Urban Forests: Environmental Health Values and Risks*

Jianwei Xing

Zhiren Hu

Fan Xia

Jintao Xu

Eric Zou

August 2023

Abstract

Forests accompany the cities we build. There are an estimated 5.5 billion urban trees in the United States. Globally, about 25 percent of urban land is covered by tree canopy. This study examines urban forests as a policy tool for air pollution mitigation. We study an afforestation program in the city of Beijing, which planted a total of 2 million *mu* of greenery – roughly the size of Los Angeles – across the city over a decade. We conduct a remote-sensing audit of the program, finding that it contributes to a substantial greening up of the city. This causes significant downwind air quality improvement, reducing average $PM_{2.5}$ concentration at city population hubs by 4.2 percent. Rapid vegetation growth, however, led to a 7.4 percent increase in pollen exposure. Analysis of medical claims data shows aeroallergens triggered emergency room visits, mirroring well-documented industrial pollution effects though less severe. We offer insight on managing urban forests' health risks, identifying harmful pollen species and susceptible population subgroups.

Keywords: urban forests, air pollution, allergy, sustainable development

JEL Codes: I18, Q23, Q53, Q56, R11

* Xing: China Center for Economic Research, National School of Development, Peking University (email: jerryxing@nsd.pku.edu.cn); Hu: Dyson School of Applied Economics and Management, Cornell University (email: zh443@cornell.edu); Xia: State Key Laboratory of Pollution Control & Resource Reuse, School of Environment, Nanjing University (email: xiafan@nju.edu.cn); Xu: China Center for Economic Research, National School of Development, Peking University (email: xujt@pku.edu.cn); Zou: Ross School of Business, University of Michigan and NBER (email: ericzou@umich.edu). We thank seminar participants at the Renmin University of China for helpful comments. All errors are our own.

1. Introduction

Urban pollution control is a universal challenge for cities worldwide. Cities have to deal with pollution from various sources, such as transportation, construction, waste, energy generation, and industrial activities (Glaeser and Kahn, 2010; Currie and Walker, 2011; Zheng and Kahn, 2013; Gendron-Carrier et al., 2022). Traditional pollution policies focus on reducing emissions from these individual pollution *sources*. Another approach is to mitigate pollution exposure at *receptors* – the people. In this paper, we study urban forests, one of the most prevalent policy instruments in that vein.

Forests accompany the cities we build. There are an estimated 5.5 billion urban trees in the United States. Worldwide, urban tree cover – the proportion of the total urban land area covered by tree canopy – is estimated to be around 25 percent (Nowak and Greenfield, 2018; 2020). Urban forests, and other forms of greenery (a term we will use interchangeably henceforth) such as shrubs and herbaceous plants, are able to absorb hazardous chemical pollutants and filter out particulate pollution as they travel through the air (Nowak, 1994). Aside from environmental benefits, urban forests also enhance landscapes, property values, and employment opportunities, making it often a politically viable policy instrument.

Despite the relevance of urban forests for residents' economic and environmental life, their effects are sparsely studied in the economics literature. We study three subjects in this paper. First, we document a substantial greening up of a mega city and examine the contribution by a government-led mass afforestation policy. Second, we quantify the impact of urban forests on downwind air quality improvement using a quasi-experimental research design. Third, we investigate a commonly wondered negative externality – pollen emissions resulting from urban afforestation – and use medical claims data to provide some of the first direct estimates of the health burdens associated with these aeroallergens. We limit this research to environmental and health effects as our first-order focus, though we expect these effects to be associated with a broader set of economic, environment, and health outcomes.

Our study setting is Beijing, the capital city of China, which has a population of over 20 million people. In response to the central government's recent focus on combating pollution, Beijing implemented a large-scale urban afforestation effort called the Million *Mu* Project (henceforth MMP), which began in 2012 and involves planting trees and shrubs throughout the city. In a decade, the project has added 2 million *mu* of greenery – equivalent to 515 square miles, roughly the size of Los Angeles – spread throughout the city, with many forest patches situated near the city's most densely populated areas.

We begin with a remote-sensing audit of Beijing's greening progress. Using a satellite measurement of vegetation growth, we document a 30 percent increase in greenery within the city between 2001 and 2020. This trend accelerated notably after 2012, with numerous densely populated areas experiencing annual vegetation index growth of over 10 percent. By cross-referencing government maps of MMP planting sites with high-resolution satellite data, we demonstrate that much of this green growth in hotspot areas is indeed driven by the MMP program. The vegetation growth is highly relevant to city residents, as many planting sites are located right on the outskirts of densely populated areas. We estimate that over 1.3 million people (7 percent of the city population) live in areas directly influenced by MMP planting, and the vast majority of the city's population resides within a few kilometers of these sites – which, as we will show later, are significantly impacted by the environmental effects of the new greenery through spatial spillover.

