are also the basis of public and private insurance reimbursements, creating a strong incentive for the hospitals to comply with reporting.

Our research data file is an extract of the Home Page data. Each observation in the database corresponds to a unique hospital stay, allowing us to observe (scrambled) patient identity, basic demographics, surgery information, and insurance payment information. The layout of the data resembles the typical inpatient records dataset available in the U.S. setting, such as the Healthcare Cost and Utilization Project (HCUP) State Inpatient Database. Below we discuss several points on variable construction that are important for our empirical analysis.

*Admissions, discharges, and patient deaths.* In the database, admission time is recorded as the date the patient entered the ward and began receiving treatment. Discharge time is recorded as the time when the treatment was terminated, and the patient exited the ward. For patients who died during the hospitalization, discharge time is recorded as the date of death, with the manner of discharge flagged as “death.” Our primary hospitalization mortality outcome variable is constructed from the discharge time and manner-of-discharge fields. Our data do not contain any cause-of-death information.<sup>9</sup>

*Diagnosis and procedure codes.* Disease diagnoses are coded in the International Classification of Diseases 10<sup>th</sup> revision (ICD-10). Operations are coded in the ICD-9-CM Volume 3 codes. These are

---

<sup>7</sup> HL7 is one of the American National Standards Institute (ANSI)-accredited international standards for the transfer of clinical and administrative data between healthcare providers (<https://www.hl7.org/implement/standards/index.cfm>).

<sup>8</sup> WS 445.10-2014 Basic Dataset of Electronic Medical Record—Part 10: Home Page of Inpatient Medical Records.

<sup>9</sup> Death reporting is administered under the Death Certificate Program, which is a separate program that is not directly linked to inpatient records.

standard coding practices used, for example, by the U.S. Agency for Healthcare Research and Quality (AHRQ) in maintaining state inpatient databases and the Centers for Medicare & Medicaid Services (CMS) in medical insurance claims data. We use these ICD codes to construct diagnosis and procedure fixed effects control variables in our main regressions.

*Multiple surgeries.* A patient may undergo multiple surgeries during a hospital stay. In these cases, we focus on the “primary” surgery defined by the Home Page data protocol as the operation most directly related to the primary diagnosis. The primary surgery is often the most technically challenging and risky operation. This focus implies that each observation in our estimation sample corresponds to a unique hospital stay.

*Medical personnel information.* Our data use agreement precludes us from accessing any hospital personnel information. Thus, we do not observe the attending physicians, nurses, or any medical assistants involved at the admission or with the surgery. This is a shortcoming that limits our ability to test a physician-side response and mechanism. We do attempt to examine a potential pollution-induced medical error channel by inferring potential medical errors from injury-related patient discharge codes that were not present at admission ([Van Den Bos et al., 2011](#); [David et al., 2013](#)). These exercises are reported in Section 3.4.

Our dataset consists of all hospitalization records submitted by 23 “3-A” hospitals (N = 2,233,969). These 3-A hospitals are the major healthcare facilities in the city, and they have the best-quality data. To improve statistical power, we restrict to surgery categories with an average death rate of at least 1 per 1,000 patients (remaining N = 1,381,283).<sup>10</sup> Appendix Table A.1 reports that our main mortality effect estimation coefficients are smaller, yet still statistically significant, when including these low-risk procedures. The average post-surgery death rate in our final estimation sample is 12.1 deaths per 1,000 patients; this rate represents 44.4% of all in-hospital deaths and 28.5% of all deaths in the city of Guangzhou during the study period. Table 1 contains more summary statistics of our estimation sample.

**Air Pollution Monitoring Data.** Guangzhou began real-time broadcast of air quality in 2013, following China’s ambient air pollution disclosure reform ([Greenstone et al., 2020](#); [Barwick et al., 2020](#)). We obtain daily air pollution concentration data between 2014 and 2017 from all eight monitoring sites in Guangzhou. Appendix Figure A.1 panel B plots monitoring sites and hospital locations. Note that data coverage is sufficient; most hospitals have monitors within several miles of the facility. In practice, we

---

<sup>10</sup> The excluded categories are operations on the endocrine system (ICD-9-CM V3 code block: 06-07), the eye (08-16), the ear (18-20), the nose, mouth and pharynx (21-29), the male genital organs (60-64), the female genital organs (6571), and obstetrical procedures (72-75).

assign each hospital the pollution readings from the nearest monitor. Appendix Figure A.2 panel A shows the distribution of daily  $PM_{2.5}$  values in Guangzhou (mean =  $36.5 \text{ ug/m}^3$ , s.d. =  $19.8 \text{ ug/m}^3$ ).

One common concern with government-provided air quality data in China is about political tampering due to city official's promotion motives ([Andrews, 2008](#); [Chen, Jin, Kumar, and Shi, 2012](#); [Ghanem and Zhang, 2014](#); [Ghanem, Shen, and Zhang, 2020](#)). Accuracy of air quality data since the 2013 reform has improved substantially as sampling is now automated and reported in real time ([Greenstone et al., 2020](#)). In Appendix Figure A.2 panel B, we compare the government  $PM_{2.5}$  data with  $PM_{2.5}$  data collected independently by the U.S. Consulate in Guangzhou from 2012 to 2017. The figure shows the two time series are consistent with each other.

Section 3 uses air pollution *monitoring* data to study the link between air quality and post-surgery survival. Section 4 studies surgery scheduling based on *forecasted* air pollution. We discuss issues related to the accuracy of air pollution forecasts in Section 4.4.

### 3. Evidence

#### 3.1 Raw Data Pattern: Surgery-Day Pollution and Patient Survival

We first investigate patient survival patterns in a standard Kaplan-Meier framework, considering high pollution levels on the surgery day as a “treatment.” In Figure 1, we compare the 30-day survival functions of patients in two groups: those whose surgeries took place on highly polluted days, which we define as those days when  $PM_{2.5}$  levels were within the highest 20 percentile of measurements recorded in our sample (i.e., days with  $PM_{2.5} > 50 \text{ ug/m}^3$ ;  $N=298,292$ ); and those received surgeries on best 20 percent pollution days (i.e., days with  $PM_{2.5} < 20 \text{ ug/m}^3$ ;  $N=264,120$ ). The raw data pattern shows that those who underwent surgery on high pollution days exhibit lower survival probability, and the effect appear to manifest gradually throughout the month following the operation. By day 30, the observed gap in survival in the high and low pollution groups is 0.151 percentage points, which is about a 15 percent difference in mortality. The patterns are robust to alternative definitions of “high” and “low” air pollution days, such as above- and below- median  $PM_{2.5}$  values. Appendix Figure A.3 reports these robustness checks.

Figure 1 suggests a clear difference in survival trajectories for patients whose surgeries took place on days with high versus low pollution. How much of this difference is due to pollution exposure *alone*, as opposed to other factors that are correlated with pollution? Are patients scheduled to undergo surgeries on highly polluted days fundamentally different (e.g., in terms of baseline health) than patients who were scheduled to undergo surgeries on days with low pollution levels? Do physicians treat patients differently

on days with high or low levels of pollution? In the next subsection, we describe a regression approach that allows us to investigate these questions.

### 3.2 Regression Framework

Our regression model links a patient’s post-surgery survival to the ambient air pollution concentration on the day of surgery. The workhorse regression equation is

$$Y_i = \alpha + \beta \cdot \text{Pollution}_i + X_i\gamma + \varepsilon_i \quad (1)$$

$Y_i$  is the outcome of interest for patient  $i$ . In our primary mortality analysis,  $Y_i$  is an indicator variable for whether the patient died during the hospitalization following the surgery (“hospital mortality”). The regressor of interest is  $\text{Pollution}_i$ , which is the logged ambient PM<sub>2.5</sub> concentration recorded at the air monitoring site closest to patient  $i$ ’s hospital on the day of surgery.<sup>11</sup>  $X_i$  denotes a series of control variables that may potentially correlate with both pollution and mortality. These controls are of three types: (1) patient-level controls that include age in five-year bins, a gender indicator, a marital status indicator, and an indicator for any history of allergies, (2) weather controls that include daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared, and (3) fixed effects control for primary diagnosis, hospital, department, procedure, year (of surgery), month of year, and day of week.  $\varepsilon_i$  is the error term. To account for serial correlation both within hospital and across different hospitals on the same day, we two-way cluster standard errors at the hospital and the date levels.<sup>12</sup>

Though simple, this regression framework allows us to test a variety of hypotheses on whether and how pollution severity affects post-surgery survival. Here we briefly describe four groups of questions we can assess through variants of equation (1). We provide more details throughout sections 3.3 and 3.4 when we present the corresponding results.

---

<sup>11</sup> Appendix Figure A.1 panel B shows that most hospitals in our study sample have air pollution monitoring sites within two miles.

<sup>12</sup> Note equation (1) is essentially a cross-sectional regression in which an observation is a surgery, and the primary outcome of interest is post-surgery death. But because surgeries occur on different dates, we can use time fixed effects to exploit different sources of air pollution variation. An equivalent formulation of equation (1) is a panel estimation model

$$Y_{jkt} = \alpha + \beta \cdot \text{Pollution}_{jkt} + \alpha_{jk} + \alpha_t + \varepsilon_{jkt}$$

in which the regression links the fraction of patients died in hospital  $k$  following a procedure  $j$  done on date  $t$  ( $Y_{jkt}$ ) to the surgery-day pollution ( $\text{Pollution}_{jkt}$ ), controlling for procedure-by-hospital fixed effects ( $\alpha_{jk}$ ) and time fixed effects ( $\alpha_t$ ). Equation (1) is essentially a micro-data version of this panel formulation.

**Are Patients Scheduled on High-Pollution Days Fundamentally Different?** One might worry about two potential types of patient selection. The first is about patient selection at admission: sicker patients are more likely to be admitted to the hospital on high-pollution days, thus driving up the mortality rate in some ways. Notice, however, that our analysis is conditioned on patients who have *already* been admitted to the hospital; these patients have spent, on average, four days in hospital between the admission and the surgery. Because our treatment of interest is pollution on the (pre-scheduled) day of surgery, rather than the day of admission, the scope for endogenous selection is likely low. In Appendix Table A.1, we report that the estimation results are similar, although less precise, when excluding patients who received surgery on the same day they were admitted, or in the two days following the day of admission. In addition, Section 3.3 shows that, for the average patient, pollution levels around hospital admission day does not predict post-surgery survival. Second, higher inpatient volume may add staffing stress and lead to worse outcomes for patients scheduled to undergo surgery on that day ([Aiken et al., 2002](#); [Needleman et al., 2011](#); [Ball et al., 2018](#)). In Appendix Table A.1, we show that the findings are robust to controlling for hospital or hospital-department patient volume (or surgery volume) in equation (1).

A second type of concern is whether sicker patients are more likely to be *scheduled* to undergo surgery on high pollution day. If patients scheduled to undergo surgery on highly polluted days are fundamentally different than patients who undergo surgeries on low-pollution days (e.g., due to scheduling practices unbeknownst to the researchers), then  $\beta$  may capture a “compositional difference” in patients, rather than the effect of pollution. We address this concern by using patient’s *pre-surgery* characteristics – such as basic demographic characteristics, health conditions, socioeconomic status – as outcome variable  $Y_i$  in equation (1). In the spirit of a balance test in randomized experiment, this exercise tells us if there is evidence of selection into high and low pollution day surgeries based on patients’ observable characteristics. In addition to the balance test on observable characteristics, we have also tackled selection on unobservables using time-invariant fixed effects controls (for hospital, procedure, etc.) and temporal fixed effects controls (month, day of week, etc.). The evidence suggests that the degree of patient selection is low.

**Does  $\beta$  Capture the Causal Effect of Pollution?** A more general endogeneity concern is that some variation in air pollution may be correlated with factors that independently affect post-surgery survival, leading to the omitted variable bias. We use a fixed effects approach to tease out plausibly exogenous, day-to-day pollution variation. Our baseline specification includes surgery-year, surgery-month-of-year, and surgery-day-of-week fixed effects to parse out annual, seasonal, and within-week factors that may confound the pollution-survival relationship. Besides the time fixed effects, all regressions

control for hospital, department, diagnosis, and procedure (or procedure-by-hospital) fixed effects, as well as for patient-level characteristics to account for time-invariant determinants of post-surgery survival.

We also assess the robustness of results using alternative sources of pollution variation. For example, we present results with month-of-sample (i.e., year-by-month-of-year) fixed effects; these exploit variation in pollution across different days of the same month. In yet another specification, we include hospital-by-procedure-by-month-of-year fixed effects; these exploit variation among patients who receive the same procedure in the same hospital during the same time of year, but across different years when pollution levels are different. In Appendix Table A.1, we further report that our effect estimates are similar with or without the weather covariates (daily temperature and precipitation), which suggests a limited role played by atmospheric conditions.

To further alleviate concerns about potential endogeneity and measurement errors in pollution, we implement a wind-transport instrumental variable (IV) approach in the spirit of [Barwick et al. \(2018\)](#), [Deryugina et al. \(2019\)](#), and [Anderson \(2020\)](#). The goal is to tease out variation in Guangzhou’s air pollution attributable to transported pollutants from upwind cities. We begin with a daily panel dataset of PM<sub>2.5</sub> in all 305 prefecture-level cities that are at least 100 kilometers away from Guangzhou. For each city  $c$  and day  $t$ , we calculate the (radian) angle  $\phi_{ct}$  between city  $c$ ’s local wind direction and the vector pointing from city  $c$  to Guangzhou (e.g.,  $\phi_{ct} = 0$  if city  $c$  is exactly upwind from Guangzhou on day  $t$ ). The IV is a time-series variable constructed using the following formula:

$$IV_t = (1/305) \sum_{c \in \{1, \dots, 305\}} \max\{0, \cos(\phi_{ct})\} \cdot PM2.5_{ct} \cdot \left( \frac{1/\text{distance}_c}{1/\sum_i (1/\text{distance}_c)} \right) \quad (2)$$

where the  $\max\{0, \cos(\phi_{ct})\} \cdot PM2.5_{ct}$  term – which we call “upwind pollution” – is the vector component of air pollution in city  $c$  on day  $t$  that is expected to move toward Guangzhou. We assume upwind pollution is zero if  $\phi_{ct}$  is an obtuse angle, i.e., winds in city  $c$  on day  $t$  are blowing *away* from the direction toward Guangzhou. On any date  $t$ , the IV is the average of individual cities upwind pollution terms, inversely weighted by city  $c$ ’s distance to Guangzhou ( $\text{distance}_c$ ). We then estimate a two-stage least squares (2SLS) version of equation (1) where we instrument for patient  $i$ ’s surgery-day pollution ( $\text{Pollution}_i$ ) using the corresponding  $IV_t$  value.

The identifying assumptions of the IV strategy are (a) *first stage*: upwind pollution is a strong predictor of air pollution variation in Guangzhou, and (b) *exclusion restriction*: except for its influence on local air quality, transported pollution from distant cities does not otherwise affect patients’ post-surgery survival in Guangzhou. In a set of robustness checks, we explore alternative IV constructions that vary the balance between (a) and (b): exploiting pollution variation in very far-away cities helps the exclusion



restriction, but it necessarily weakens the predictive power of the first stage. We have experimented with excluding cities over 1,000 km away, using inverse-distance-squared weighting, or employing a data-driven method that selects the most predictive upwind cities in a “zero-stage” Lasso regression.<sup>13</sup> Our IV estimation yield similar findings with the OLS results. This further supports the view that endogeneity and measurement error concerns are limited in our study context. For the rest of the analysis, we use OLS outlined in equation (1) as our preferred estimation strategy for the sake of efficiency.

Overall, our effect estimates are robust to a variety of specification checks; many augmented models we estimate (many discussed in this subsection) appear to be well behaved in relation to prior studies. We therefore interpret  $\beta$  as capturing the effect of exogeneous pollution changes.

**Does  $\beta$  Capture the General Effect of Pollution or the Special Effect of Surgery-Day Pollution?** A deeper concern is to what degree  $\beta$  captures the *special* effect of high pollution on the day of surgery, as opposed to the adverse effects that high pollution may *generally* have on survival. We approach this concern by altering the time window during which  $\text{Pollution}_i$  is defined. In one exercise, we replace surgery-day pollution with pollution levels prior to hospital admission, between admission and surgery, or after the surgery. We then estimate the effects of pollution that occur in these alternative time frames. In Appendix Table A.2, we also report a distributed leads and lags model which shows similar evidence that pollution on the surgery day, rather than its leads and lags, leads to higher post-surgery mortality.

In a second, more generalized exercise, we *randomly* assign surgery dates to patients, and estimate the effect of pollution under the “placebo” scenario. Repeating this process many times yields a distribution of placebo effect sizes under the null hypothesis that “surgery-day pollution is just as bad as pollution on any other day.” Comparing the true effect size with these alternative scenarios, we conclude that high levels of pollution on the day of surgery is particularly bad for post-surgery survival.

**Why Does Surgery-Day Pollution Affect Patient Survival?** In the absence of patient selection on days with different levels of pollution, what are some of the other channels through which surgery-day pollution influences post-surgery survival? We first test a hypothesis that pollution may affect how physicians treat patients. Such an effect may have numerous causes, such as the possibility that doctors

---

<sup>13</sup> Specifically, before constructing the IV variable, we estimate the following estimation equation with linear Lasso

$$\text{PM2.5}_{\text{Guangzhou},t} = \lambda_0 + \sum_{c \in \{1, \dots, 305\}} \lambda_c \cdot \max\{0, \cos(\phi_{ct})\} \cdot \text{PM2.5}_{ct} + \epsilon_t$$

which selects a subset of 119 upwind cities. In Appendix Figure A.4, we map out the location of the selected cities and report the  $\hat{\lambda}_c$  coefficients from the post-Lasso regression. We then conduct the IV construction outlined in equation (2) using these 119 selected cities (rather than all 305 cities) as upwind cities.

face increased cognitive stress on heavily polluted days. Without physician-level information, we cannot pin down a physician effect per se, but we can examine related procedures and outcomes that can serve as proxies to measure behavioral changes that could be linked to pollution levels. With this in mind, we examine changes in procedures and outcomes that have been used elsewhere in the literature: the use of antimicrobial agents (which are commonly used to prevent and combat wound infection), incidences of non-healing surgical wounds, and a measurement of medical error based on post-hospitalization injuries (Van Den Bos et al., 2011; David et al., 2013). In each case, we find limited evidence that pollution affects these treatment and performance indicators.

Turning to the patient side, we next estimate a series of augmented models that are commonly presented in prior research of pollution in epidemiology and economics. We estimate a concentration-response function model, a multiple pollutants model, and a heterogeneous treatment effect model with respect to subgroup vulnerability. Overall, we find results that are qualitatively similar to prior studies, suggesting that the pollution effects we observe likely share common toxicological and physiological mechanisms that have been evidenced in prior studies, while the surgeries might have left the patients exposed to greater risk of pollution's effects.

**Measurement Errors, “True” Exposure Effects, and Observed Effects.** Before proceeding, we discuss two sources of measurement errors that pertain specifically to our setting, and we offer several thoughts on how they might affect the interpretation of our estimates.

One view is that the “ideal” effect estimate should reflect the true *exposure effect* of pollution on surgery patients’ subsequent death – that is, the effect of PM<sub>2.5</sub> exposure inside the hospital on patient deaths that occur either during or after hospitalization. Our estimation departs from this ideal in two ways. First, the pollution measure we use is *outdoor* PM<sub>2.5</sub>. Outdoor PM<sub>2.5</sub> is likely higher than indoor levels, which may bias the effect estimate downward relative to the true exposure effect of PM<sub>2.5</sub>. That said, our review of prior studies on hospital air quality (Section 2.1) suggests that indoor air pollution correlates strongly with outdoor levels, and that the indoor-to-outdoor ratio is near one for some Chinese hospitals, suggesting the magnitude of this measurement error may not be substantial.

Second, our effect estimate only reflects the impact of surgery-day pollution up to the point of hospital discharge. The effect of this measurement error on our estimates is *ex ante* ambiguous: on one hand, we do not observe deaths after patients leave the hospital, which may lead us to understate the total effect of pollution; on the other hand, sicker patients may stay longer in hospital so that the pool of patients that remain in our data have higher average mortality rates compared to those already discharged; this may lead us to overstate the true effect of pollution. One piece of empirical pattern we see, however, is that the

curvature of the survival curve flattens out over time, and so does the adverse effect of pollution (Figure 1 and Table 2). This pattern does not support the view that the health composition of longer-staying patients is substantially worse.

While our estimate departs from the “ideal”, it puts forward a relevant estimate. Most hospitals have information available (through PM<sub>2.5</sub> readings and forecasts) about outdoor, rather than indoor, air quality levels. At the same time, hospitals likely focus on patient’s outcome to the greatest extent during the period in which patients are in their care. Thus, our estimated  $\beta$  is a more relevant piece of knowledge than the “ideal” estimate. Empirically, we find the IV estimates are larger but of the same order of magnitude compared to the OLS estimates, suggesting the magnitude of the measurement errors is unlikely to be huge. We provide more discussion on these points in Section 3.3 and Section 4.

### 3.3 Main Results

**Balance Test of Patient Characteristics.** In the spirit of a balance test in randomized experiment, we first use equation (1) to test whether surgery-day pollution ( $Pollution_i$ ) can predict patients’ *pre-surgery* characteristics. Table 1 reports the results. We test a wide variety of characteristics ranging from basic demographic and health condition (age, gender, marital status, history of allergies), surgical information (delay, number of procedures, anesthesia methods, levels of operations) and payment (whether the stay is reimbursed through the City Workers Health Insurance Program, the New Rural Cooperative Insurance Scheme, or out of pocket).<sup>14</sup> While this is not an exhaustive list, we believe they are broad enough to encompass a patient’s overall characteristics.

Results in Table 1 show that patients are well-balanced in observable characteristics with respect to surgery-day pollution. The estimated coefficients are generally small and statistically insignificant. One exception is that patients who underwent surgery on high-pollution days appear to be statistically more likely to be married. However, the magnitude of the correlation is small: we can reject a 0.5% effect using the largest estimate from column 3, i.e.,  $(1.970 + 2 \times 0.651) / 802.8 = 0.41\%$  out of the mean level of the indicator of married patient.

These results suggest that patients who undergo surgeries on high pollution days are on average no different from those undergo surgeries on low pollution days, conditional on the control variables in equation (1). Another implication of these results is that hospitals are not systematically scheduling patients

---

<sup>14</sup> The City Workers Health Insurance Program and the New Rural Cooperative Insurance Scheme are also proxies for the patient’s urban and rural status, respectively.

according to pollution levels: older patients or patients undergoing more intensive procedures are, on average, equally likely to be assigned to high-pollution or low-pollution days, etc. This determination serves as one basis of the rescheduling structural analysis we conduct in Section 4, and it informs our position that changing existing scheduling practices to offer targeted patients surgeries on days with lower levels of pollution may benefit patients' overall survival chances.

**The Effect of Pollution on Post-Surgery Mortality.** Table 2 presents our main mortality estimation. Like Table 1, the columns are organized by the fixed effects choices; the rows are organized by mortality at different post-surgery time horizons. We first examine “1-day mortality”, defined as whether the patient died on the day of surgery. The mean of that variable is 1.322, meaning the average rate at which patients die on the day of surgery is 1.322 per 1,000 patients. While the point estimates are positive across specifications, we find no statistical evidence that pollution is associated with any significant increase in 1-day mortality, i.e., pollution does not increase the odds of deaths during or immediately following the surgery. The rest of the rows show that the effect of pollution manifests as we expand the post-surgery window to 7-day, 14-day, 28-day, and overall in-hospital deaths (“hospital mortality”). This evidence is consistent with the raw trends we presented in Section 3.1: high- and low-pollution surgery survival gap does not surface on the day of operation itself, but rather emerges over time.

The average of the hospital mortality regression coefficients (shown across the four columns in Table 2) is 0.428, which means that one log increase in surgery-day  $PM_{2.5}$  concentration increases the post-surgery hospital mortality rate by 0.428 per 1,000 patients; this is an increase of about 3.5 percent in the average hospital mortality rate.<sup>15</sup> As discussed in section 3.2, we interpret this effect size as a lower bound on the true *exposure effect* of  $PM_{2.5}$ . That said, from a practical perspective, the  $\beta$  coefficient contains the most immediately relevant knowledge for the hospital because (a) the outdoor (rather than indoor)  $PM_{2.5}$  forecast is what is available for use in planning and scheduling, and (b) minimizing mortality that occurs during hospitalization is an important goal for the hospital. This logic also underlies our structural modeling of surgery schedule in Section 4.

**Instrumental Variable Estimation.** Table 3 summarizes the IV estimation results. Upwind pollution is a strong predictor of air pollution in Guangzhou. Our first-stage Kleibergen-Paap F-statistics range from a smallest 169.7 in panel A that uses the baseline IV construction outlined in equation (2), to a largest 559.8 in panel D that uses a “zero-stage” Lasso to select the most important upwind cities. The most

---

<sup>15</sup> Given the average  $PM_{2.5}$  level of 36.5  $\mu g/m^3$ , our effect size translates to about 0.95% increase in hospital mortality per 10  $\mu g/m^3$  increase in  $PM_{2.5}$ .

conservative IV results (panel D) are over 20 percent larger in magnitude and noisier than the OLS estimates of Table 2, and the qualitative findings are consistent across the two estimation approaches.

Readers familiar with the quasi-experimental economics literature on the air pollution – mortality link may find it surprising that our IV estimates are only moderately larger than their OLS counterparts. For example, some recently published studies including [Deschenes, Greenstone, and Shapiro \(2017\)](#), [Ebenstein et al. \(2017\)](#), and [Deryugina et al. \(2019\)](#) all find that the IV estimates on the effect of ambient pollution on adult mortality is an order of magnitude larger than their OLS estimates. We believe this difference has several explanations that pertain to the special setting of our study context. First, the nature of potential endogeneity concern in our setting is different from prior studies. We consider patients already admitted to the hospital for a few days, and we exploit variation in air pollution on the (pre-scheduled) day of surgery. This means the usual endogeneity concerns, such as confoundedness by road traffic, may be muted in our setting.

Second, most hospital locations in our study sample have air pollution monitors within several miles (Appendix Figure A.1). This might have helped reduce measurement errors in pollution compared to prior studies that, due to larger geographic scope and the need to cover non-urban areas where monitoring network coverage is sparse, often have to rely on imputed pollution values from monitors tens of miles away. That we have pollution measurement from much more nearby monitors may help better capture intra-city variation in air pollution exposure ([Hoek et al., 2008](#); [Anderson 2020](#)).

Finally, we study the inpatient surgery population who likely spend the vast majority of time in hospital wards. This feature has two implications on measurement errors: (a) our pollution measurement is less likely to suffer from conventional sources of errors where people’s activity locations (and thus pollution exposure) are remarkably different across daytime and nighttime ([Ratti et al., 2006](#); [Bhaduri et al., 2007](#)); (b) as we discussed in Section 3.2, that patients staying mostly *indoor* may increase measurement errors compared to prior studies because we only observe *outdoor* pollution. However, some institutional knowledge suggests this source of error might not be enormous considering a high outdoor-to-indoor penetration of fine particulate matter pollutants (Section 2.1).

For the sake of estimation efficiency, in the rest of the analysis we will use the OLS approach in equation (1) as our preferred estimation strategy.

**The Importance of Surgery-Day Pollution.** While our analysis focuses on the effect of pollution on the day a surgical procedure takes place, literatures in both epidemiology and economics have documented the *general* effect of pollution on health outcomes such as hospital visits and mortality (e.g., [Schlenker and Walker, 2016](#); [Deryugina et al., 2019](#); [Xia, Xing, Xu, and Pan, 2020](#)). The question then

arises concerning the extent to which  $\beta$  picks up a “special” effect of pollution exposure on the surgery day.

We approach this concern by altering the time window during which pollution is defined. Figure 2 panel A presents the results. We first replace  $\text{Pollution}_i$  in equation (1) with average  $\text{PM}_{2.5}$  concentrations in the second week prior to the hospital admission. The estimate is close to zero (0.095) and statistically insignificant. We find similar conclusion for pollution during the week prior to admission (“-7 to -1 days to admission”). Together, these estimates suggest that the post-surgery survival effect is not a consequence of any lagged effects of exposure prior to hospitalization. We then examine other post-hospitalization exposure windows, such as average  $\text{PM}_{2.5}$  concentrations between the admission to the hospital and the date of the surgery, and the weeks following the surgery. We find generally positive but statistically insignificant effects. The lone exception is the positive and statistically significant “surgery day” coefficient, suggesting that surgery-day exposure is particularly detrimental to survival. In Appendix Table A.2, we further report distributed lag models in which we include current, lag, as well as lead terms pollution simultaneously in the regression specification. We show that, once condition on current (surgery-day) pollution, lagged pollution has no additional effects on patient mortality.

Another way to test whether surgery-day pollution plays a unique role in patient mortality is to implement a permutation-style inference. For each surgery patient, we pick a random date during the patient’s hospitalization as the “placebo” surgery day, and then re-estimate equation (1). We repeat this exercise 1,000 times and obtain a set of  $\beta$  estimates under each of the 1,000 “placebo” scenario. By construction, the placebo  $\beta$  estimates reflect the distribution of the mortality effect of  $\text{PM}_{2.5}$  under the null hypothesis that surgery-day exposure is just as bad as exposure on any other day during one’s hospitalization period. We can thus calculate a  $p$ -value for the observed effect (with true surgery dates) by comparing it with the placebo effect distribution. Figure 2, panel B shows that the observed effect falls outside the 95% range of the placebo estimates. This again points to a special effect of exposure to pollution on the day of the surgery itself on patient survival.

**Lagged Effect of Pollution versus the Effect of Lagged Pollution.** An alternative, perhaps more concise way to summarize our finding is that (a) the patient mortality effect is driven by the lagged effect of pollution that occurs on the surgery day (Figure 1; Table 2), and (b) there is no evidence of an effect of lagged pollution – that is, the pollution that occurs on days following the surgery – on patient mortality (Figure 2; Appendix Table A.2). Note that (a) is much expected as the vast majority of surgical patient deaths occur *after* the operation (Spence et al., 2019); it is also consistent with the prior literature that suggests transient changes in air pollution can have lagged effect on health which manifests over time (e.g., Schwartz, 2000; Deryugina et al., 2019). (b) is a key finding of our paper: days with high  $\text{PM}_{2.5}$  are linked

to a substantial reduction of survival among patients who undergo surgeries *on that day*, while  $PM_{2.5}$  changes on other days have no similar effect on the same patients. This result is the basis of the second part of this paper, which considers alternative surgery scheduling to reduce surgery-day pollution’s adverse impact.

### 3.4 Mechanisms

Why does high pollution on the surgery day reduce subsequent survival? We present additional evidence to shed light on the underlying mechanisms.

**Physician Effects.** Evidence in sections 3.2 and 3.3 suggest that surgery patients on high pollution days are unlikely to be significantly different from those who undergo operations on low pollution days. We next explore the possibility that similar patients are treated differently on days with varying pollution levels. One such possibility is that pollution exposure reduces physicians’ ability to perform cognitively demanding operation tasks.<sup>16</sup> The physicians in our study samples (3-A hospitals) are among China’s most distinguished and experienced surgeons. Recall, too, that the operating rooms in which the surgeries are performed are designed to be nearly particulate free (Section 2.1). Of course, these factors cannot not exclude the possibility that physicians’ *cumulative* pollution exposure during the day outside the operating room may have adversely impacted surgical performance.

As we mentioned in Section 2.2, we cannot observe physician identifiers and characteristics from our dataset. Instead, we exploit three pieces of information available from the surgery records to indirectly test for any effects on physicians. First, we examine the utilization of antimicrobial agents, a common practice to prevent and combat wound infection. We observe the monetary value of antimicrobial agents used related to the surgery, and we use equation (1) to test if higher surgery-day  $PM_{2.5}$  predicts more usage. Second, we test if higher  $PM_{2.5}$  levels on the date of the surgery increases the odds that patients suffer from non-healing surgical wounds, which can be related to surgical site infection ([Mathieu, Linke, and Wattel, 2006](#)). About 1.7 per 1,000 patients in our study sample experienced such non-healing surgical wounds. Finally, we examine a proxy for medical error. We follow [Van Den Bos et al. \(2011\)](#) and [David et al. \(2013\)](#) to construct an umbrella proxy for medical error based on injury and infection-related patient discharge diagnosis codes that were not present at admission.<sup>17</sup> These diagnoses do not necessarily imply

---

<sup>16</sup> For example, exposure to air pollution has been causally linked to reduced work productivity in sectors that require various levels of cognitive abilities ([Graff Zivin and Neidell, 2012](#); [Archsmith, Heyes, and Saberian, 2018](#); [Adhvaryu, Kala, and Nyshadham, 2019](#); [Chang, Graff Zivin, Gross, and Neidell, 2019](#); [He, Liu, and Salvo, 2019](#)). We are unaware of any existing evidence that pertains specifically to physicians.

<sup>17</sup> The companion dataset of this paper contains a list of ICD-10 codes we use to identify medical errors.

medical errors, but they are more likely to occur in the presence of one. We examine whether higher levels of surgery-day pollution influence this measure of “medical error.”

Table 4 reports the results. We find no evidence that these indicators for treatment style and medical errors change as a consequence of PM<sub>2.5</sub> variation on the day of surgery; the estimates can rule out small effect size. In Appendix Table A.3, we further test heterogeneous effect by surgery complexity levels (level I being the easiest, and level IV being the hardest), finding no evidence that the impact of pollution is systematically larger for harder procedures.

To be clear, while physician effects do not appear to be the main channel, we cannot completely *rule out* the possibility due to the limited information we have on physician characteristics. Future research may better exploit this channel using other physician-level mediating factors, such as experience ([Sosa et al., 1998](#)).

**Patient Effects.** Now turning to the patient side. A natural interpretation of our findings is that surgical patients are not shielded from air pollution after all. One way to tease out a patient effect is to look at intermediate outcomes that are well known to be directly influenced by pollution exposure. We are unaware of evidence from the medical literature on the effect of *post-surgery* air pollution exposure on patient health. However, we may draw some testable hypotheses from the literature that considers the effect of *intra-surgery* air quality and surgical complications. Since the 1970s, airborne bacteria in the operating room has been recognized as an important source of surgical site infection ([Gryska and O’Dea, 1970](#); [Charnley, 1972](#); [Lidwell et al., 1982](#)), which prompted the use of high-efficiency air-filtration technologies in operating room ([Mangram et al., 1999](#); [American Institute of Architects, 2001](#); [Chinn and Schulster, 2003](#)). We show with our data that such pollution-infection link holds in the post-surgery context as well: Appendix Table A.4 panel A suggests an increase of incidents of infection as outdoor pollution level rises. This result may arise as the typical surgical wound takes 24 to 48 hours to close after the operation ([Mangram et al., 1999](#)), leaving the patient susceptible to exogenous infections from in-hospital exposure to pollution. Panel B shows the impact of pollution appears to be larger among patients with diabetes and hypertension, consistent with prior findings that these chronic conditions amplify the impact of air pollution on inflammatory responses ([Dubowsky et al., 2006](#)). Panel C suggests the same heterogeneity pattern holds for hospital mortality as well.

Our data also allow us to benchmark our findings with those from several other commonly performed tests in prior studies on the effect of ambient air pollution on health of the general population. First, we estimate a concentration-response relationship by replacing the [Pollution<sub>i</sub>](#) term in equation (1) with quintile bins of PM<sub>2.5</sub> concentrations. This model thus allows for a flexibly nonlinear relationship



between post-surgery mortality and  $PM_{2.5}$  levels. Figure 3 presents the results. The first quintile bin (average  $PM_{2.5} = 14.8 \text{ ug/m}^3$ ) is the omitted category, and the other estimates show the differences in mortality rate on higher-pollution days relative to days when pollution falls in the first quintile bin. Note that the relationship between this figure and the point estimate  $\beta$  in Table 2:  $\beta$  is essentially the slope of a linear fit across the five point estimates in the concentration-response figure, while the figure allows the relationship to be nonlinear. Figure 3 suggests that the mortality damage of pollution rises quickly at first, and then flattens out when it passes the third quintile (average  $PM_{2.5} = 32.4 \text{ ug/m}^3$ ). A similar “concave” pattern between general and cardiovascular mortality and both short- and long-term  $PM_{2.5}$  exposure has been reported in other epidemiological studies (e.g., [Pope III et al., 2011](#); [Crouse et al., 2012](#); [Pope III, Cropper, Coggins, and Cohen, 2015](#)). The dosage relationship suggests that pollution is the culprit underlying the mortality damage, and that even exposure at the low end of the distribution deserves attention.