Does a greener city result in cleaner air? We first extend upon the spatial analysis described above, this time using a satellite pollution measurement as the outcome variable. We document sharp improvements in air quality within MMP's planting areas, indicating a distinct gradient that was not observed in the same locations prior to the implementation of the policy.

Focusing solely on the MMP planting areas overlooks important spillover benefits. Forests serve as natural absorbers and filters that may reduce pollution that would otherwise reach population centers. To assess the broader effect on the population, we use pollution monitoring data from *in-situ* real-time air quality monitors installed by the government near population hubs, combined with prevailing wind and satellite vegetation data, to estimate the impact of urban forests on air quality experienced at these densely-populated locations. Exploiting plausible day-to-day variations in wind directions, we demonstrate that upwind vegetation reduces particulate air pollution. Forests have a large effect on particulate pollution reduction, and also reduces trace gas pollutants such as sulfur dioxides. Placebo tests show that none of these effects are observed when using vegetation conditions at *downwind* locations as the source of shocks instead.

Next, we examine the impact of urban forests on ambient pollen concentration. We use *in-situ* monitoring data from pollen monitors, also located at population hubs, that capture daily pollen counts. Using the same wind-direction-based estimation method described earlier, we demonstrate that upwind vegetation leads to a sharp increase in pollen spikes. The impact of vegetation on pollen levels is sizable, if not larger and more statistically precise in terms of elasticity, than the pollution reduction effect. For example, given an upwind vegetation shock of the same size, the response of the pollen spike in log scale is over twice as large as that of the reduction in fine particulate matter ($PM_{2.5}$).

Our estimates suggest that, over the first eight-year period since MMP started, the project reduces the average population $PM_{2.5}$ exposure in the city by 4.2 percent (about 2.9 ug/m^3 from 2012 baseline). This effect is significant: the city of Beijing has achieved a 40 percent reduction in since China's War on Pollution campaign (Greenstone et al., 2021); our findings suggest a sizable share of that reduction is contributed by urban afforestation. On the other hand, the MMP has increased the average population pollen exposure by 7.4 percent.

Naturally, these findings raise questions about the health implications. In particular, can pollen exposure trigger health events that are severe enough to require substantial medical attention and resources and, if so, how these effects compare to the known impacts of air pollution exposure? The final part of this research casts light on these questions, providing direct estimates using administrative medical claims data covering all city residents. To begin, we highlight two facts about pollen and pollution variation: First, while the timing of pollen seasons is relatively fixed, there is substantial variation in pollen exposure across pollen seasons of different years; Second, there is a significant amount of independent variation in daily fluctuations of pollen and air pollution. These variations allow us to provide credible estimates on the effect of day-to-day pollen fluctuations on healthcare utilization, and to compare them with the same outcomes but looking at pollution fluctuations. We find that daily pollen shocks lead to significant increases in the number of emergency room (ER) visits, driven mostly by those due to respiratory or sensory system emergencies. The elasticity is on par with the effect of daily $PM_{2.5}$, both using estimates from our own data and those borrowed from the prior literature.

The health evidence suggests that the negative health externality associated with pollen exposure should not be overlooked: it is significant enough to trigger ER visits, with the magnitude of the elasticity similar to the effects of pollution on ER visits. However, several characteristics of the effects indicate that the overall health cost of pollen exposure is likely small compared to the benefits of pollution reduction: First, although both pollen exposure and $PM_{2.5}$ contribute to an increase in the *number* of ER visits to a similar extent, the impact on ER *spending* is three times smaller for pollen exposure. This suggests that ER visits due to pollen exposure tend to be less severe compared to those triggered by $PM_{2.5}$. Second, pollen primarily leads to increased ER visits for patients who do not require further inpatient care. In contrast, it is well-documented that $PM_{2.5}$ can result in severe health emergencies and can lead to death even at relatively low exposure levels (Deschenes, Greenstone, and Shapiro, 2017; Deryugina et al., 2019; Huang, Xing, and Zou, 2023). Third, unlike industrial pollution, which tends to disproportionately affect the elderly population, pollen has an equal or potentially stronger effect on the non-elderly population. We also find that pollen's impact appears to be concentrated among individuals with prior allergy histories, which may facilitate targeted prevention efforts.

Using both our own empirical findings and relevant estimates from the prior literature, we estimate that the annual healthcare cost savings resulting from the reduction of $PM_{2.5}$ due to the MMP project range from 229 to 916 million CNY (32 to 88 million USD). This corresponds to about 0.25 percent of Beijing's annual total reported health spending. Additionally, the annual mortality benefits, measured in terms of the value of statistical life (VSL), amount to an estimated value of 5 billion CNY (710 million USD). The healthcare costs associated with the increased pollen exposure related to the MMP project are about one-ninth of the magnitude of the pollution benefits, totaling 25 to 102 million CNY (3.5 to 14 million USD). We acknowledge that further research is necessary to fully account for the impacts of pollen on a broader range of well-being measures, such as cognitive health and productivity impacts.

This paper joins a nascent economics literature on the environmental impacts of forests and their downstream effects on health and economic prosperity, including the impacts of large-scale deforestation or afforestation on downwind precipitation and agricultural outcomes (Araujo, 2023; Grosset, Papp, and Taylor, 2023), the effects of urban trees on microclimate and property values (Han et al., 2021; Li, 2023), and the effect of urban afforestation on infant health improvement (Jones and Goodkind, 2019). The dearth of economic studies contrasts with a vast, multi-disciplinary effort to understand urban forests and their potentials in air quality improvement, climate moderation, energy conservation, carbon sequestration, noise reduction, stormwater management, wildlife habitat provision, among many other ecosystem services that are critical for sustainable urban development and climate change policies more broadly (Kahn and Walsh, 2015). While the natural sciences offer valuable insights into the interactions between forests and the environment, numerous questions remain open regarding the real-world effectiveness, efficiency, and equity of urban afforestation policies from an econometric standpoint.[1]

In addition to the environmental impacts, a unique feature of urban afforestation worth studying as a policy tool is its track record of implementability at large scale. We have studied the city of Beijing, but there are numerous examples of cities having been able to implement urban afforestation at mega scale, such as Los Angeles' Million Trees Initiative, New York City's Million Trees NYC initiative, Toronto's Every Tree Counts program, and Singapore's City in a Garden project. This level of implementation is noteworthy when considering urban afforestation within the context of environmental protection policies, which often encounter feasibility challenges, enforcement and compliance issues, and political resistance (Gray and Shimshack, 2011; Meng 2017; Mastini et al., 2021; Giles, 2022; Browne et al., 2023). We expect that similar urban afforestation programs and other forms of green infrastructure will play an increasingly important role in future sustainable urban development (Thacker et al., 2019).

[1] For example, on the topic of urban forestry and air quality, Nowak, Crane, and Sevens (2006) uses computational modeling approach combined with parameter calibration to predict impact of urban trees on air pollution removal. In atmospheric science, Abhijith et al. (2017) provides a detailed review of the atmospheric interactions between vegetation and surrounding, and the implications for pollution exposure. In epidemiology, Rojas-Rueda et al. (2019) conducts a meta-analysis of recent cohort studies linking NDVI to mortality.

On the health side, this research also adds to a small health economics literature on seasonal allergy and the economics costs in terms of cognitive outcomes such as test performance, crimes, and accidents (Marcotte, 2015; Bensnes, 2016; Chalfin, Danagoulian, and Deza, 2019; Akesaka and Shigeoka, 2022).

On the method front, we join a growing econometric literature on treatment effect with unit interference (Sävje, Aronow, and Hudgens, 2021): although the MMP policy causes forests to be planted in a given area, the environmental (and thus health) effects of those forests are not confined locally to those areas. We leverage structural knowledge on the nature of the interference – that the spatial spillover effects are likely driven by wind transport – and build a wind directivity design to capture the phenomenon. We then combine this design with causal estimates on two other marginal effect estimates that can be obtained through local treatment effect estimation – the impact of policy on local vegetation growth, and the impact of local pollen on health – to deliver the overall causal effect of the program on the environment and health. Our analysis is facilitated by a comprehensive dataset we compiled combining government surveys, high-resolution remote sensing data, real-time ground monitoring, and administrative medical claims data, which allows us to directly estimate each link on the causal chain from policy implementation to endpoint health outcomes.

The rest of the paper is organized as follows: Section 2 discusses policy background and data. Section 3 outlines our empirical framework. Section 4 documents Beijing's greening up and the role of the urban afforestation policy. Section 5 studies environmental effects of urban forests. Section 6 studies health effects. Section 7 ties together various estimates and discusses implications, and concludes the paper.

2. Background and Data

2.1 Background

A Brief Chronicle of Urban Afforestation Policies in Beijing. The 1977 World Conference on Desertification Control in Nairobi, Kenya, identified Beijing as a city on the brink of desertification, raising serious ecological alarms to the Chinese government. In reaction, China inaugurated Arbor Day in 1979 and the National Capital's Obligatory Tree Planting Day in 1985, encouraging public participation in the country's green transformation; the country also introduced numerous initiatives in the two decades between 1991 and 2010 including the Three-North Shelter Forest program, the Taihang Mountain afforestation project, among others. Many cities followed suit with local programs focused on sand control, afforestation, water source protection, and the establishment of green corridors and green separation zones.

Beijing, the capital city, has been particularly proactive in this movement. This culminated when the city committed to a "Green Olympics" as the host of the 2008 Beijing Olympics, propelling major regional afforestation and ecological initiatives such as the Beijing-Tianjin Sand Source Control Project and the Beijing-Hebei Ecological Water Source Protection Forest Construction Project.