Next, we augment equation (1) with additional air pollutant covariates, including ozone ( $O_3$ ), nitrogen dioxide ( $NO_2$ ), sulfur dioxide ( $SO_2$ ), and carbon monoxide (CO).<sup>18</sup> Table 5 columns 1 through 4 sequentially add these pollutants as an additional covariate in equation (1), whereas column 5 includes all five pollutants simultaneously. The coefficient estimates from these regressions should be interpreted with caution because these pollutant terms are interdependent. For example, both nitrates and sulfates act as precursors to particulate matter and ozone formation; high particulate concentration scavenges hydroperoxyl radicals, which suppress ozone production. However, Table 5 does suggest that  $PM_{2.5}$  is a robust and statistically significant predictor of post-surgery survival. This finding echoes a strand of the epidemiology literature that suggests that ozone does not have a first-order impact on cardiovascular morbidity and mortality once  $PM_{2.5}$  is controlled for ([Jerrett et al., 2009](#); [Brook et al., 2010](#)). In the related causal inference literature, [Deryugina et al. \(2019\)](#) examine a two-pollutant model in which  $PM_{2.5}$  and  $O_3$  are simultaneously instrumented for using variations in local wind directions. They find that the estimated effect of  $PM_{2.5}$  on elderly mortality is five times as large as the  $O_3$  effect, with the  $O_3$  effect showing the “wrong” sign (mortality reduction). Table 5 column 1 shows that a similar conclusion holds in the context of patients’ post-surgery survival.

Finally, we try to identify high-risk patients. The impact of pollution likely varies across different groups of patients. Table 6 reports estimates from an augmented version of regression equation (1) in which we fully interact  $PM_{2.5}$  with an indicator for a proposed high-risk group of patients – respiratory and neoplasm surgeries for patients aged over 60. This regression specification allows us to estimate the impact

---

<sup>18</sup> The same monitoring site that samples  $PM_{2.5}$  is also equipped with samplers for these other air pollutants.

of surgery-day  $PM_{2.5}$  separately for the high-risk group patients and all other patients.<sup>19</sup> Table 6 suggests the impact of pollution concentrates in the high-risk group, where a log unit increase in  $PM_{2.5}$  leads to a 2.91 per 1,000 patients increase in post-surgery mortality (the average coefficient estimate across four columns). The corresponding increase for other patients is 0.17 per 1,000 patients. Part of this difference is because the high-risk group has a higher *average* post-surgery mortality rate: 38.35 per 1,000 patients, compared to 8.74 per 1,000 patients for those who are not in this high-risk group. Note, however, that the effect of  $PM_{2.5}$  is also larger even in *percentage* terms for the high-risk group: a 7.59 percent increase in mortality surfaces per one log unit increase in pollution, compared with a 1.96 percent increase among other patients.

This is an important finding we will leverage in our structural exploration in Section 4. High-risk patients constitute a small fraction (6 percent) of surgery cases but account a majority (60 percent) of the mortality effect. This provides us with opportunity to achieve survival improvement by targeting better surgery scheduling to address this particular group of patients. We discuss this point in greater detail in Section 4.

To further examine plausibility, we assess the exact surgical procedures performed with the high-risk group. For respiratory patients, the five most common procedures are *fiber-optic bronchoscopy* (14.9% of all respiratory procedures, ICD-9-CM Volume 3 code: 33.22), *other bronchoscopy* (6.5%, code 33.23), *insertion of intercostal catheter for drainage* (6.3%, code 34.04), *insertion of endotracheal tube* (4.9%, code 96.04), and *venous catheterization* (4.8%, code 38.93). For neoplasm patients, the most common procedures are *other gastroscopy* (2.1%, code 44.13), *endoscopic destruction of other lesion or tissue of large intestine* (2.0%, code 45.43), *endoscopic polypectomy of large intestine* (2.0%, code 45.42), *other transurethral excision or destruction of lesion or tissue of bladder* (1.9%, code 57.49), and *closed percutaneous needle biopsy of lung* (1.9%, code 33.26). Our findings suggest that it is possible that pollution is the most detrimental for those who have recently undergone endoscopic procedures.

**Summary.** We briefly summarize our findings about underlying mechanisms. First, statistical selection of surgery patients is unlikely to be the explanation. This is supported by various analyses of patient characteristics, which do not differ in the surgeries performed on days with high or low levels of pollution. An instrumental-variable design exploiting transported pollution from upwind cities also supports the claim. Second, while we do not fully know whether the effect works through pollution’s effect of physician treatment decision and/or surgery performance, we are able to reject some specific channels that can serve as proxies for greater errors or issues that might occur on highly polluted days. A lack of

---

<sup>19</sup> We came up with this high-risk group by first estimating similar heterogeneous effect specifications by diagnosis categories and by age groups separately. These first-step results are reported in Appendix Tables A.4 and A.5.

information on attending surgeons and other medical personnel precludes us from exploring further possibilities. Third, the same toxicological and biological mechanisms that underpin general mortality effects of pollution may also be at play in the perioperative context. This is reflected by some similar effect characteristics reported in previous studies such as those that examine concentration response, pollutant-specific effects, and subpopulation heterogeneity.

## 4. (Re)scheduling Pollution Exposure

The empirical evidence in Section 3 invites a question: is it possible to improve a patient’s post-surgery survival by avoiding operating on days when pollution is high? In this section, we use structural methods to characterize the hospitals’ status quo decision-making processes that govern patient surgery scheduling. We then analyze how hospitals could incorporate knowledge about the potential adverse impact of pollution on patients’ mortality risks.

We keep a “Pareto” principle in mind throughout the structural exercise. That is, while some patients are made better off when surgery is rescheduled to a day with better air quality, no patient should be made worse off under the counterfactual schedule. With this in mind, we seek to switch surgery dates for only a small group of patients who would benefit the most from having surgeries on days with lower levels of pollution, and we only consider small deviations from originally scheduled dates. This conservative approach leads to a small impact on the hospital’s overall surgery capacity utilization, while still able to discover substantial opportunity for survival improvement.

### 4.1 A Model of Surgery Scheduling

**Setup and Parameterization.** We model the decision of a hospital to schedule surgeries for newly admitted patients. The hospital chooses the operation date for each patient as  $d$  days after the admission date. For simplicity, we consider a static version of the problem, where we take the pool of all  $N$  patients ever admitted to the hospital during our study period, and model how the hospital optimally chooses the admission-to-surgery delay profile  $\{d_i\}_{i=1}^N$ . Below, we use subscript  $i$  to denote a patient. The hospital maximizes its utility function for scheduling an operation for patient  $i$  on the  $d$  days following admission:

$$u_{id} = -\alpha h_{id} + \lambda_{id} + e_{id} \quad (3)$$

where  $h_{id}$  is the *perceived* patient mortality hazard from performing a procedure,  $\alpha$  represents the scaling weight of the mortality hazard when the hospital makes scheduling decisions,  $\lambda_{id}$  captures non-health-related payoffs, and  $e_{id}$  captures idiosyncratic considerations of the hospital in scheduling surgeries, which we assume follow an i.i.d. type-I extreme value distribution.

The perceived mortality hazard  $h_{id}$  follows the exact same specification as in our reduced form regression equation (1), except that the **Pollution** term is omitted because we assume hospitals do not internalize the mortality hazard of pollution in the baseline model. We parameterize non-health-related payoffs  $\lambda_{id}$  as a stepwise function of admission-to-surgery delays  $d$  (days), and a dummy variable indicating whether the surgery day is on a workday:

$$\lambda_{id} = \sum_g \beta_g \cdot 1(d_{id} = g) + \phi_{\text{workday}} \quad (4)$$

The workday indicator variable  $\phi_{\text{workday}}$  is important because both air pollution and surgery mortality exhibit strong weekday-weekend cycles. Our data suggest that the odds of post-surgery death are higher for procedures done on the weekend. On the other hand, air pollution tends to be higher during the weekdays.  $\phi_{\text{workday}}$  may also capture differences in physician availability between weekdays and weekends, which matters for hospitals' cost considerations.

**Patient Pool.** Next, we impose several constraints to our modeling exercise to maximize the practical relevance of our counterfactual analysis. First, we focus on patients receiving respiratory and neoplasm surgeries who were over 60 years old at the time of admission. As we report in Section 3, the pollution effect concentrates among this vulnerable group, which constitutes of 6% of the overall surgery sample but explains 60% of the total effect size. From a practical perspective, restricting to a relatively small group of procedures helps avoid causing periods of extremes in surgical capacity. Only 1.3% of counterfactual case assignments exceed capacity, which can be handled by reassigning to the second-best day (more below).

Second, we set the maximum value of delay  $d$  to be three days, meaning that we only model scheduling for patients whose observed surgery dates are up to three days after the date of hospital admission. That is, we assume that the hospital chooses delay  $d \in \{0, 1, 2, 3\}$ . About 41% of age 60+ respiratory/neoplasm patients fall in this situation ( $N = 130,440$ ). This restriction also implies that all counterfactual surgery schedules occur within this window. This specification choice is motivated by several considerations. First and foremost, air quality forecasting is the most accurate in the near term. Section 4.4 presents evidence on short-term accuracy of pollution forecast in Guangzhou. Our modeling assumption that delays *per se* have no effect on mortality hazards ( $h_{id}$ ) is also more likely to hold if we

consider short-term rescheduling. Other practical issues may arise as well, such as patient agreement and the deterioration of patient’s condition, when long delays are involved.

In summary, in subsequent analysis we assume that (a) it is feasible for surgeries in the high-risk group to deviate from its original schedule within a three-day window – for example, for a patient whose observed surgery occurred three days after hospital admission, it was feasible for the hospital to schedule the patient’s surgery to an earlier day such the admission day, and vice versa; and (b) such short-term deviation has negligible impact on patient’s post-surgery mortality. In appendix, we document supportive evidence for these assumptions. First, we show that, in the *observed* surgery schedules, it is common for elderly patients to receive respiratory and cancer surgeries on each day of the  $d \in \{0, 1, 2, 3\}$  window (Appendix Figure A.5). To further address the concern that switching surgeries to earlier days may be costly due to pre-surgery preparation etc., in Section 4.2 we provide an example surgery *postponing* rule in which the hospital preserves the status quo scheduling practice while pushing high-risk surgeries back if pollution on the surgery day turns out to be exceptionally high. Second, we re-estimate our primary regression equation (1) among our patient pool using admission-to-surgery delay as the independent variable. We find statistically insignificant effect of delay on post-surgery mortality with a “wrong” sign (longer delay is associated with a statistically insignificant decrease in patient mortality), suggesting short-term differences in surgery delay has limited impact on patient survival.

**Estimation.** Our estimation therefore boils down to a maximum likelihood estimation of parameters  $\theta = (\alpha, \beta_0, \beta_1, \beta_2, \varphi_{\text{workday}})$ . Assuming  $e_{id}$  follows a type-I extreme value distribution, the probability that the patient  $i$  is assigned a delay of  $d$  is

$$P_{id}(\theta) = \frac{\exp(\alpha h_{id} + \lambda_{id})}{\sum_{g=0}^3 \exp(\alpha h_{ig} + \lambda_{ig})} \quad (5)$$

where  $h_{id}$  is the predicted mortality risk for the patient  $i$  from our reduced form estimation of equation (1).  $\lambda_{id}$  follows the exact definition from equation (4), with an additional  $d$  index to make explicit that the term depends on the delay. The patient’s individual likelihood is given by

$$R_i(\theta) = P_{id}(\theta) \prod_{g \neq d} (1 - P_{ig}(\theta)) \quad (6)$$

Because optimal scheduling decisions are independent across patients, the model likelihood function is

$$L(\theta) = \prod_{i=1}^N R_i(\theta) \quad (7)$$

We report estimation results  $\hat{\theta}$  in Appendix Table A.7. Note we estimate the model separately for respiratory and neoplasm patients because the health and non-health trade-offs are expected to be different

across different diagnoses. In Appendix Figure A.6 we report that the predicted surgery schedule from these parameterizations matches well with the observed schedule for both respiratory and neoplasm patient groups.

**Counterfactual Analysis.** We now consider counterfactual surgery scheduling that reflects what could be devised if a hospital were to internalize the impact of air pollution on post-surgery survival (Table 2). We compute predicted mortality hazard  $h'_{id}$  using equation (1), now with the **Pollution** term “switched on.” We then substitute  $h'_{id}$  into equation (5) holding fixed  $\hat{\theta}$ , which gives us a list of probabilities  $P_{id}(\hat{\theta})$  that patient  $i$  is assigned surgery with a delay  $d_i \in \{0,1,2,3\}$ . Patient  $i$ ’s counterfactual surgery delay is therefore corresponding to the day with the maximum (counterfactual) surgery probability, namely  $\text{argmax}_{d \in \{0,1,2,3\}} P_{id}(\hat{\theta})$ . We repeat this exercise for every patient, and generate a counterfactual surgery schedule  $\{d'_i\}_{i=1}^N$ .

Figure 4 summarizes the results. Panel A displays the observed and counterfactual surgery-day  $PM_{2.5}$  distributions among patients in the high-risk group (respiratory/neoplasm patients aged 60+). The counterfactual distribution is much more skewed to the right, compared to the observed distribution, suggesting that a significant share of high-pollution surgeries are shifted towards days with better air quality.

With many cases shifting, one might worry that the counterfactual schedule may exceed the hospital department’s surgery capacity, especially during days with particularly low levels of air pollution. To gauge the likelihood of such situation, we estimate each hospital’s respiratory and neoplasm surgery capacity as the maximum single-day surgery patient volume during the study period (mean = 10.0, s.d. = 5.0 for respiratory surgeries; mean = 33.4, s.d. = 26.8 for neoplasm surgeries). We calculate that in about 1.3% cases, the counterfactual scheduling causes an assignment of the patient to a date when the hospital department’s surgery capacity is already full. Panel B of Figure 4 shows average hospital respiratory and neoplasm surgery capacity utilization rate by  $1\text{-ug}/\text{m}^3$  pollution bins. The pattern suggests that while the counterfactual scheduling makes the surgery departments a little busier on low-pollution days, such scheduling does not come close to jeopardizing overall capacity. This result stems largely from the focus tailoring the rescheduling to target the vulnerable groups, which constitute a small share of patients, but a large share of health damage.

In panel C of Figure 4, we plot the distribution of improvement in survival probability for “switcher” patients whose observed surgery dates are different from the counterfactual. Overall, our method switched the date of surgery for 20,853 patients out of a total of 52,406 patients in the vulnerable group. A vast-majority of switchers (98.6%) improved survival probability. This is a desirable result:

although hospitals could in principle switch patients to better air quality days with higher revenue potential but *worse* mortality potential, such behavior does not often arise according to the revealed health and revenue trade-offs for most hospitals in our data. The mortality reduction for the average switcher is 1.6 deaths per 1,000 patients, or about a 4.2 percent improvement upon the baseline mortality rate of 38.4 per 1,000 patients in the high-risk group.

The big picture takeaway from Figure 4 is that for nearly half of the scheduled surgeries, there exists an alternative, lower-pollution day within three days of the originally scheduled day such that moving the surgery would (a) not surpass the overall surgery capacity of the alternative day, (b) yield an average 4 percent better post-surgery survival, and (c) meet the basic cost-and-benefit trade-offs according to revealed hospital preference from historical surgery records.

## 4.2 A Practical Example

The structural exercise suggests there are potential gains from *taking pollution into account* when scheduling surgeries. This could mean many things in practice: more salient reminders of pollution levels, better education on the adverse effect of pollution on patients, or an established protocol on avoiding surgeries on vulnerable patients on high-pollution days, for example. Exactly which approach to adopt depends on the individual hospital's constraints and preferences. Here we consider the feasibility of a simple rescheduling (postponing) process that responds to a high level of pollution on a given day by delaying scheduled surgery to the near future. The intention of this exercise is to set a concrete example of how hospitals *might* consider pollution in scheduling decisions.

(a) We begin with 6,093 surgeries among the high-risk group that are scheduled on days when the Air Quality Index is over 100 ( $PM_{2.5} > 75 \text{ ug/m}^3$ ) with Air Quality Index below 100 during at least one of the following three days.<sup>20</sup> Our goal is to reschedule these surgeries scheduled on days with high pollution levels to one of the subsequent days with lower pollution to improve post-surgery survival (according to the health function estimated in Section 3).

(b) Consider rescheduling all 6,093 patients to the day with the *lowest* level of pollution in the upcoming three-day period. Although exposure is reduced in every case, survival does not necessarily improve because pollution is not the only determinant of mortality hazards ( $h_{id}$ ). For example, a patient could be reassigned from a weekday that has high pollution to a weekend day with low pollution – even though, as we have previously discussed, the mortality hazard is on average higher during the weekend.

---

<sup>20</sup>  $PM_{2.5}$  concentration exceeds  $75 \text{ ug/m}^3$  in about 5 percent of days in Guangzhou.

Our analysis shows that under this scenario, 1,935 of the 6,093 patients would actually see a decline in survival.

(c) Thus, to make sure that no patient is harmed by rescheduling, we reassign the cases that experience negative effects to the day with the *second-lowest* level of pollution day that occurs in the following three-day period, or, if that scenario still results in negative effects, we reschedule again to the day with the *third-lowest* pollution day. With such adjustments, we are able to obtain positive survival benefits for 902 out of the 1,935 patients who would have potentially experienced negative effects from a blanket approach to rescheduling. For the rest of 1,033 patients, the second / third lowest pollution day over the next three days, in fact, have mortality risks that are higher than those on the originally scheduled day of surgery. Therefore, for these patients, the surgery schedule would remain the same.