Historically, the plain area of Beijing, the hub for the city's population and industries (Figure A.3, areas with <100m altitude), has been ecologically under-resourced, with lower forest coverage compared to the mountainous regions. In an effort to remedy this, in 2012, the Beijing Municipal Government introduced the "Million *Mu* Afforestation Project in Plain Areas." We call this project the Million *Mu* Project (MMP) in this paper. This ambitious initiative aimed to augment the forest coverage in these plain areas by an additional one million mu (approximately 165,000 acres) over five years, with a goal to attain over 25% forest coverage rate. This marked a significant milestone in Beijing's ongoing journey of urban afforestation and ecological restoration.

The Million *Mu* Project. The MMP project was discussed in a January 2012 meeting of the Beijing Municipal Government as part of the *Beijing Municipal Air Pollution Control Plan 2012-2020*. The stated goal of the MMP project is to improve the ecological environment of the capital city, reduce $PM_{2.5}$ pollution, and promote green and ecological development. The MMP project involves both the conversion of existing construction land and the reclamation of abandoned sand and gravel pits for afforestation purposes. It focused on key areas such as ecologically sensitive zones, the peripheries of roads and rivers, water source protection areas, and agriculturally non-viable land for focused forest creation.

The spatial layout of these urban forests was planned in alignment with the broader urban and green space system planning. The guidelines are known as "Two Rings, Three Belts, Nine Wedges, and Multiple Corridors."[2] Specifically, "Two Rings" refers to the creation of continuous green belts alongside the Fifth Ring Road, each extending 100 meters in width. These green belts act as the first line of ecological defense in the plain area. Further out, beyond the Sixth Ring Road, two additional green belts have been established – an outer one extending 1,000 meters in width and an inner one spanning 500 meters, jointly serving as the second ecological protection layer in the plain area. "Three Belts" denotes the establishment of permanent green belts along the banks of the Yongding River, Beiyun River, and Chaobai River, each stretching at least 200 meters wide. These belts function as vital ecological preservation zones. "Nine Wedges" is the creation of four functionally-defined, moderately-sized suburban parks within nine wedge-shaped areas of limited development. Together with large, contiguous forested areas, these parks provide crucial green spaces that bridge the urban core with the city's outskirts. "Multiple Corridors" encompasses

[2] http://yllhj.beijing.gov.cn/zwgk/cwgk/jbcwgg/202103/t20210319_2311459.html

the development of green passages along key roads, riversides, and railways, as well as health-oriented greenways that interconnect different regional forest landscapes and park green spaces.

The MMP shapefiles we use in the analysis (Figure 2, for example) represent actual planting sites that fall under these forest creation guidelines. See Section 2.2 for more discussion.

Program Financing. Funding for the MMP construction was sourced from both city and sub-city (district) levels. Data published by the city government indicate that the initial investment during the first three years of the project (2012 to 2015) amounted to 34.3 billion CNY, of which 25.5 billion CNY was contributed by the city government and the rest by district governments. Investment and construction policies varied across regions, but major cost items included purchasing fixed assets, building conservation areas, and performing maintenance and management duties post-planting. As the program expanded between 2015-2022, it incurred an additional cost of 40.6 billion CNY. We estimate the total recorded cost of the program to be around **75 billion CNY** (about 10 billion USD). We will compare this figure with potential health benefits in Section 7.

Program Outcomes. From 2012 to 2017, a total of 1.17 million *mu* (780,000 hectares) of afforestation and greening was completed in the city. Following the success of these efforts, a new phase of the million *mu* project was launched by the municipal government in late 2017 and concluded in 2022. This second phase of the project had a more extensive coverage, spanning the central urban area, new towns, and low mountainous regions. Through two consecutive five-year rounds of the MMP project, a total of **2.07 million *mu* (135,000 hectares)** of new afforestation and greening has been added to the city. In the plain areas of Beijing, the forest coverage rate increased from 15% in 2011 to 31% in 2022, while the overall forest coverage rate in the city improved from 38% to 45% during the same period.

Due to the lack of data regarding the exact planting sites established in the first five years versus the second five-year phase of the project, our analysis considers the entire decade starting from 2011 as the MMP policy period.

The Science of Urban Forests and Pollution. Vegetation leaves reduce pollution through a process known as phytoremediation. This happens in two main ways: First, small openings on leaves (stomata) absorb trace gas pollutants such as nitrogen dioxide and sulfur dioxide (Harris and Manning, 2010; Yin et al., 2011); Second, leaves "filter out" particle pollutants through dry deposition (McDonald et al., 2007; Nowak, Crane, and Sevens, 2006). In atmospheric science and urban planning literature, forests and urban greenery are widely recognized as an effective tool for pollution reduction (e.g., Baldauf, 2017; Kumar et al., 2019).

There is one notable exception: the implications for ground-level ozone, where forests and urban greenery can have both positive and negative effects. On one hand, planting trees can increase the rate of ozone deposition and absorption, reducing near-ground ozone concentrations. On the other hand, certain tree species, such as poplar, willow, and oak, release volatile organic compounds (VOCs) to repel pests and attract pollinating insects. These VOCs are precursors to ozone formation – that is, they react with nitrate compounds and sunlight to form ozone. Our empirical results below indeed show that the impact of urban forests on ozone is not as clearcut as their impacts on other criteria air pollutants.

Clinical Evidence on Pollen and Health. In Section 6, we econometrically quantifies the link between ambient pollen concentration and emergency room visits. Here we review some of the clinical foundation of that link.

Exposure to allergenic pollen has become an increasingly concerning environmental health issue in urban areas in recent decades (Biedermann et al., 2019; D'Amato et al., 2015). The prevalence rates of seasonal allergies caused by pollen range from 10% to 40% in developed countries, with an estimated 400 million sufferers worldwide (Greiner et al., 2011; Meltzer et al., 2012). The rise in allergies is particularly concerning in the context of global warming, as it is expected to prolong the plant growing season and increase the overall pollen production per season (Ziska et al., 2019).

Inhalation of airborne pollen can lead to seasonal allergies, often referred to as hay fever or pollinosis. This condition is a widespread chronic issue and a global health concern. Unlike the year-round threat posed by air pollution, such as $PM_{2.5}$, which impacts the respiratory and cardiovascular health of a large portion of the population, exposure to pollen exhibits strong seasonal variations and differentially affects individuals with allergy histories. Allergic reactions are triggered by the immune system. When a person who is allergic to substances like dust, mold, or pollen encounters these substances, their immune system might react excessively, producing antibodies that attack the allergen aggressively. A number of typical allergic reactions are linked with the production of a specific antibody known as immunoglobulin E (IgE) by the body. Allergens can be introduced into the body through inhalation, consumption, or contact with the skin.

Symptoms of seasonal allergies caused by pollen exposure involves the sensory system, including allergic rhinitis (sneezing, runny and stuffy nose) and allergic conjunctivitis (itchy eyes and tears). In rare cases, asthma and atopy may also occur (Sun et al., 2016). Severe allergic reactions can lead to bronchitis, bronchial asthma, pulmonary heart disease, and even life-threatening situations (Brunekreef et al., 2000). Pollen has a particularly noticeable impact on individuals with respiratory allergies, which affect approximately 10-30% of the global population (Sierra-Heredia et al., 2018). Seasonal allergies not only worsen physical and mental health but also decrease productivity, increase medical expenses, and reduce daily activities, thereby impacting people's quality of life. Higher pollen counts lead to increased visits to asthma emergency departments and more sales of over-the-counter allergy medications. Allergy medications used to alleviate symptoms can have side effects such as drowsiness, dry mouth, lethargy (Jáuregui et al., 2009; Meltzer et al., 2012), which may negatively impact cognitive performance and productivity.

Different types of pollen vary in allergenicity, which depends on the strength of allergenic pollen antigens and the pollen concentration. In Beijing, herbaceous plants like *Artemisia* exhibit relatively high allergenicity but low pollen levels during the autumn season, while deciduous trees like *Cupressaceae* have relatively weak allergenicity but high pollen levels during the spring season.

In the 1980s, the total amount of airborne pollen, primarily from herbaceous plants, was higher during summer and autumn than during the spring peak period. Since 2000, the pollen content from deciduous tree species has been increasing each spring, making spring the season with the largest proportion of the total annual pollen content. This shift became even more pronounced since 2010, with a significant increase in pollen from spring-flowering deciduous trees like *Ginkgo*, *Platanus*, and *Cupressaceae* (Zhao et al., 2021). Currently, the predominant pollen types in Beijing are from cypress and poplar trees. Chinese juniper pollen has also become an important allergen for spring pollen allergies, with the peak allergy season running from March to May.[3]

2.2 Data Sources

This project's data sources are tabulated in Table 1. Here we provide more details about each source.

Planting Sites. We get location information on the universe of MMP planting sites from a policy document published in 2022 by the Beijing Municipal Commission of Development and Reform. To be clear, the map shows where trees and other greenery ended up being planted (rather than where planting was planned).[4] We digitize and geo-reference these maps to create polygon files that represent locations of MMP sites. The total area of MMP sites according to our digitized data is 2.38 million *mu* (about 1,589 km^2), which is close to the official number of 2.07 million *mu* (1,380 km^2). Our analysis Section 4 provides

[3] See http://bj.cma.gov.cn/xwzx/mtjj/202103/t20210326_3023644.html

[4] *14th Five-Year Period Land Resources Protection and Utilization Plan in Beijing* (京政发 [2022] 26 号: 北京市"十四五"时期土地资源保护利用规划, last accessed January 5th, 2023.) This document outlines land resources protection and utilization planning for the five-year period of 2021-2025. The document provides MMP maps as a part of its summary of achievement in the previous decade.