(d) Over the course of the rescheduling exercise, there are 238 incidents in which the hospital would have found its daily surgery capacity would be exceeded. We treat these cases in the same way we did with negative survival improvement cases, assigning them to the next-best air quality day. As in Section 4.1, surgery capacity overall is not a binding constraint because we focus on a relatively small group of patients.

Appendix Figure A.7 presents the effect of rescheduling process on patients' pollution exposure, hospital capacity utilization, and survival improvements. Panel A shows the observed surgery-day  $PM_{2.5}$  value for the 6,093 patients; by design, these values all lie above AQI 100 ( $75 \mu g/m^3$ ). Panel A also contrasts this with the counterfactual exposure that would have occurred if steps (a) through (d) were followed. Panel B shows that rescheduling increases surgery capacity utilization of the associated hospital departments on low pollution days, but not to such an extent that the rescheduling would approach capacity constraints. Panel C illustrates the distribution of survival improvements after step (a) shown by the hollow bar, and after steps (a) through (d) shown by the filled bar. Overall, the process switched surgery days for 5,060 patients out of a total of 6,093 patients. The mortality reduction rate for the average switcher is 2.8 deaths per 1,000 patients, which is about a 7.3 percent improvement upon the baseline mortality rate of 38.4 per 1,000 patients in the high-risk group.<sup>21</sup>

---

<sup>21</sup> The *per-patient* (rather than per-switcher) improvement is 0.11 per 1,000 patients. This is much smaller than the structural results in Section 4.1. This is both because the rescheduling process only focuses on a subset of days with very high pollution levels, and because the process does not seek to strike the optimal cost-benefit trade-offs as does the structural method.



### 4.3 Compare Effect Sizes to Existing Surgical Improvement Programs

From Sections 4.1 and 4.2, the mortality benefits of the rescheduling exercises among the switchers range from **1.6 to 2.8 deaths per 1,000 patients**. How large are these effect sizes? To put numbers in context, here we briefly review two surgical quality improvement programs that are widely studied in the medical literature.

**The Safe Surgery Saves Lives Program.** In 2008, the World Health Organization (WHO) published a surgical safety checklist identifying multiple recommended practices to ensure the safety of surgical patients ([WHO, 2008](#)). Some example entries of the checklist include “Before induction of anesthesia, members of the team orally confirm that the patient has verified his or her identity, the surgical site and procedure, and consent”; “Before skin incision, the entire team orally confirms that prophylactic antibiotics have been administered  $\leq 60$  min before incision is made or that antibiotics are not indicated”; “Before the patient leaves the operating room, nurse reviews items aloud with the team that the needle, sponge, and instrument counts are complete (or not applicable)”. [Haynes et al. \(2009\)](#) and [de Vries et al. \(2010\)](#) studied a pilot intervention program that instituted such a checklist in eight hospitals of different countries. Results indicate patient rate of death decreased from 1.5 percent (i.e., 15 deaths per 1,000 patients) before the program to 0.8 percent after the program. The absolute mortality gain is thus about **7 deaths per 1,000 patients**.

**The National Surgical Quality Improvement Program.** The National Surgical Quality Improvement Program (NSQIP) was developed in the 1990s by the Veterans Health Affairs to improve surgical outcomes in the Veterans Affairs (VA) Health System. The program involves annual reports of patient characteristics and outcomes, peer conducted site visits, among other measures, and is shown to have led to substantial improvements in surgical quality ([Ingraham et al., 2010](#)). NSQIP is since introduced to non-VA settings. [Hall et al. \(2009\)](#) studied the American College of Surgeons NSQIP program that was implemented in 118 private-sector hospitals since 2004. The results indicate a mean improvement of observed/expected mortality by 0.174 based on an average mortality rate of about 1.8 percent, which works to about an absolute mortality gain of **3 deaths per 1,000 patients**.

In terms of absolute mortality reduction, the pollution rescheduling exercise generates an effect size with same order of magnitude compared to these well-known surgical quality improvement programs. Because we study a high-risk patient group with higher baseline mortality rate (about 38 deaths per 1,000 patients), the relative effect size of pollution rescheduling can be up to an order of magnitude smaller than these programs (e.g., 4.2 percent mortality improvement in Section 4.1, compared to about 47 percent improvement in the Safe Surgery Saves Lives Program).

## 4.4 Limitations and Future Research

Before concluding the paper, we discuss several issues that we are not able to fully pin down with this study, and potential directions for future work.

**Static versus Dynamic Scheduling Problem.** For simplicity, we employ a retrospective, static approach in the structural scheduling exercise. In reality, hospitals schedule patients sequentially based on dynamically changing surgery capacity. Another state variable in a dynamic scheduling is the patient’s health condition, which changes stochastically over time. The scheduling decision is thus based on the patient’s health condition upon arrival *and* the expectation of future health changes. Certain patients may even need multiple operations and frequent adjustments of treatment plans as a function of surgical outcomes.

Although we abstract away from these dynamic details, static models may still be a reasonable approximation because we only consider a local change in observed surgery scheduling. In particular, we have shown that surgery capacity is far from binding on a vast majority of days both with the observed and the counterfactual scheduling. Our focus on the short-term (i.e., next three-day) scheduling horizon also limits the scope for stochastic changes in patient health changes.

**Accuracy of Pollution Forecasts.** Throughout the counterfactual scheduling exercise we assume that hospitals can perfectly anticipate air pollution levels in the next three days. Here we survey the current forecasting technology used by Guangzhou and its performance.

Air pollution forecast in Guangzhou is primarily based on outputs from four Eulerian Chemical Transport Models: NAQPMS, CMAQ, CAMx, and WRF-Chem.<sup>22</sup> These models combine weather predictions with data from ambient pollution monitoring, emission monitoring, and emission inventory to provide simulations of atmospheric chemistry, and to generate numerical forecasts of air quality in the next 72 hours. The automated model outputs are then post-processed by scientists at the Guangzhou Environmental Monitoring Center and the Weather Bureau. The scientists use statistical methods and ad hoc knowledge on pollution events to improve upon the numerical forecasts.

We do not have access to historical pollution forecasts in Guangzhou. Instead, we rely on several reports from the city, province, and national environmental agencies that compare forecasts and observed

---

<sup>22</sup> These are the Nested Air Quality Prediction Modeling System developed by the Chinese Academy of Sciences, the Community Multiscale Air Quality Model developed by the U.S. Environmental Protection Agency, the Comprehensive Air Quality Model with Extensions by the Randoll U.S. Corporation, and the Weather Research and Forecasting Model Coupled to Chemistry by the U.S. National Oceanic and Atmospheric Administration.

pollution. Appendix Figure A.8 shows daily time-series plots of the 24-hour forecasted and observed Air Quality Index (AQI) in Guangzhou throughout the year of 2016 (taken from [Zhang et al., 2017](#) and [Zhang et al., 2018](#)). These plots suggest that forecasted AQIs match well overall with observed values. The correlation coefficient between the two time series is 0.72. A simple linear regression of observed AQI on forecasted AQI yields a slope coefficient of 0.95. A separate, province-wide analysis by [Shen et al. \(2017\)](#) shows 24-hour forecasted AQI categories (below 50, the “Green-Good” category; between 50 and 100, the “Yellow-Moderate” category, etc., as shown in Appendix Figure A.2) are correct over 70 percent of the time in Guangzhou, with almost perfect accuracy when  $PM_{2.5}$  is the predominant pollutant of the day. The only report of longer-term, 72-hour forecast accuracy we can find is in the data appendix of the central government’s Technical Guideline for Numerical Forecasting of Ambient Air Quality of 2020 (HJ 1130-2020). Using forecasting data from 17 cities, the Technical Guideline reports that forecasted AQI categories coincide with observed categories 75 percent of the time, with the hit rates similar across 24-hour, 48-hour, and 72-hour time horizons.

Both in China and internationally, air pollution forecasting is a growing and high-interest research field. As we argue in this paper, endowing the general public with better abilities to anticipate air pollution events may have important health values. In Appendix Figure 9, we repeat the main rescheduling exercise of Figure 4 but infuse the true  $PM_{2.5}$  with mean-zero, normally-distributed errors  $N(0, \sigma)$ . We show that more accurate forecasts (smaller  $\sigma$ ) lead to better patient survival through rescheduling.

**Physician Mechanisms.** One question we did not fully address in Section 3 is whether a physician effect may explain the main mortality finding. We excluded some channels, such as injury-related medical errors, but a lack of information on attending surgeons and other medical personnel precludes us from exploring further possibilities.

The rescheduling exercise per se does not fully hinge on knowledge of the mechanism: so long as surgery-day pollution is a cause for higher mortality, rescheduling to alternative surgery days may provide benefit – no matter what the exact underlying mechanisms at work. That said, a fuller understanding on the physician effect may help create even better surgery scheduling practices. For example, if the pollution effect only changes performance of a certain group of physicians, it might be more efficient to re-optimize surgeon schedules rather than patient schedules.

## 5. Conclusion

We use data from more than 1 million surgical records from a major city in China to study the link between day-to-day variation in fine particulate matter pollution (mean = 36.5  $\mu g/m^3$ , s.d. = 19.8  $\mu g/m^3$ )

and post-surgery mortality. We find that undergoing an operation on a high pollution day is associated with a significant reduction in post-surgery survival – especially elderly patients undergoing respiratory and neoplasm operations. The excess mortality is linked to pollution on the day of the surgery, rather than pollution exposure prior to hospital admission, during the waiting period, or following the date of the operation. Building upon these empirical patterns, we build and analyze a structural model of hospital surgery scheduling, and consider counterfactual scheduling that could improve patient survival. We demonstrate that relatively small changes in surgical dates for a small group of patients can improve survival rates. Such changes in scheduling can accommodate the patients who are likely to benefit most from undergoing such procedures on a day with better air quality. We conclude that there may be important potential for hospitals to consider air pollution forecasts in day-to-day surgery scheduling.

## References

- Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. Management and shocks to worker productivity. No. w25865. National Bureau of Economic Research, 2019.
- Aiken, Linda H., Sean P. Clarke, Douglas M. Sloane, Julie Sochalski, and Jeffrey H. Silber. "Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction." *JAMA* 288, no. 16 (2002): 1987-1993.
- Alves, Arnaud, Yves Panis, Pierre Mathieu, Georges Mantion, Fabrice Kwiatkowski, and Karem Slim. "Postoperative mortality and morbidity in French patients undergoing colorectal surgery: results of a prospective multicenter study." *Archives of Surgery* 140, no. 3 (2005): 278-283.
- American Institute of Architects. Academy of Architecture for Health, and Facilities Guidelines Institute. Guidelines for design and construction of hospital and health care facilities. Aia Press, 2001.
- Anderson, Michael L. "As the wind blows: The effects of long-term exposure to air pollution on mortality." *Journal of the European Economic Association* 18, no. 4 (2020): 1886-1927.
- Andrews, Steven Q. "Inconsistencies in air quality metrics: 'Blue Sky' days and PM10 concentrations in Beijing." *Environmental Research Letters* 3, no. 3 (2008): 034009.
- Archsmith, James, Anthony Heyes, and Soodeh Saberian. "Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation." *Journal of the Association of Environmental and Resource Economists* 5, no. 4 (2018): 827-863.
- Bainbridge, Daniel, Janet Martin, Miguel Arango, Davy Cheng, and Evidence-based Peri-operative Clinical Outcomes Research (EPiCOR) Group. "Perioperative and anaesthetic-related mortality in developed and developing countries: a systematic review and meta-analysis." *The Lancet* 380, no. 9847 (2012): 1075-1081.

- Ball, Jane E., Luk Bruyneel, Linda H. Aiken, Walter Sermeus, Douglas M. Sloane, Anne Marie Rafferty, Rikard Lindqvist, Carol Tishelman, Peter Griffiths, and RN4Cast Consortium. "Post-operative mortality, missed care and nurse staffing in nine countries: a cross-sectional study." *International Journal of Nursing Studies* 78 (2018): 10-15.
- Barron, Manuel, and Maximo Torero. "Household electrification and indoor air pollution." *Journal of Environmental Economics and Management* 86 (2017): 81-92.
- Barwick, Panle Jia, Shanjun Li, Liguang Lin, and Eric Zou. *From fog to smog: The value of pollution information*. No. w26541. National Bureau of Economic Research, 2019.
- Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur. *The morbidity cost of air pollution: evidence from consumer spending in China*. No. w24688. National Bureau of Economic Research, 2018.
- Bhaduri, Budhendra, Edward Bright, Phillip Coleman, and Marie L. Urban. "LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics." *GeoJournal* 69, no. 1 (2007): 103-117.
- Bozic, Kevin J., Edmund Lau, Steven Kurtz, Kevin Ong, and Daniel J. Berry. "Patient-related risk factors for postoperative mortality and periprosthetic joint infection in medicare patients undergoing TKA." *Clinical Orthopaedics and Related Research* 470, no. 1 (2012): 130-137.
- Brook, Robert D., Sanjay Rajagopalan, C. Arden Pope III, Jeffrey R. Brook, Aruni Bhatnagar, Ana V. DiezRoux, Fernando Holguin et al. "Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association." *Circulation* 121, no. 21 (2010): 2331-2378.
- Cardoen, Brecht, Erik Demeulemeester, and Jeroen Beliën. "Operating room planning and scheduling: A literature review." *European Journal of Operational Research* 201, no. 3 (2010): 921-932.
- Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell. "The effect of pollution on worker productivity: evidence from call center workers in China." *American Economic Journal: Applied Economics* 11, no. 1 (2019): 151-72.
- Charnley J. *Post-operative infection after total hip replacement with special reference to contamination in the operating room*. Internal Publication 38, Centre for Hip Surgery, Wrightington Hospital, Wigan, Lancs., UK; 1972.
- Chay, Kenneth, Carlos Dobkin, and Michael Greenstone. "The Clean Air Act of 1970 and adult mortality." *Journal of Risk and Uncertainty* 27, no. 3 (2003): 279-300.
- Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. "Gaming in air pollution data? Lessons from China." *The BE Journal of Economic Analysis & Policy* 13, no. 3 (2012).
- Chen, Chun, and Bin Zhao. "Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor." *Atmospheric Environment* 45, no. 2 (2011): 275-288.
- Chinn, Raymond YW, and Lynne Schulster. "Guidelines for environmental infection control in health-care facilities; recommendations of CDC and Healthcare Infection Control Practices Advisory Committee (HICPAC)." (2003).

Clifford, Angela, Linda Lang, Ruoling Chen, Kaarin J. Anstey, and Anthony Seaton. "Exposure to air pollution and cognitive functioning across the life course—a systematic literature review." *Environmental Research* 147 (2016): 383-398.

Cohen, Aaron J., Michael Brauer, Richard Burnett, H. Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan et al. "Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015." *The Lancet* 389, no. 10082 (2017): 1907-1918.

Crouse, Dan L., Paul A. Peters, Aaron van Donkelaar, Mark S. Goldberg, Paul J. Villeneuve, Orly Brion, Saeeda Khan et al. "Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study." *Environmental Health Perspectives* 120, no. 5 (2012): 708-714.

Currie, Janet, and Matthew Neidell. "Air pollution and infant health: what can we learn from California's recent experience?" *The Quarterly Journal of Economics* 120, no. 3 (2005): 1003-1030.

Cutter, W. Bowman, and Matthew Neidell. "Voluntary information programs and environmental regulation: Evidence from 'Spare the Air'." *Journal of Environmental Economics and Management* 58, no. 3 (2009): 253-265.

Cyrus, Josef, Mike Pitz, Wolfgang Bischof, H - Erich Wichmann, and Joachim Heinrich. "Relationship between indoor and outdoor levels of fine particle mass, particle number concentrations and black smoke under different ventilation conditions." *Journal of Exposure Science & Environmental Epidemiology* 14, no. 4 (2004): 275-283.

David, Guy, Candace L. Gunnarsson, Heidi C. Waters, Ruslan Horblyuk, and Harold S. Kaplan. "Economic measurement of medical errors using a hospital claims database." *Value in Health* 16, no. 2 (2013): 305-310.

de Vries, Eefje N., Hubert A. Prins, Rogier MPH Crolla, Adriaan J. den Outer, George van Andel, Sven H. van Helden, Wolfgang S. Schlack et al. "Effect of a comprehensive surgical safety system on patient outcomes." *New England Journal of Medicine* 363, no. 20 (2010): 1928-1937.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. "The mortality and medical costs of air pollution: Evidence from changes in wind direction." *American Economic Review* 109, no. 12 (2019): 4178-4219.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro. "Defensive investments and the demand for air quality: Evidence from the NOx budget program." *American Economic Review* 107, no. 10 (2017): 2958-89.

Dominici, Francesca, Michael Greenstone, and Cass R. Sunstein. "Particulate matter matters." *Science* 344, no. 6181 (2014): 257-259.

Dubowsky, Sara D., Helen Suh, Joel Schwartz, Brent A. Coull, and Diane R. Gold. "Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation." *Environmental Health Perspectives* 114, no. 7 (2006): 992-998.

Duflo, Esther, Michael Greenstone, and Rema Hanna. "Indoor air pollution, health and economic well-being." *SAPI EN. S. Surveys and Perspectives Integrating Environment and Society* 1.1 (2008).

- Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114, no. 39 (2017): 10384-10389.
- Ghaferi, Amir A., John D. Birkmeyer, and Justin B. Dimick. "Variation in hospital mortality associated with inpatient surgery." *New England Journal of Medicine* 361, no. 14 (2009): 1368-1375.
- Ghanem, Dalia, Shu Shen, and Junjie Zhang. "A Censored Maximum Likelihood Approach to Quantifying Manipulation in China's Air Pollution Data." *Journal of the Association of Environmental and Resource Economists* 7, no. 5 (2020): 965-1003.
- Ghanem, Dalia, and Junjie Zhang. "'Effortless Perfection:' Do Chinese cities manipulate air pollution data?." *Journal of Environmental Economics and Management* 68, no. 2 (2014): 203-225.
- Gong, Qing. Physician learning and treatment choices: Evidence from brain aneurysms. (2018). Working paper.
- Graff Zivin, Joshua, and Matthew Neidell. "Days of haze: Environmental information disclosure and intertemporal avoidance behavior." *Journal of Environmental Economics and Management* 58, no. 2 (2009): 119-128.
- Graff Zivin, Joshua, and Matthew Neidell. "The impact of pollution on worker productivity." *American Economic Review* 102, no. 7 (2012): 3652-73.
- Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution. No. w27502. National Bureau of Economic Research, 2020.
- Greenstone, Michael, Kenneth Lee, and Harshil Sahai. "Indoor Air Quality, Information, and Socioeconomic Status: Evidence from Delhi." In *AEA Papers and Proceedings*, vol. 111, pp. 420-24. 2021.
- Gryska, Paul F., and Arthur E. O'Dea. "Postoperative streptococcal wound infection: the anatomy of an epidemic." *JAMA* 213, no. 7 (1970): 1189-1191.
- Guidetti, Bruna, Paula Pereda, and Edson Severnini. "Health Shocks under Hospital Capacity Constraint: Evidence from Air Pollution in Sao Paulo, Brazil." (2020). Working Paper.
- Hall, Bruce L., Barton H. Hamilton, Karen Richards, Karl Y. Bilimoria, Mark E. Cohen, and Clifford Y. Ko. "Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals." *Annals of Surgery* 250, no. 3 (2009): 363-376.
- Hanley, James T., David S. Ensor, Daryl D. Smith, and L. E. Sparks. "Fractional aerosol filtration efficiency of in - duct ventilation air cleaners." *Indoor Air* 4, no. 3 (1994): 169-178.
- Haynes, Alex B., Thomas G. Weiser, William R. Berry, Stuart R. Lipsitz, Abdel-Hadi S. Breizat, E. Patchen Dellinger, Teodoro Herbosa et al. "A surgical safety checklist to reduce morbidity and mortality in a global population." *New England Journal of Medicine* 360, no. 5 (2009): 491-499.
- He, Jiaxiu, Haoming Liu, and Alberto Salvo. "Severe air pollution and labor productivity: Evidence from industrial towns in China." *American Economic Journal: Applied Economics* 11, no. 1 (2019): 173-201.

- Hoek, Gerard, Rob Beelen, Kees De Hoogh, Danielle Vienneau, John Gulliver, Paul Fischer, and David Briggs. "A review of land-use regression models to assess spatial variation of outdoor air pollution." *Atmospheric Environment* 42, no. 33 (2008): 7561-7578.
- Ingraham, Angela M., Karen E. Richards, Bruce L. Hall, and Clifford Y. Ko. "Quality improvement in surgery: the American College of Surgeons national surgical quality improvement program approach." *Advances in Surgery* 44, no. 1 (2010): 251-267.
- Jerrett, Michael, Richard T. Burnett, C. Arden Pope III, Kazuhiko Ito, George Thurston, Daniel Krewski, Yuanli Shi, Eugenia Calle, and Michael Thun. "Long-term ozone exposure and mortality." *New England Journal of Medicine* 360, no. 11 (2009): 1085-1095.
- Jeuland, Marc, Subhrendu K. Pattanayak, and Randall Bluffstone. "The economics of household air pollution." *Annual Review of Resource Economics* 7, no. 1 (2015): 81-108.
- Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. "Caution, drivers! Children present: Traffic, pollution, and infant health." *Review of Economics and Statistics* 98, no. 2 (2016): 350-366.
- Landrigan, Philip J., Richard Fuller, Nereus JR Acosta, Olusoji Adeyi, Robert Arnold, Abdoulaye Bibi Baldé, Roberto Bertollini et al. "The Lancet Commission on pollution and health." *The Lancet* 391, no. 10119 (2018): 462-512.
- Li et al. 李燕, 贺征, 陈坤才, 梁颖茹, 刘鹏达, 张旭, 林云万, 邓小冰. "广州市医院洁净手术室运行及管理现状调查." *中国消毒学杂志* 6 (2014).
- Lidwell, O. M., E. J. Lowbury, W. Whyte, R. Blowers, S. J. Stanley, and D. Lowe. "Effect of ultraclean air in operating rooms on deep sepsis in the joint after total hip or knee replacement: a randomised study." *British Medical Journal* 285, no. 6334 (1982): 10-14.
- Mathieu, Daniel, Jean-Christophe Linke, and Francis Wattel. "Non-healing wounds." In *Handbook on Hyperbaric Medicine*, pp. 401-428. Springer, Dordrecht, 2006.
- Mangram, Alicia J., Teresa C. Horan, Michele L. Pearson, Leah Christine Silver, William R. Jarvis, and Hospital Infection Control Practices Advisory Committee. "Guideline for prevention of surgical site infection, 1999." *Infection Control & Hospital Epidemiology* 20, no. 4 (1999): 247-280.
- Meara, John G., Andrew JM Leather, Lars Hagander, Blake C. Alkire, Nivaldo Alonso, Emmanuel A. Ameh, Stephen W. Bickler et al. "Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development." *The Lancet* 386, no. 9993 (2015): 569-624.
- Molitor, David. "The evolution of physician practice styles: evidence from cardiologist migration." *American Economic Journal: Economic Policy* 10, no. 1 (2018): 326-56.
- Moretti, Enrico, and Matthew Neidell. "Pollution, health, and avoidance behavior evidence from the ports of Los Angeles." *Journal of Human Resources* 46, no. 1 (2011): 154-175.
- Needleman, Jack, Peter Buerhaus, V. Shane Pankratz, Cynthia L. Leibson, Susanna R. Stevens, and Marcelline Harris. "Nurse staffing and inpatient hospital mortality." *New England Journal of Medicine* 364, no. 11 (2011): 1037-1045.
- Neidell, Matthew. "Information, avoidance behavior, and health the effect of ozone on asthma hospitalizations." *Journal of Human Resources* 44, no. 2 (2009): 450-478.



- Pearse, Rupert M., Rui P. Moreno, Peter Bauer, Paolo Pelosi, Philipp Metnitz, Claudia Spies, Benoit Vallet, Jean-Louis Vincent, Andreas Hoeft, and Andrew Rhodes. "Mortality after surgery in Europe: a 7 day cohort study." *The Lancet* 380, no. 9847 (2012): 1059-1065.
- Pope III, C. Arden, Richard T. Burnett, Michelle C. Turner, Aaron Cohen, Daniel Krewski, Michael Jerrett, Susan M. Gapstur, and Michael J. Thun. "Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships." *Environmental Health Perspectives* 119, no. 11 (2011): 1616-1621.
- Pope III, C. Arden, Maureen Cropper, Jay Coggins, and Aaron Cohen. "Health benefits of air pollution abatement policy: Role of the shape of the concentration–response function." *Journal of the Air & Waste Management Association* 65, no. 5 (2015): 516-522.
- Ratti, Carlo, Dennis Frenchman, Riccardo Maria Pulselli, and Sarah Williams. "Mobile landscapes: using location data from cell phones for urban analysis." *Environment and Planning B: Planning and Design* 33, no. 5 (2006): 727-748.
- Riley, William J., Thomas E. McKone, Alvin CK Lai, and William W. Nazaroff. "Indoor particulate matter of outdoor origin: importance of size-dependent removal mechanisms." *Environmental Science & Technology* 36, no. 2 (2002): 200-207.
- Schlenker, Wolfram, and W. Reed Walker. "Airports, air pollution, and contemporaneous health." *The Review of Economic Studies* 83, no. 2 (2016): 768-809.
- Schwartz, Joel. "The distributed lag between air pollution and daily deaths." *Epidemiology* 11, no. 3 (2000): 320-326.
- Shen et al. 沈劲, 嵇萍, 周亦凌, 叶斯琪, 陈多宏, 谢敏, 汪宇, 潘月云, 许凡, 蔡日东. "广东省空气质量等级预报准确性评估." *环境监控与预警* 9, no. 3 (2017): 15-18.
- Shrime, Mark G., Stephen W. Bickler, Blake C. Alkire, and Charlie Mock. "Global burden of surgical disease: an estimation from the provider perspective." *The Lancet Global Health* 3 (2015): S8-S9.
- Sosa, Julie Ann, Helen M. Bowman, James M. Tielsch, Neil R. Powe, Toby A. Gordon, and Robert Udelsman. "The importance of surgeon experience for clinical and economic outcomes from thyroidectomy." *Annals of Surgery* 228, no. 3 (1998): 320.
- Spence, Jessica, et al. Vascular Events in Noncardiac Surgery Patients Cohort Evaluation (VISION) Study Investigators. "Association between complications and death within 30 days after noncardiac surgery." *CMAJ* 191, no. 30 (2019): E830-E837.
- Stafford, Tess M. "Indoor air quality and academic performance." *Journal of Environmental Economics and Management* 70 (2015): 34-50.
- Van Den Bos, Jill, Karan Rustagi, Travis Gray, Michael Halford, Eva Ziemkiewicz, and Jonathan Shreve. "The \$17.1 billion problem: the annual cost of measurable medical errors." *Health Affairs* 30, no. 4 (2011): 596-603.
- World Health Organization. "World Alliance for Patient Safety. WHO guidelines for safe surgery." <http://www.gawande.com/documents/WHOGuidelinesforSafeSurgery.pdf> (2008).

Xia, Fan, Jianwei Xing, Jintao Xu, and Xiaochuan Pan. "The Impact of Air Pollution on Medical Expenditure: Evidence from Beijing." (2020). Working paper.

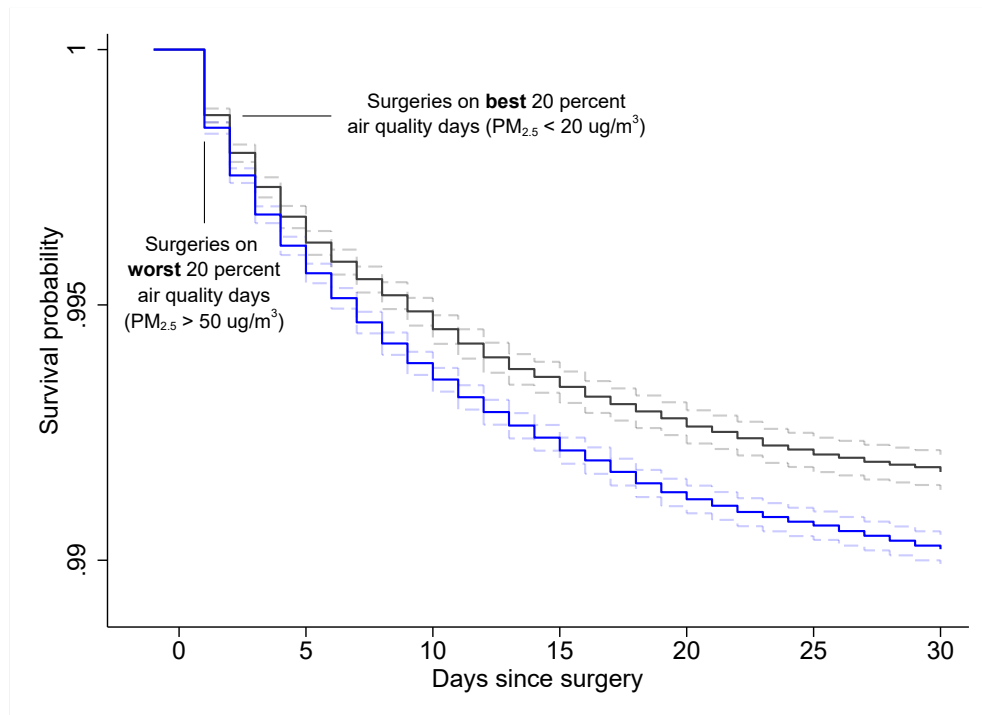
Zhang et al. 张金谱, 梁桂雄, 冯彪, 邱晓暖, 陈瑜. "广州市 2016 年空气质量预报效果评估." 广州环境科学 32, no. 2 (2017): 1-4.

Zhang et al. 张金谱, 梁桂雄, 冯彪, 邱晓暖, 陈瑜. "广州市污染季节空气质量预报效果评估及误差分析." 环境监控与预警 10, no. 2 (2018): 40-44.

Zhang, Xin, Xi Chen, and Xiaobo Zhang. "The impact of exposure to air pollution on cognitive performance." Proceedings of the National Academy of Sciences 115, no. 37 (2018): 9193-9197.

Zheng. 郑聃. "医院室内 PM<sub>2.5</sub> 的污染特点和影响因素研究." 环境卫生学杂志 4, no. 5 (2014): 504-507

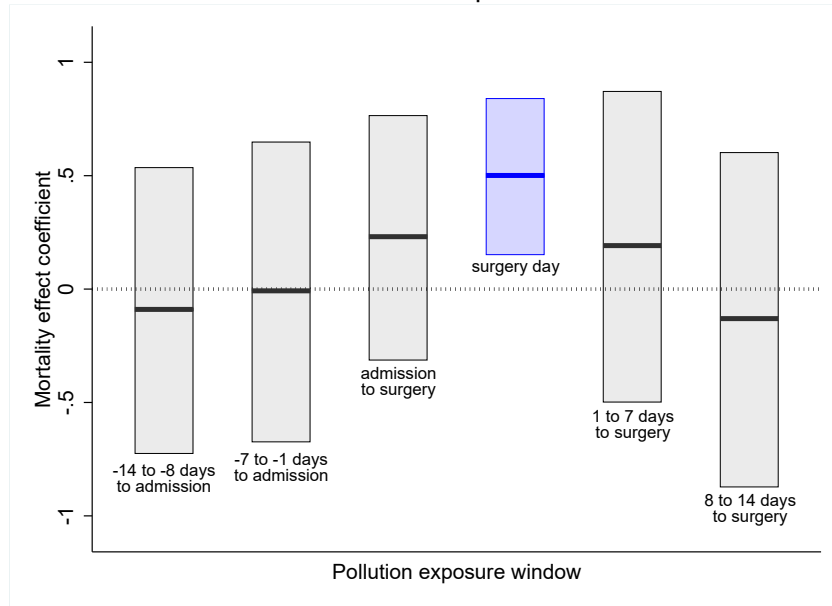
Figure 1. Patient Survival after Surgeries on High versus Low Pollution Days



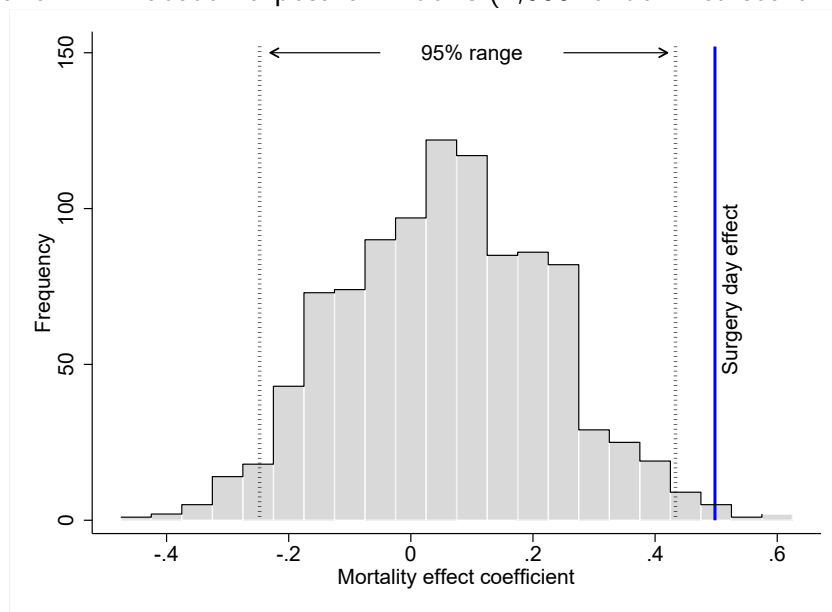
Notes: This graph reports Kaplan-Meier survival estimates and 95% confidence intervals among surgeries conducted on days with the worst quintile  $PM_{2.5}$  concentration ( $>50 \text{ ug/m}^3$ ) and days with the best quintile  $PM_{2.5}$  concentration ( $<20 \text{ ug/m}^3$ ).

Figure 2. The Importance of Surgery-Day Pollution Exposure

Panel A. Alternative exposure windows

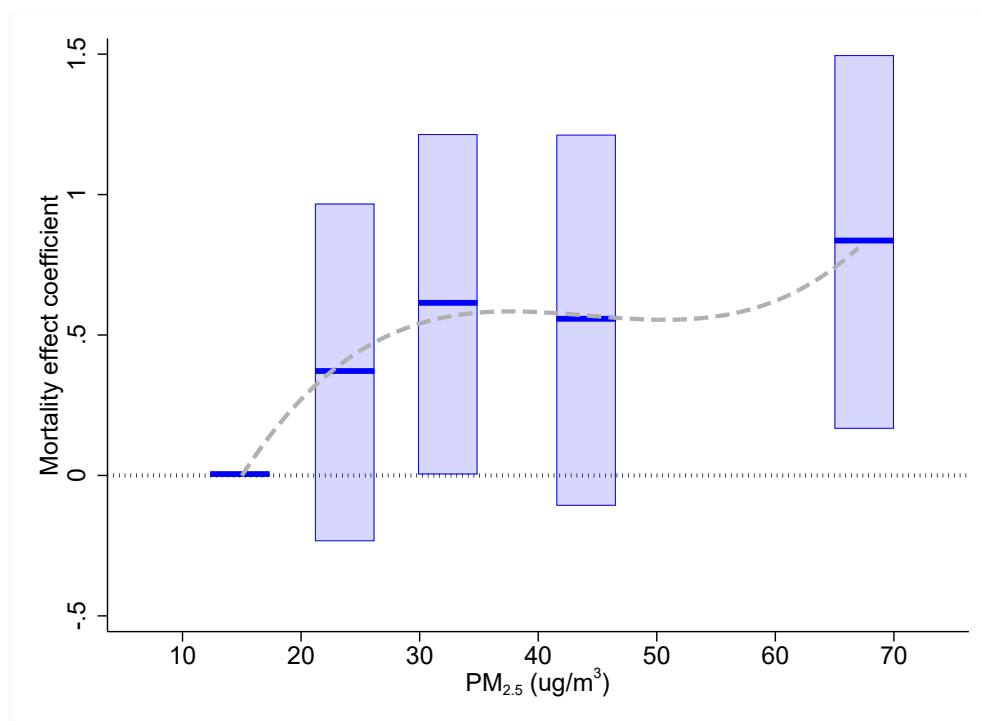


Panel B. "Placebo" exposure windows (1,000 randomized scenarios)



Notes: Panel A reports the alternative estimates of pollution effect when alternative exposure windows are used. Bars represent 95% confidence intervals. Panel B compares the observed, surgery-day pollution effect with the placebo distribution of effect sizes generated from 1,000 placebo estimation using the same data and the same regression specification but with randomly-dated surgeries.

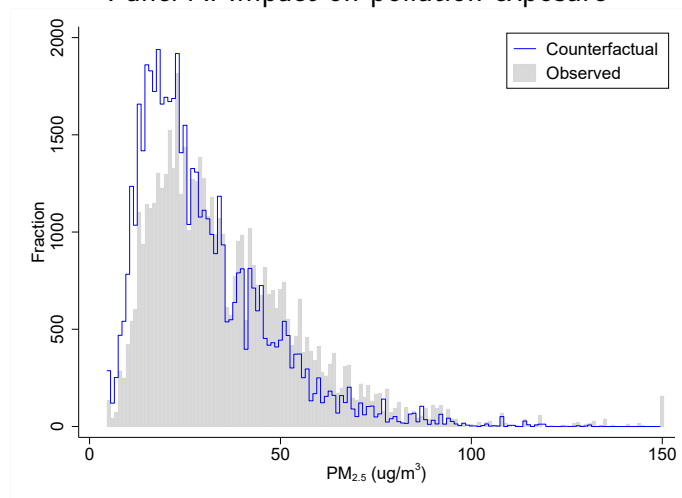
Figure 3. Concentration Response of Surgery-Day Pollution



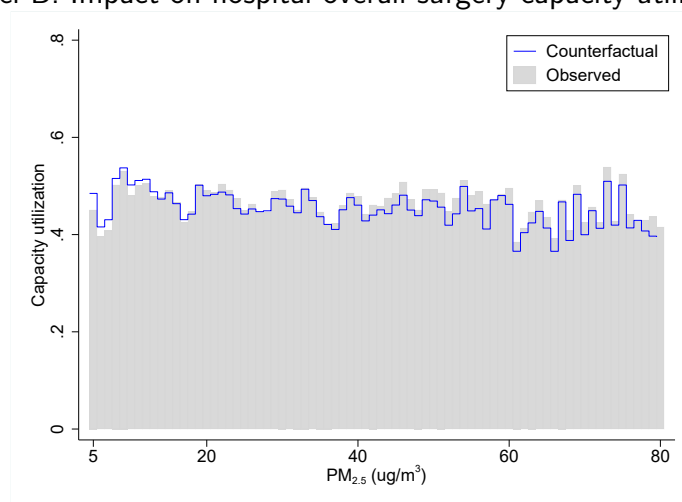
Notes: This graph reports the effect of surgery-day PM<sub>2.5</sub> on post-surgery mortality by quintile bins. Bars represent 95% confidence intervals. The first pollution quintile bin is the reference category. Dashed line shows a cubic fit.

Figure 4. Counterfactual Surgery Scheduling that Internalizes Pollution Damages

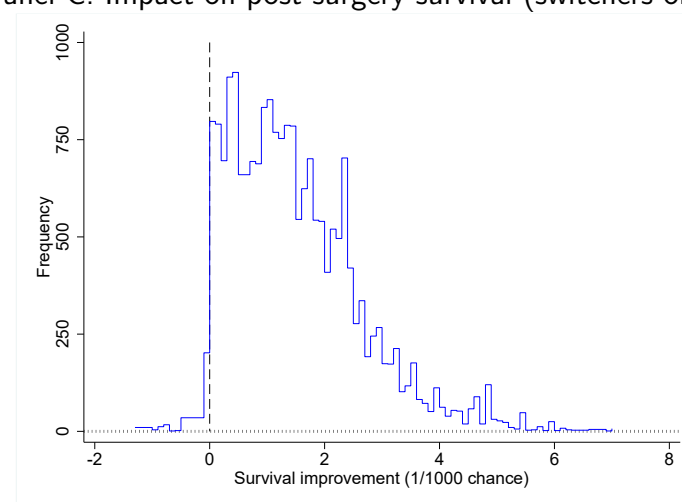
Panel A. Impact on pollution exposure



Panel B. Impact on hospital overall surgery capacity utilization



Panel C. Impact on post-surgery survival (switchers only)



*Notes:* Panel A compares observed and counterfactual surgery profile's PM<sub>2.5</sub> distribution among the high-risk patient group. Panel B compares hospitals' daily overall surgery capacity utilization rates averaged by 1-ug/m<sup>3</sup> bins. We omit few observations for days above 80 mg/m<sup>3</sup>. Panel C plots the distribution of survival improvements among "switcher" patients whose counterfactual surgery day is different from the observed day.

Table 1. Surgery-Day Pollution and Pre-surgery Patient/Surgery Characteristics

		(1)	(2)	(3)	(4)
		Indep. var.: Log PM <sub>2.5</sub> concentration			
Age	mean = 47,964	30.264 (28.337)	21.331 (29.190)	51.694* (28.049)	38.744 (29.745)
1(male)	mean = 563.3	-0.788 (0.996)	-0.664 (1.027)	-0.349 (0.887)	-0.888 (1.152)
1(married)	mean = 802.8	1.490** (0.621)	1.505** (0.623)	1.970*** (0.651)	1.640*** (0.638)
1(allergy history)	mean = 51.0	0.456 (0.421)	0.565 (0.450)	0.564 (0.541)	0.772 (0.525)
Days of delay	mean = 3,944	-6.435 (14.563)	-6.502 (13.819)	-33.683** (14.002)	-13.556 (13.005)
Number of procedures	mean = 2,050	-0.849 (2.106)	-2.682 (2.458)	-0.577 (1.822)	-0.368 (1.953)
1(general anesthesia)	mean = 491.5	-3.866 (3.953)	-3.722 (4.188)	-3.993* (1.949)	-2.657 (3.165)
1(level-1 operation - easiest)	mean = 283.7	-0.453 (3.038)	-0.477 (2.922)	2.318 (3.250)	0.237 (3.504)
1(level-2 operation)	mean = 321.5	1.325 (2.737)	0.663 (2.812)	-0.632 (2.654)	0.639 (2.689)
1(level-3 operation)	mean = 250.9	0.298 (1.021)	0.375 (0.988)	-0.562 (0.819)	-0.380 (0.916)
1(level-4 operation - hardest)	mean = 144.0	-1.170 (0.776)	-0.561 (0.660)	-1.124** (0.512)	-0.496 (0.633)
1(insurance program: City Workers)	mean = 416.1	-0.110 (2.001)	0.263 (1.884)	-0.487 (1.430)	0.206 (1.672)
1(insurance program: New Rural Cooperative)	mean = 61.7	1.289 (0.821)	0.831 (0.751)	0.742 (0.827)	0.600 (0.754)
1(insurance program: none)	mean = 295.0	0.135 (1.155)	0.498 (1.251)	0.377 (0.812)	0.320 (1.028)
FEs: diagnosis		✓	✓	✓	✓
FEs: department		✓	✓	✓	✓
FEs: procedure		✓			
FEs: hospital		✓			
FEs: year		✓	✓		✓
FEs: month		✓	✓		
FEs: day-of-week		✓	✓	✓	✓
FEs: procedure×hospital			✓	✓	
FEs: year×month				✓	
FEs: procedure×hospital×month					✓

Notes: Each cell reports a separate regression of a characteristics on surgery-day pollution. All characteristics variables are multiplied by 1,000 to increase readability. "Days of delay" is number of days between hospitalization and surgery. Numbers of observations are 1,317,099 (column 1), 1,307,307 (columns 2 and 3), and 1,236,997 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table 2. Surgery-Day Pollution and Patient Mortality

		(1)	(2)	(3)	(4)
		Indep. var.: Log PM <sub>2.5</sub> concentration			
1-day mortality	mean = 1.322	0.060 (0.074)	0.083 (0.073)	0.087 (0.074)	0.137 (0.080)
7-day mortality	mean = 5.987	0.200 (0.127)	0.248* (0.128)	0.230 (0.136)	0.231* (0.118)
14-day mortality	mean = 8.791	0.241* (0.139)	0.303** (0.145)	0.286** (0.133)	0.257* (0.135)
28-day mortality	mean = 10.28	0.339** (0.136)	0.405** (0.149)	0.456*** (0.147)	0.381** (0.176)
Overall hospital mortality	mean = 12.11	0.378** (0.176)	0.434** (0.182)	0.498*** (0.172)	0.400* (0.218)
FEs: diagnosis		✓	✓	✓	✓
FEs: department		✓	✓	✓	✓
FEs: procedure		✓			
FEs: hospital		✓			
FEs: year		✓	✓		✓
FEs: month		✓	✓		
FEs: day-of-week		✓	✓	✓	✓
FEs: procedure×hospital			✓	✓	
FEs: year×month				✓	
FEs: procedure×hospital×month					✓

Notes: Each cell reports a separate regression of a measure of post-surgery mortality on surgery-day pollution. Each mortality variable is an indicator for whether the patient died in hospital following k-day since surgery, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,099 (column 1), 1,307,307 (columns 2 and 3), and 1,236,997 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .



Table 3. Surgery-Day Pollution and Patient Mortality: Instrumental Variable Estimation

	(1)	(2)	(3)	(4)
	Dep. var.: Hospital mortality			
Panel A. IV = upwind pollution from all cities (distance <sup>-1</sup> weighted)				
Log PM <sub>2.5</sub>	0.471 (0.429)	0.462 (0.435)	0.705* (0.410)	0.562 (0.392)
Kleibergen-Paap F-stat.	211.8	226.8	169.7	236.8
Panel B. IV = upwind pollution from cities≤1,000 km (distance <sup>-1</sup> weighted)				
Log PM <sub>2.5</sub>	0.655* (0.333)	0.740** (0.319)	0.908*** (0.281)	0.841*** (0.281)
Kleibergen-Paap F-stat.	337.4	345.4	290.2	368.5
Panel C. IV = upwind pollution from all cities (distance <sup>-2</sup> weighted)				
Log PM <sub>2.5</sub>	0.532* (0.263)	0.595** (0.254)	0.806*** (0.191)	0.619*** (0.218)
Kleibergen-Paap F-stat.	454.5	464.9	391.3	497.0
Panel D. IV = upwind pollution from 119 cities ("0-stage" Lasso, distance <sup>-1</sup> weighted)				
Log PM <sub>2.5</sub>	0.448* (0.255)	0.493* (0.262)	0.656*** (0.222)	0.490** (0.235)
Kleibergen-Paap F-stat.	503.9	515.1	420.5	559.8
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

Notes: Each cell reports a separate two-stage least squares (2SLS) regression of post-surgery mortality on surgery-day pollution, using upwind pollution from cities at least 100km away from Guangzhou as the instrumental variable (IV). All IV regressions are exactly identified with one endogenous variable (log PM<sub>2.5</sub>) and one excluded instrument (upwind pollution). In panel A, the IV is logged inverse-distance-weighted average upwind PM<sub>2.5</sub> vector from all cities over 100km away. In panel B, the IV is logged inverse-distance-weighted average upwind PM<sub>2.5</sub> vector from all cities between 100-1,000km away. In panel C, the IV is logged inverse-distance-squared-weighted average upwind PM<sub>2.5</sub> vector from all cities over 100km away. In panel D, a "zero-stage" linear Lasso of Guangzhou's PM<sub>2.5</sub> on upwind PM<sub>2.5</sub> vectors are run to first select 119 contributing cities (also restricted to those over 100km away from Guangzhou). The IV is then constructed as logged inverse-distance-squared-weighted average upwind PM<sub>2.5</sub> vector from the selected cities. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are some 1.3 million. \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table 4. Surgery-Day Pollution and Treatment Differences

		(1)	(2)	(3)	(4)
		Indep. var.: Log PM <sub>2.5</sub> concentration			
Log(antimicrobial agents use)	mean = 3.381	-0.009 (0.007)	-0.006 (0.007)	0.001 (0.005)	-0.004 (0.006)
1(non-healing surgical wounds)×1,000	mean = 1.664	-0.021 (0.088)	0.005 (0.084)	0.051 (0.084)	0.076 (0.067)
1("medical error")×1,000	mean =6.138	0.189 (0.226)	0.198 (0.230)	0.256 (0.238)	0.198 (0.257)
FEs: diagnosis		✓	✓	✓	✓
FEs: department		✓	✓	✓	✓
FEs: procedure		✓			
FEs: hospital		✓			
FEs: year		✓	✓		✓
FEs: month		✓	✓		
FEs: day-of-week		✓	✓	✓	✓
FEs: procedure×hospital			✓	✓	
FEs: year×month				✓	
FEs: procedure×hospital×month					✓

Notes: Each cell reports a separate regression of a measure of treatment/performance on surgery-day pollution. The indicator variables are multiplied by 1,000 to increase readability. "Medical errors" is a proxy built from patient injury at discharge following Van Den Bos et al. (2011) and David et al. (2013). All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,099 (column 1), 1,307,307 (columns 2 and 3), and 1,236,997 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table 5. Surgery-Day Pollution and Patient Mortality: Multiple Pollutants Model

	(1)	(2)	(3)	(4)	(5)
	Dep. var.: Hospital mortality				
Log PM <sub>2.5</sub>	0.545*** (0.163)	0.887*** (0.264)	0.470** (0.194)	0.405** (0.167)	0.906*** (0.300)
Log O <sub>3</sub>	-0.166 (0.125)				-0.213 (0.137)
Log NO <sub>2</sub>		-0.843* (0.448)			-1.017** (0.476)
Log SO <sub>2</sub>			0.010 (0.198)		0.135 (0.191)
Log CO				0.311 (0.263)	0.393 (0.243)
FEs: diagnosis	✓	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓	✓
FEs: day-of-week	✓	✓	✓	✓	✓
FEs: procedure×hospital	✓	✓	✓	✓	✓
FEs: year×month	✓	✓	✓	✓	✓

Notes: Each column is a separate regression. Mortality variable is an indicator for whether the patient died in hospital, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Number of observations are 1,307,307. \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table 6. Surgery-Day Pollution and Patient Mortality: High-Risk Patients

		(1)	(2)	(3)	(4)
		Dep. var.: Hospital mortality			
Log PM <sub>2.5</sub> × 1(high-risk patients)	mean mortality = 38.35	2.769** (1.263)	2.993** (1.173)	3.044** (1.156)	2.837** (1.233)
Log PM <sub>2.5</sub> × 1(other patients)	mean mortality = 8.74	0.134 (0.191)	0.176 (0.205)	0.224 (0.207)	0.151 (0.217)
FEs: diagnosis		✓	✓	✓	✓
FEs: department		✓	✓	✓	✓
FEs: procedure		✓			
FEs: hospital		✓			
FEs: year		✓	✓		✓
FEs: month		✓	✓		
FEs: day-of-week		✓	✓	✓	✓
FEs: procedure×hospital			✓	✓	
FEs: year×month				✓	
FEs: procedure×hospital×month					✓

Notes: Each column reports a separate regression that allows the effect of PM<sub>2.5</sub> on hospital mortality to vary by patient groups. The outcome variable is an indicator for whether the patient died in hospital following the surgery, multiplied by 1,000 to increase readability. “High-risk” group consists of respiratory and neoplasm patients aged over 60. “Mean mortality” shows average hospital mortality rate among the two groups of patients. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,099 (column 1), 1,307,307 (columns 2 and 3), and 1,236,997 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

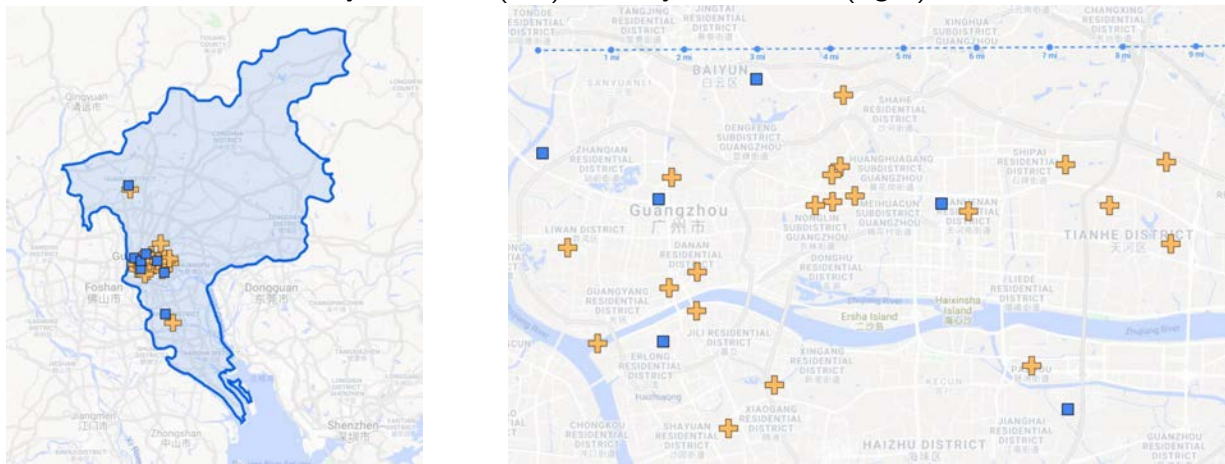
## **Additional Figures and Tables**

Figure A.1. Study Location

Panel A. City location



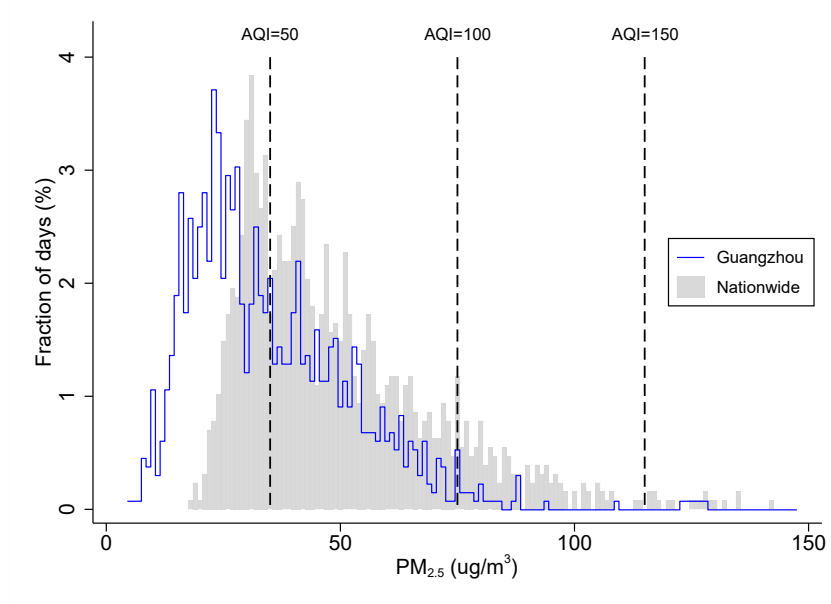
Panel B. Hospital and pollution monitor locations  
citywide view (left) and city center view (right)