further cross-reference checks using remote-sensing based vegetation index data, showing sharp increase in vegetation growth immediately starting the MMP boundaries.[5]

Land Use. We obtain land-use information from China's Land-Use/Cover Datasets (CLUD) from the Institute of Geographic Sciences and Natural Resources Research at the Chinese Academy of Sciences. CLUD are mainly based on Landsat images, and generated using a human-computer interaction (HCI) interpretation process. The data are available at 30-meter spatial resolution and five-year intervals, and we obtain data for the 2000, 2010, and 2020 cross sections. Our analysis uses level-1 classifications which contain six categories: cropland, forest, grassland, water bodies, construction land, and unutilized land. CLUD's level-1 classification is estimated to have an accuracy rate of over 94 percent (Yang and Huang, 2021). We use this data to characterize the type of places where MMP plantings areas were sited.

Population. We use population estimates data for the city of Beijing from the WorldPop 100-meter resolution product for year 2020. WorldPop combines satellite data, census, and machine learning to predict population distribution at fine spatial scale. We use WorldPop mainly to demonstrate the location of population hubs in the city.

Vegetation Index. Normalized Difference Vegetation Index (NDVI) is a widely-used remote-sensing measure of the density of vegetation (Pettorelli et al., 2005). NDVI is a function of the reflectance of two specific wavelengths of light: near-infrared (NIR) and visible red (RED). Healthy vegetation reflects a large amount of NIR light and absorbs most of the RED light. In contrast, non-vegetated surfaces, such as bare soil or water, reflect less NIR light and more RED light. The formula for NDVI for a given location i is:

$$\mathrm{NDVI}_i = \frac{\mathrm{NIR}_i - \mathrm{RED}_i}{\mathrm{NIR}_i + \mathrm{RED}_i}$$

The formula normalizes the values to a range of -1 to 1, with higher values indicating more vegetation. Negative NDVI values occur when an area reflects more red light than near-infrared light. This mostly occurs for water bodies in our sample, which we drop from our analysis.

The NDVI data are obtained from NASA Moderate Resolution Imaging Spectroradiometer (MODIS) product MOD13Q1.061, which provides 16-day composite of vegetation indices at a spatial resolution of 250 meters. We obtain data from January 2001 to December 2020 for the city of Beijing, and we aggregate up the temporal frequency from 16-day to monthly.

[5] To alleviate further concerns about measurement errors, particularly those associated with the digitization of small forest patches, in Appendix Figure A.4 we report a robustness check where we restrict our analysis to planting areas that exceed 1km^2 in size.

Remote-Sensed Air Pollution. A part of our analysis uses a satellite measure of Aerosol Optical Depth (AOD) to measure air quality. The data are from MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm which provides AOD at daily frequency and 1-km spatial resolution from 2001-2020 (product MCD19A2.006; Lyapustin, 2018). AOD is an index measure of sunlight scattering and absorbance, which provides a good proxy for the concentration of particulate pollution in the atmosphere. A downside of AOD is that it is technically a measure of column (ground to top-of-atmosphere) pollution instead of ground-level pollution. The upside is that, being a remote-sensing measurement, the data allow us to look at pollution changes where in-situ monitoring data are not available, such as places near or inside planting sites.

***In-Situ* Air Pollution.** Our main air pollution analysis uses data from 35 ground-level monitoring stations in the city from 2014 to 2019. Each station performs real-time monitoring of six criteria pollutants: fine particulate matter ($PM_{2.5}$), coarse particulate matter (PM_{10}), ozone (O_3), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO). Our data are at the station-daily level. The data are sourced from the Beijing Municipal Ecological and Environmental Monitoring Center.[6]

Pollen Counts. We have access to four years (2013-2016) of daily pollen counts data from 20 stations from the Beijing Meteorological Service Center. The pollen monitoring season spans from March 1 to September 30 each year, with some station-years extending to October 15. Pollen counts are measured using a filter-based method in which the pollen deposited on a filter is observed and the concentration is calculated by dividing the counts by the amount of air that passes through (with the unit of measurement being counts per 1,000 mm^3 of air).

Each station monitors six pollen species: *Cupressaceae* (Cypress), *Salicaceae* (Willow), *Pinaceae* (Pine), *Moraceae* (Mulberry), *Artemisia* (Sagebrush/Wormwood), *Chenopodiaceae* (Goosefoot). In our main analysis, we define pollen count as the sum of pollen counts across all six species. We also provide disaggregated analysis in Section 6 to examine species-specific health effect estimates, and compare those to clinical evidence. For example, pine tree pollen allergy is known to be relatively uncommon, whereas other species such as cypress tree pollen are known to be more allergenic (Charpin et al., 2005; AAAI, 2018).