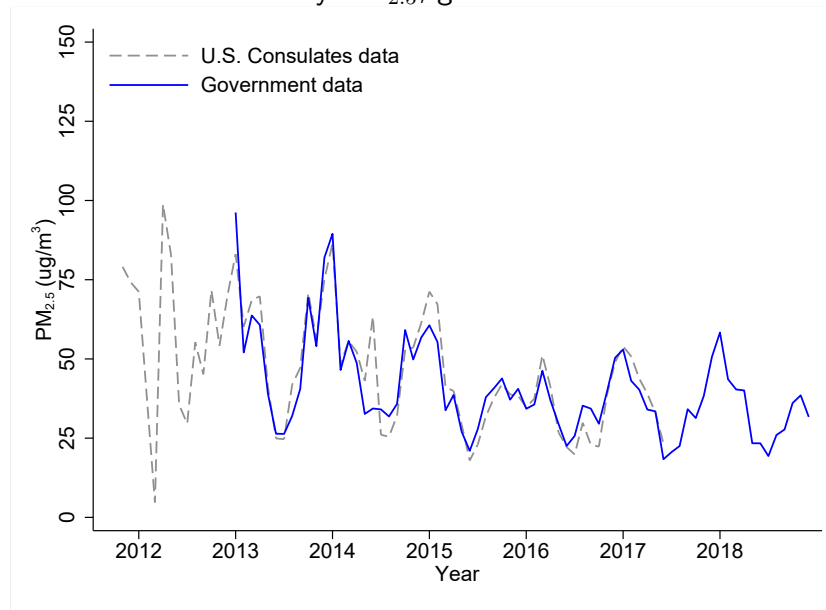
Notes: Panel A highlights the province of Guangdong (light blue) and the city of Guangzhou (dark blue). Panel B shows the location of hospitals (crosses) and air pollution monitors (squares) in our study sample.

Figure A.2. PM<sub>2.5</sub> Pollution Monitoring Data

Panel A. Distribution of daily PM<sub>2.5</sub>, Guangzhou versus nationwide



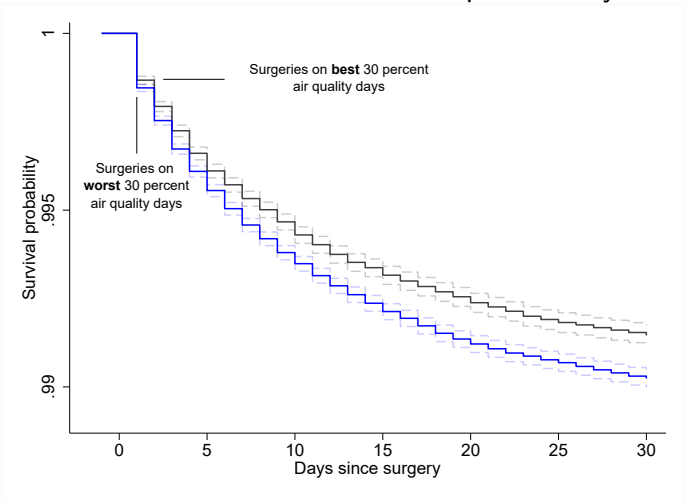
Panel B. Time series of monthly PM<sub>2.5</sub>, government versus U.S. consulate data



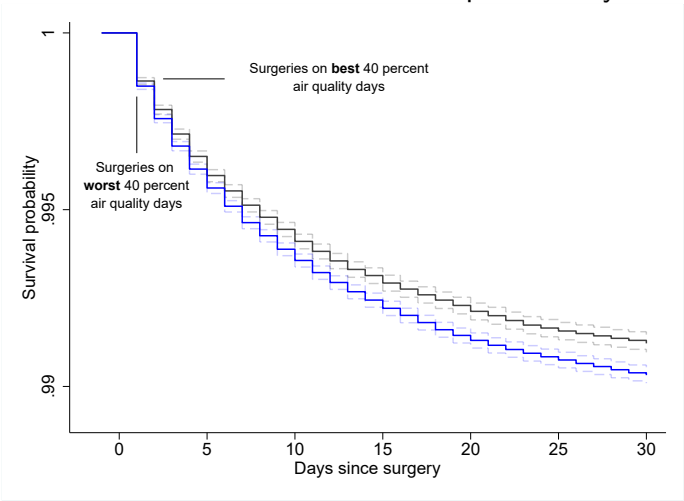
Notes: Panel A compares distribution of daily PM<sub>2.5</sub> in the city of Guangzhou and nationwide. Vertical dashed lines correspond to Air Quality Cutoffs for Good, Moderate, Unhealthy for Sensitive Groups, and Unhealthy. Panel B compares monthly PM<sub>2.5</sub> using data from the Chinese government and independent monitoring data from the U.S. consulate at Guangzhou.

Figure A.3. Patient Survival after Surgeries on High versus Low Pollution Days: Alternative Cutoffs

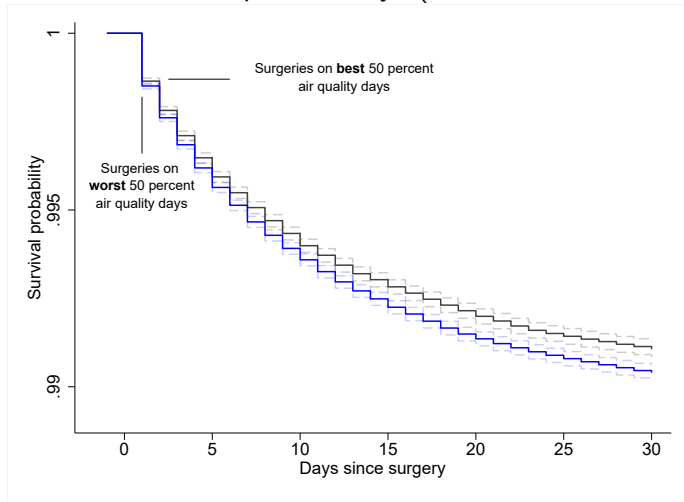
Panel A. Best versus worst 30 percent days



Panel B. Best versus worst 40 percent days



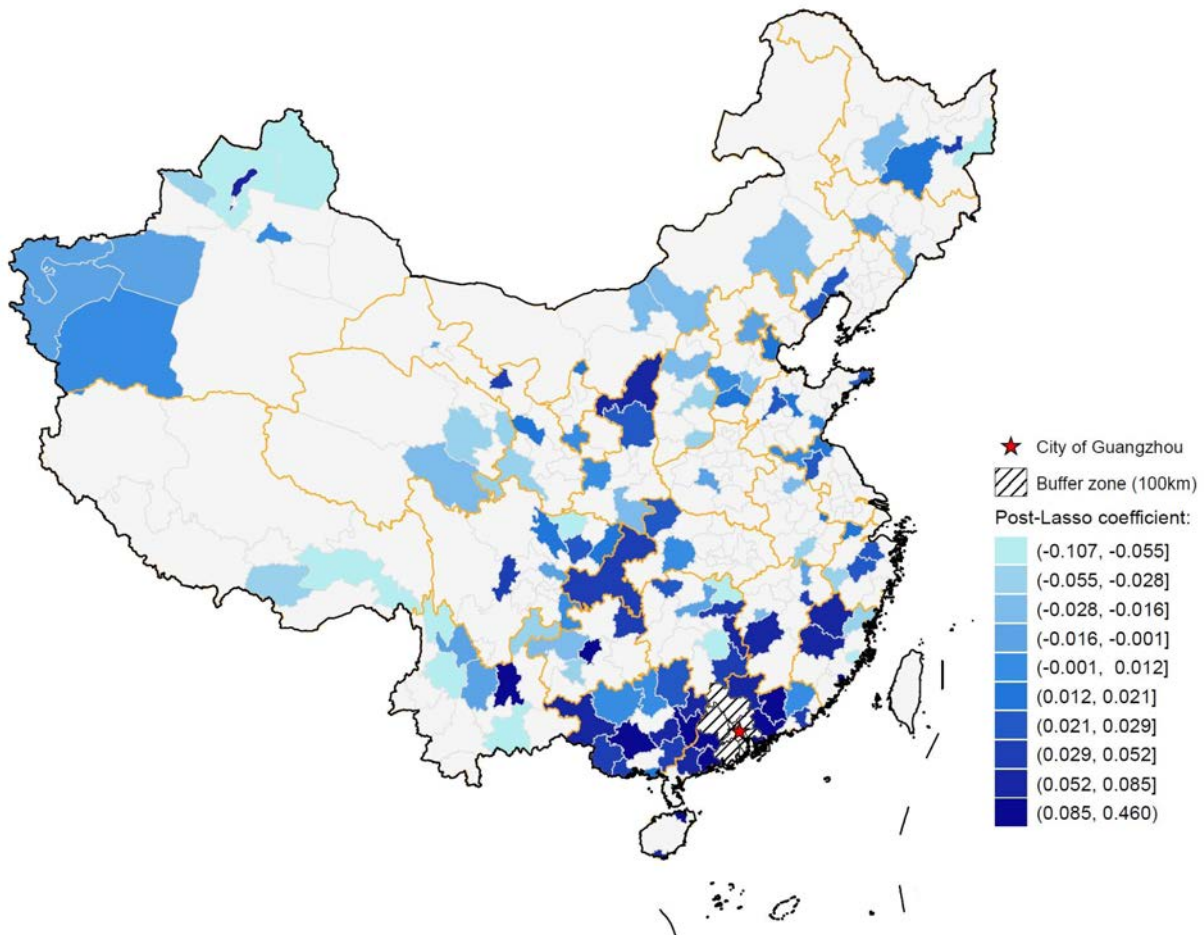
Panel C. Best versus worst 50 percent days (i.e., above versus below median)



Notes: This graph reports robustness of the Kaplan-Meier survival estimates with respect to alternative high versus low pollution day cutoffs.



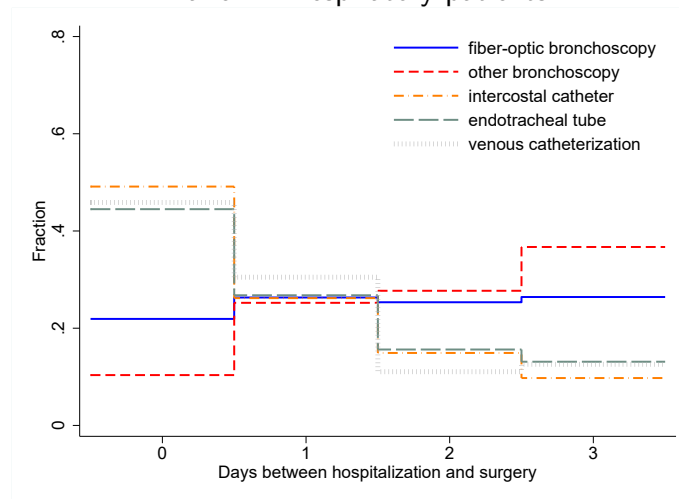
Figure A.4. Upwind Cities Selected by the “Zero-Stage” Lasso Regression



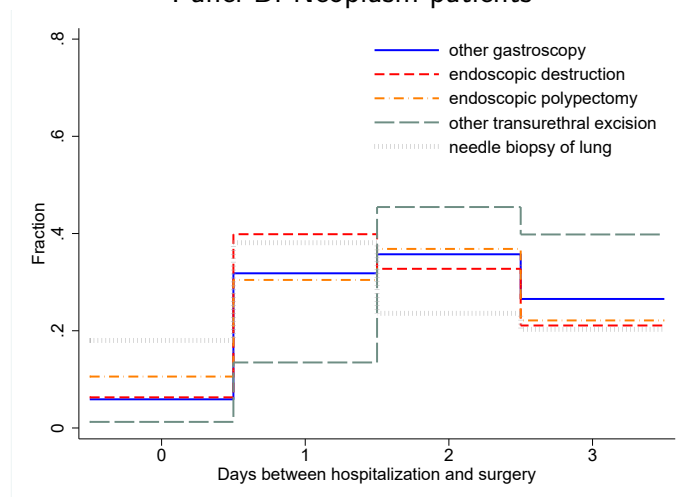
*Notes:* This map highlights 119 cities selected by a “zero-stage” Lasso regression of Guangzhou’s daily  $PM_{2.5}$  on all other 305 cities’ upwind component vector  $PM_{2.5}$ . See text for details.

Figure A.5. Distributions of Admission-to-Surgery Delays among High-Risk Patients

Panel A. Respiratory patients



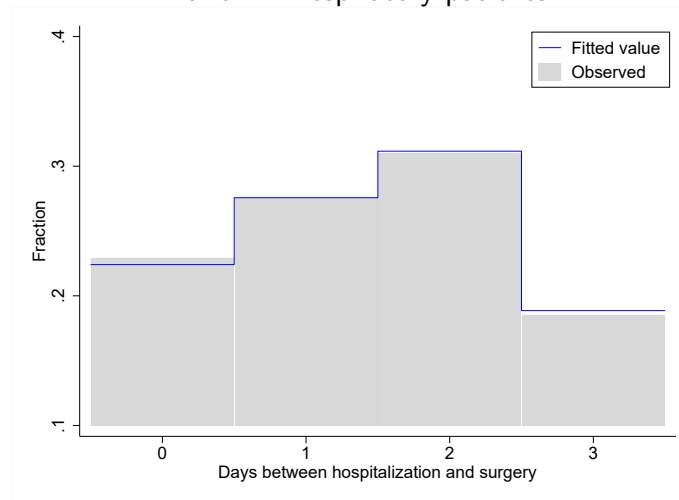
Panel B. Neoplasm patients



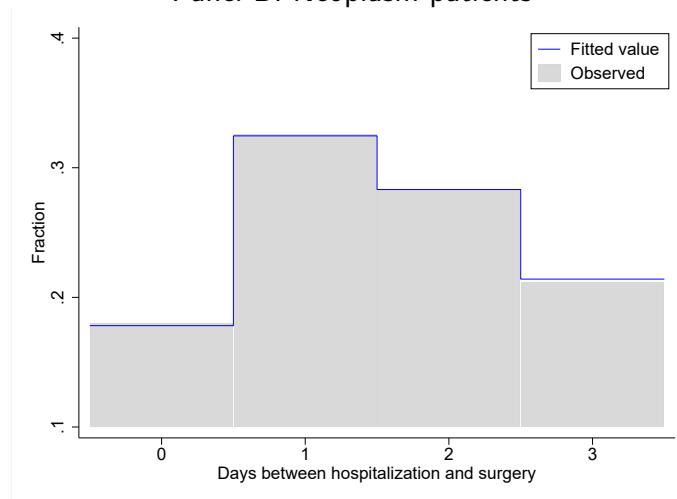
Notes: Samples restrict to respiratory patients (panel A) and neoplasm patients (panel B) aged over 60 and those scheduled to receive surgeries within three days of hospital admission. For each diagnosis group, the graph shows distributions for the five most common procedures.

Figure A.6. Structural Models of Surgery Scheduling: Fitted versus Observed Schedules

Panel A. Respiratory patients



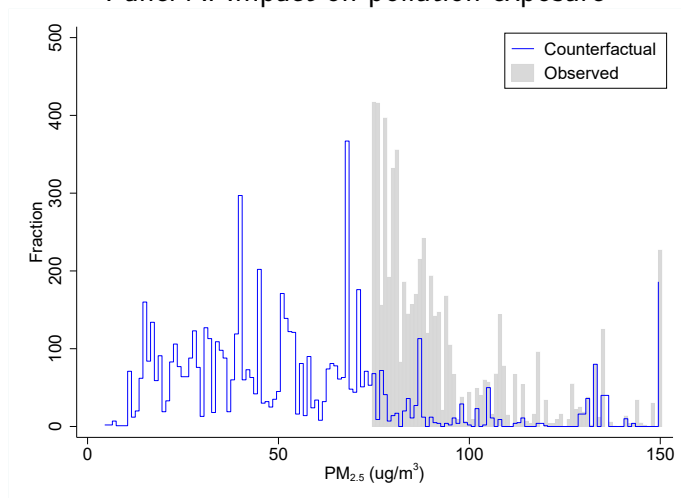
Panel B. Neoplasm patients



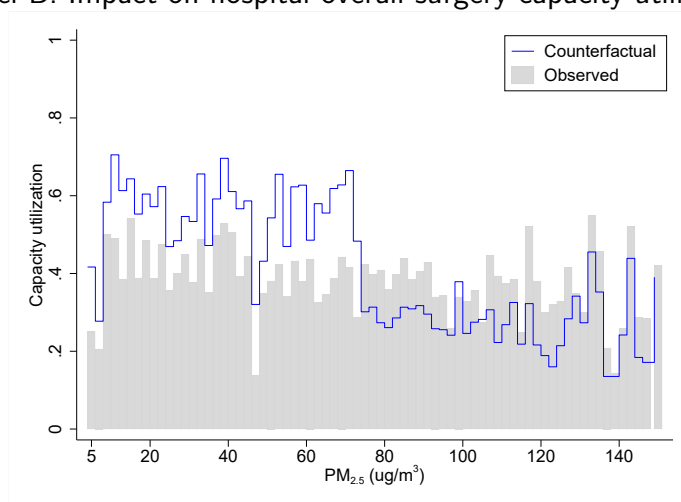
Notes: Samples restrict to patients aged over 60 and those scheduled to receive surgeries within three days of hospital admission.

Figure A.7. Counterfactual Surgery Scheduling that Internalizes Pollution Damages: Practical Example

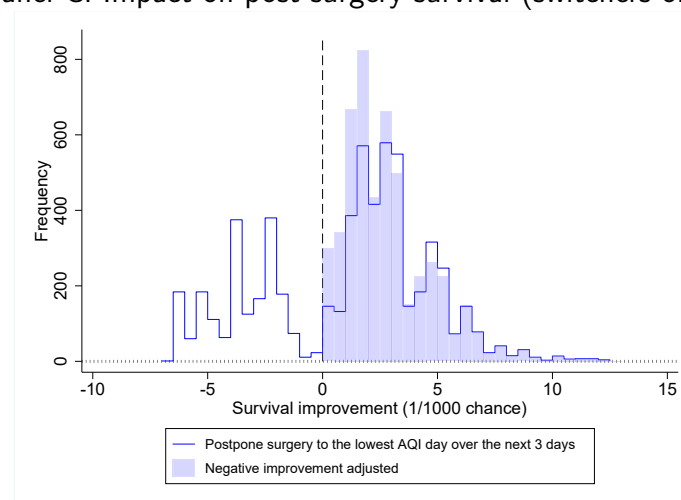
Panel A. Impact on pollution exposure



Panel B. Impact on hospital overall surgery capacity utilization



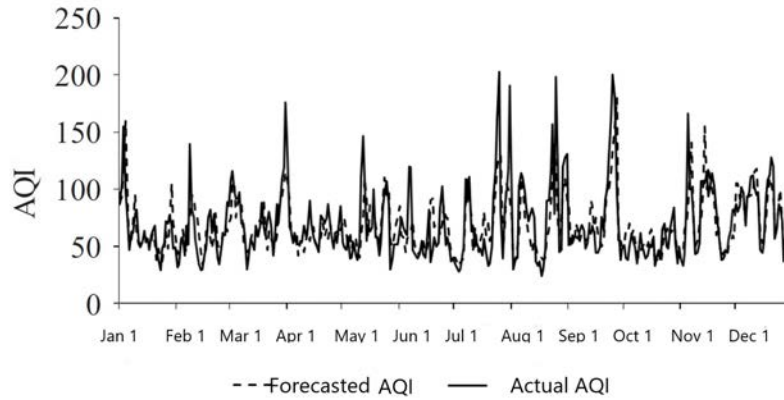
Panel C. Impact on post-surgery survival (switchers only)



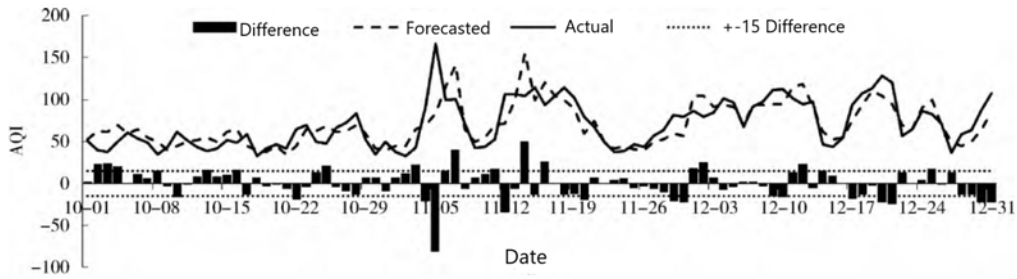
Notes: Panel A compares observed and counterfactual surgery profile's  $PM_{2.5}$  distribution among the high-risk patient group. Panel B compares hospitals' daily overall surgery capacity utilization rates averaged by  $1\text{-ug}/m^3$  bins. We omit few observations for days above  $80\text{ mg}/m^3$ . Panel C plots the distribution of survival improvements among "switcher" patients whose counterfactual surgery day is different from the observed day.

Figure A.8. Forecasted versus Observed Air Quality Index in Guangzhou (Zhang et al., 2017, 2018)

Panel A. Year 2016

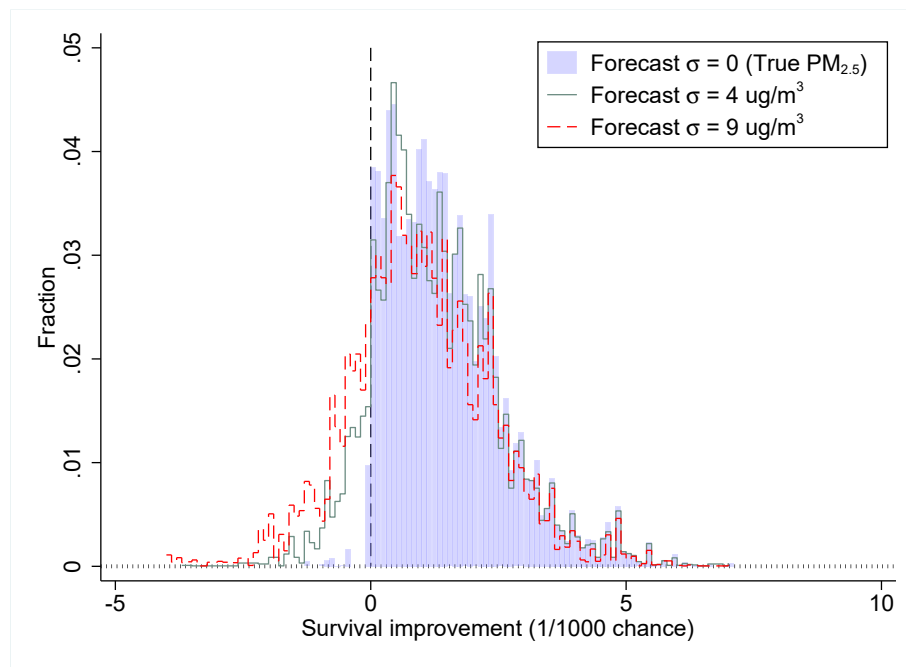


Panel B. October to December, 2016



Notes: These figures are taken from Zhang et al.(2017) and Zhang et al. (2018) respectively, with superimposed English translation. Panel A reports 24-hour-ahead forecasted (dashed line) and observed (solid line) Air Quality Index throughout the year of 2016. Panel B zooms in to October to December of 2016. Bars on panel B represent differences between observed and forecasted Air Quality Index values.

Figure A.9. Pollution Forecasting Accuracy and Patient Survival Improvement



*Notes:* This graph plots the distribution of survival improvements among “switcher” patients whose counterfactual surgery day is different from the observed day. Blue histogram repeats Panel C of Figure 4. Green and red histograms show survival improvement when true  $PM_{2.5}$  values are infused with  $N(0, \sigma = 4 \text{ ug/m}^3)$  and  $N(0, \sigma = 9 \text{ ug/m}^3)$  noise.

Table A.1. Surgery-Day Pollution and Patient Mortality: Robustness

	(1)	(2)	(3)	(4)
	Indep. var.: Log PM <sub>2.5</sub> concentration			
No patient/weather controls	0.311* (0.155)	0.354** (0.159)	0.443** (0.162)	0.368** (0.176)
Add low-risk surgeries	0.226* (0.110)	0.291** (0.115)	0.328** (0.119)	0.228* (0.126)
Drop admission-day surgeries	0.284 (0.168)	0.325* (0.164)	0.361** (0.141)	0.283 (0.195)
Drop admission-day & next-two-day surgeries	0.341 (0.251)	0.430 (0.269)	0.447* (0.229)	0.368 (0.339)
Control for daily hospital admission volume	0.352** (0.168)	0.357** (0.168)	0.362** (0.166)	0.366** (0.168)
Control for daily inpatient surgery volume	0.464** (0.167)	0.465** (0.167)	0.465** (0.167)	0.470** (0.166)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

Notes: Each cell reports a separate regression of a measure of post-surgery mortality on surgery-day pollution. Each mortality variable is an indicator for whether the patient died in hospital following k-day since surgery, multiplied by 1,000 to increase readability. Unless noted otherwise, all regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are (from top to bottom rows) some 1.3 million, 2.1 million, 1.1 million, 0.6 million, 1.3 million, and 1.3 million. \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table A.2. Surgery-Day Pollution and Patient Mortality: Distributed Leads &amp; Lags Models

	(1)	(2)	(3)	(4)
	Dep. var.: Hospital mortality			
Surgery-day PM <sub>2.5</sub> (t)	0.499*** (0.170)	0.499*** (0.173)	0.514*** (0.178)	0.469** (0.182)
Lag PM <sub>2.5</sub> (t+1 to t+3)	-0.066 (0.242)	-0.074 (0.274)	-0.087 (0.238)	-0.115 (0.283)
Lag PM <sub>2.5</sub> (t+4 to t+6)		0.112 (0.291)		0.098 (0.283)
Lead PM <sub>2.5</sub> (t-1 to t-3)			-0.032 (0.284)	0.053 (0.275)
Lead PM <sub>2.5</sub> (t-4 to t-6)				-0.234 (0.242)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital	✓	✓	✓	✓
FEs: year×month	✓	✓	✓	✓

Notes: “Lag PM<sub>2.5</sub>” is pollution observed in days after the surgery. “Lead PM<sub>2.5</sub>” is pollution observed in days before the surgery. Each column is a separate regression. Mortality variable is an indicator for whether the patient died in hospital, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Number of observations are is some 1.3 million depending on leads/lags included. \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .



Table A.3. Surgery-Day Pollution and Patient Mortality: Surgery Complexity Level

	(1)	(2)	(3)	(4)
	Dep. var.: Hospital mortality			
Log PM <sub>2.5</sub> × 1(level-1 operation - easiest)	0.814** (0.387)	0.536* (0.274)	0.601** (0.279)	0.486 (0.288)
Log PM <sub>2.5</sub> × 1(level-2 operation)	0.168 (0.219)	0.297 (0.216)	0.356 (0.217)	0.236 (0.251)
Log PM <sub>2.5</sub> × 1(level-3 operation)	0.235 (0.152)	0.372* (0.191)	0.434** (0.180)	0.324 (0.238)
Log PM <sub>2.5</sub> × 1(level-4 operation - hardest)	0.246 (0.195)	0.639*** (0.220)	0.704*** (0.187)	0.731** (0.324)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

Notes: Each column reports a separate regression that allows the effect of PM<sub>2.5</sub> on hospital mortality to vary by the surgery's reported complexity level. The outcome variable is an indicator for whether the patient died in hospital following the surgery, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,105 (column 1), 1,307,313 (columns 2 and 3), and 1,237,003 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table A.4. Surgery-Day Pollution and Surgical Site Infection

	(1)	(2)	(3)	(4)
Panel A. Average SSI effect				
Log PM <sub>2.5</sub>	0.076 (0.046)	0.076* (0.042)	0.102** (0.047)	0.100* (0.050)
Panel B. SSI Effects by pre-existing conditions				
Log PM <sub>2.5</sub> × 1(hypertensive or diabetic patients)	0.096 (0.072)	0.111 (0.067)	0.137* (0.067)	0.138 (0.091)
Log PM <sub>2.5</sub> × 1(others)	0.070 (0.053)	0.065 (0.046)	0.091* (0.052)	0.088* (0.049)
Panel C. Mortality Effects by pre-existing conditions				
Log PM <sub>2.5</sub> × 1(hypertensive or diabetic patients)	1.270*** (0.448)	1.268*** (0.380)	1.332*** (0.387)	1.400*** (0.458)
Log PM <sub>2.5</sub> × 1(others)	0.103 (0.145)	0.174 (0.169)	0.238 (0.152)	0.081 (0.183)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

Notes: Each column reports a separate regression. In panel A, the outcome variable is an indicator for whether the patient had surgical site infection (SSI), multiplied by 1,000 to increase readability. The mean of the dependent variable is 0.563 incidents per 1,000 patients. Panel B allows the effect of PM<sub>2.5</sub> on SSI to vary by patient's pre-existing hypertension or diabetes conditions. Panel C repeats the same regressions in panel B, but use hospital mortality as the outcome variable. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,407 (column 1), 1,307,616 (columns 2 and 3), and 1,237,297 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table A.5. Surgery-Day Pollution and Patient Mortality: Diagnosis Categories

	(1)	(2)	(3)	(4)
	Dep. var.: Hospital mortality			
Log PM <sub>2.5</sub> × 1(infectious)	3.389* (1.963)	3.547* (1.945)	3.593* (1.942)	2.985 (2.228)
Log PM <sub>2.5</sub> × 1(neoplasms)	0.814 (0.491)	0.908* (0.477)	0.962** (0.452)	0.913 (0.554)
Log PM <sub>2.5</sub> × 1(blood)	-2.107 (1.265)	-2.335* (1.236)	-2.310* (1.233)	-2.046 (1.464)
Log PM <sub>2.5</sub> × 1(metabolic)	-0.773 (1.641)	-0.570 (1.639)	-0.523 (1.624)	-0.565 (2.012)
Log PM <sub>2.5</sub> × 1(mental)	-0.756 (0.800)	-0.758 (0.775)	-0.676 (0.768)	-0.681 (0.813)
Log PM <sub>2.5</sub> × 1(nervous)	1.014 (1.112)	0.488 (1.079)	0.537 (1.091)	1.177 (1.256)
Log PM <sub>2.5</sub> × 1(eye)	1.204 (2.581)	1.071 (2.457)	1.109 (2.467)	1.878 (2.306)
Log PM <sub>2.5</sub> × 1(ear)	-0.440 (0.620)	0.352 (0.536)	0.480 (0.556)	0.415 (0.913)
Log PM <sub>2.5</sub> × 1(circulatory)	0.320 (0.591)	0.443 (0.584)	0.496 (0.573)	0.430 (0.686)
Log PM <sub>2.5</sub> × 1(respiratory)	3.883* (2.014)	4.059** (1.812)	4.098** (1.812)	3.889* (1.964)
Log PM <sub>2.5</sub> × 1(digestive)	0.137 (0.299)	0.231 (0.274)	0.271 (0.277)	0.306 (0.399)
Log PM <sub>2.5</sub> × 1(skin)	0.435 (0.812)	0.703 (0.866)	0.755 (0.891)	0.774 (1.018)
Log PM <sub>2.5</sub> × 1(musculoskeletal)	-0.466 (0.390)	-0.324 (0.345)	-0.283 (0.346)	-0.465 (0.457)
Log PM <sub>2.5</sub> × 1(genitourinary)	-0.646* (0.337)	-0.578 (0.344)	-0.522 (0.349)	-0.548 (0.391)
Log PM <sub>2.5</sub> × 1(childbirth)	-0.520 (1.578)	-0.680 (2.043)	-0.606 (2.046)	-3.888 (3.007)
Log PM <sub>2.5</sub> × 1(perinatal)	1.630 (3.371)	3.090 (3.382)	3.136 (3.411)	0.043 (5.037)
Log PM <sub>2.5</sub> × 1(congenital)	-0.139 (0.373)	-0.051 (0.346)	0.009 (0.384)	0.279 (0.529)
Log PM <sub>2.5</sub> × 1(laboratory)	-0.086 (2.609)	0.074 (2.692)	0.097 (2.716)	1.377 (3.030)
Log PM <sub>2.5</sub> × 1(injury)	0.211 (0.719)	0.368 (0.781)	0.419 (0.770)	-0.381 (1.199)
Log PM <sub>2.5</sub> × 1(health services)	-0.028 (0.298)	-0.335 (0.385)	-0.288 (0.386)	-0.388 (0.411)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

Notes: Each column reports a separate regression that allows the effect of PM<sub>2.5</sub> on hospital mortality to vary by patient groups. The outcome variable is an indicator for whether the patient died in hospital following the surgery, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table A.6. Surgery-Day Pollution and Patient Mortality: Elderly Patients

	(1)	(2)	(3)	(4)
	Dep. var.: Hospital mortality			
Log PM <sub>2.5</sub> × 1(< 60 years old)	0.048 (0.164)	0.159 (0.184)	0.223 (0.181)	0.149 (0.194)
Log PM <sub>2.5</sub> × 1(≥ 60 years old)	1.024** (0.381)	0.971*** (0.313)	1.032*** (0.303)	0.881** (0.379)
FEs: diagnosis	✓	✓	✓	✓
FEs: department	✓	✓	✓	✓
FEs: procedure	✓			
FEs: hospital	✓			
FEs: year	✓	✓		✓
FEs: month	✓	✓		
FEs: day-of-week	✓	✓	✓	✓
FEs: procedure×hospital		✓	✓	
FEs: year×month			✓	
FEs: procedure×hospital×month				✓

*Notes:* Each column reports a separate regression that allows the effect of PM<sub>2.5</sub> on hospital mortality to vary by patient groups. The outcome variable is an indicator for whether the patient died in hospital following the surgery, multiplied by 1,000 to increase readability. All regressions control additionally for individual characteristics including age in 5-year bins, gender indicator, marital status indicator, and an indicator for any history of allergy, and weather controls including daily average temperature in 10-degree Celsius bins, precipitation, and precipitation squared. Numbers of observations are 1,317,099 (column 1), 1,307,307 (columns 2 and 3), and 1,236,997 (column 4). \*:  $p < 0.10$ ; \*\*:  $p < 0.05$ ; \*\*\*:  $p < 0.01$ .

Table A.7. Structural Parameter Estimates

	(1)	(2)	(3)	(4)	(5)
	$\alpha$	$d_0$	$d_1$	$d_2$	$\varphi_{\text{workday}}$
Respiratory patients	0.2301 (0.0141)	0.0927 (0.0164)	0.3315 (0.0155)	0.4991 (0.0147)	-0.0600 (0.0778)
Neoplasm patients	0.3013 (0.0079)	-0.3735 (0.0075)	0.2826 (0.0065)	0.2694 (0.0065)	0.2556 (0.0433)

*Notes:* This table reports parameter estimates in our structural analysis. Sample restricts to patients aged over 60. Patients are separated by main diagnosis categories. Standard errors are in parentheses.