A part of our analysis in Section 6 compares the ER impacts of ambient $PM_{2.5}$ and pollen. One technical issue worth addressing head-on: because pollens are also particles, are they picked up by $PM_{2.5}$

[6] In 2013, China implemented a pollution-monitoring reform which replaces its old monitoring system that was both sparse and shown to have reliability issues (Ghanem and Zhang, 2014). The post-reform system adopts real-time monitoring technologies used by the U.S. EPA, and the data are believed to be much more reliable (Greenstone et al., 2022).

monitors as well? Most pollen grain diameters range from about 10 to 100 micrometers, which are too large to be picked up by $PM_{2.5}$ monitors which use a size-selective inlet to only allow particles of 2.5 micrometers in diameters or smaller to pass through the inlet before the monitor measures particle concentration. This is often realized using a cyclone inlet design where air sample is forced through a spiral pathway, during which the larger particles are left out because they cannot change direction as quickly as smaller particles due to greater inertia. Most pollen particles will be left out at this stage.

Wind. We obtain wind direction data from 20 weather stations from 2014-2019 maintained by the Institute of Geographic Sciences and Natural Resources Research at the Chinese Academy of Sciences. For each air pollution and pollen monitoring station, we assign daily wind direction based on what was recorded at that station's nearest weather station.

Emergency Room Visits. To measure ER uses, we use anonymized medical claims data maintained by the Beijing Municipal Medical Insurance Bureau. This data contains the universe of institutional (hospital) medical service transactions for all city residents enrolled in the Urban Employee Basic Medical Insurance, a state-run mandatory medical insurance program that covers all urban residents employed in the formal sector. The data covers about 91 percent of all permanent residents of Beijing, and we have access to four years of data from 2013 to 2016.

It is worth noting that our claims data are *insurance* records maintained for reimbursement purposes, rather than hospital *medical* records. Therefore, although the data allows us to observe most information about an ER visit such as patient demographics, hospital location, date, and spending, it does not contain primary diagnosis made by the ER attending physician.[7]

But identifying the medical condition underlying the ER visit is important for this research. We do so by leveraging information on the type of medications associated with the ER visit, including drugs administered during the visit and prescribed medications for post-care use. Our data cover more than 20,000 distinct medications. We categorize these into one of the 14 disease treatment categories (such as respiratory, sensory, among others) using the Beijing Reimbursement Drug List for Basic Medical Insurance crosswalk file. In principle, an ER visit may involve medications from multiple categories. However, for simplicity, we define three primary categories: respiratory ER visits (involving any respiratory drugs), sensory ER visits (involving any sensory drugs), and other ER visits (not involving any respiratory or sensory drugs). Appendix Table A.1 tabulates the top ten most frequently prescribed medications in our respiratory and sensory ER categories.

[7] We likewise do not observe patient death, which is a commonly studied outcome in the environmental economics literature.

Table 1 summarizes data sources and where they are primarily used in the paper. Generally speaking, the remote-sensing data are used in Section 4 to document the effect of MMP planting on local vegetation growth and air quality change. The monitor data are used in Section 5 to document the impact of MMP planting on air quality and pollen changes at population hubs. Both the monitor and the medical claims data are used in Section 6 to estimate health implications.

3. Research Framework

Before proceeding, we outline the overall structure of the empirical analysis, and clarify several decisions we have made and why. The goal of this paper is to study the environmental and health effects of urban forests. We organize our econometric exercises around two conceptual equations. The first is an environment equation, which links the afforestation policy (MMP) to changes in city greenery, and to changes in environmental agents (ambient air pollutants and pollen):

$$\Delta \mathrm{Agent}_{\mathrm{i}} = \overbrace{\frac{\partial \mathrm{Green}}{\partial \mathrm{Policy}}}^{\text{Section 4}} \cdot \overbrace{\frac{\partial \mathrm{Agent}_{\mathrm{i}}}{\partial \mathrm{Green}}}^{\text{Section 5}} \quad (1)$$

The second is a health equation that quantifies the health consequences of the environmental changes:

$$\Delta \mathrm{Health} = \sum\nolimits_{\mathrm{i}} \underbrace{\frac{\partial \mathrm{Health}}{\partial \mathrm{Agent}_{\mathrm{i}}}}_{\text{Section 6}} \cdot \Delta \mathrm{Agent}_{\mathrm{i}} \quad (2)$$

Our empirical analysis revolves around econometrically identifying the three set of causal parameters highlighted in brackets in these equations. These parameters are the focus of Section 4, 5, and 6, respectively. Several features of this framework worth discussing:

Why Marginalize? A notable feature of our framework is that it takes a *marginal* approach: instead of directly estimating the causal effect of the MMP policy on environmental conditions and on health in one pass, we estimate three separate links on the causal chain (Policy → Greenery → Environment → Health).

The primary motivation underlying this design is the violation of the Stable Unit Treatment Value Assumption (SUTVA) in our study context. Whereas the MMP policy causes forests to be planted in a given area, the environmental (and thus health) effects of those forests are not confined locally to those areas. For example, a location's air quality is not only influenced by the presence of greenery in that exact location, but the prevailing wind direction and how much greenery is present in the upwind direction that altered the air before it reaches the location. Conventional instrumental variables model that assumes no cross-unit interference – that is, a study unit's outcome depends *only* on policy treatment status of the unit itself – is not applicable in our context when studying environmental and health changes.

Our idea of addressing this issue is to produce treatment effect estimates in separation, and in each part, we use the most appropriate design to causally estimate the corresponding marginalized effect, given the available data.

Section 4 is the most straightforward, where we combine government-provided maps of MMP planting area and satellite data on vegetation index to document the impact of planting on local greenery growth.

Section 5 is where we need to address the SUTVA violation: the environmental condition at any given location can be influenced by nearby MMP forests. Two features help us build the research design: First, we know the nature of unit interference is likely driven by wind transport. This gives us structural knowledge of where spatial spillover comes from, and we use a up/downwind directivity design – one that is in fact quite familiar with environmental economists – to capture that. Another helpful feature is that, most city population cluster in a small number of areas, and those areas all have ground-based monitors for both pollution and pollen. To the extent that we primarily care about changes in environmental agents experienced by the typical citizen, we only need to run the estimation at the monitor level, which makes the task computationally feasible.

Once we quantified how local environmental conditions changed due to MMP, Section 6 is a standard environmental-health analysis, linking local variation in environmental conditions with health outcomes. Our medical claims data allow us to build healthcare use variables at the district-by-day level, which we merge with air pollution and pollen data to conduct a standard panel data analysis.

Why MMP? A related question is why do we focus on the MMP policy in the first place? We focus on this particular policy because, as we will demonstrate below, it is the driving force of Beijing's significant greening-up in the War on Pollution era, it is such a large-scale policy that has significant impact on environmental experience of the typical resident of the city, and it is a policy with many more aspects understudied that worth exploring for further work.

It is worth noting that, while the MMP provides a narrative anchor for this paper, much of our analysis and estimates do not solely depend on this specific policy context. For example, in Section 5 when we estimate the causal effect of vegetation growth on air quality, we find that our results are similar if we calculate shocks of vegetation growth from *all* relevant upwind areas, rather than only from those influenced by the MMP policy.

4. The Greening Up of Beijing

4.1 Raw Statistics

We begin by characterizing Beijing's greening up using satellite NDVI data. For each 250m-by-250m grid, we calculate a "rate of vegetation growth" variation defined as the annual change in NDVI of the grid, i.e., the slope of a linear time trend fit across annual NDVI observations of that grid. We do so separately for the pre-MMP period (2001-2011) and the post-MMP period (2012-2020).

Figure 1 plots the results for the pre-MMP period (panel a) and the post-MMP period (panel b). In each panel, we provide a whole-city view on the left, and a zoomed-in, city-center view on the right. In the city-center view, we overlay a map of major road networks to provide a sense of the geography of economic activities.

These maps show that the city was experiencing vegetation growth even before 2012, but the growth has accelerated after 2012. The magnitude of this acceleration is substantial. Panel (b) of Figure 1 shows numerous hotspots (grids with blue color) popping up with an annual growth rate of over 0.02 units of NDVI per year. Based on a sample average NDVI of 0.34 in year 2010, these hotspots have experienced an increase of more than 53% over the period of 2012-2020.

In the Online Appendix, we provide more companion statistics on the greening up. Appendix Figure A.1, panel (a) shows annual NDVI grew by 11.7% between 2001 and 2011, and by another 16.7% between 2012-2020. The speed of growth is therefore 28% higher over the latter period. Figure A.2, panel (a) shows satellite-based land use categorization data that are available in years 2000, 2010, and 2020. Consistent with acceleration of vegetation growth, the land use data suggests a clear deacceleration of urbanization after 2010.

4.2 The Role of the Million *Mu* Project

In Figure 2, we use polygons to highlight areas corresponding to the Million *Mu* Project (MMP) planting areas, and then we overlay the map with NDVI growth rate 2012-2020. Panel (a) provides a whole-city view, and on the right hand side of the panel we zoom in two six example areas with large MMP patches. Panel (b) once again provides the city-center view. The results suggest a high spatial correspondence between MMP planting areas and post-MMP greening up.

Next, we provide more systematic evidence on this spatial correspondence. For each grid, we calculate its distance to the nearest MMP planting area boundary. Because many planting areas are large and contain multiple grids, for those grids that fall within the MMP boundary, we assign them negative distances. About 16.3% of grids fall within MMP planting areas. Among all grids that are outside of the MMP areas, about 90% are within 2 km to the nearest MMP boundary. Appendix Figure 2, panel (b) shows the distribution of grids by distance. We then estimate the relationship between NDVI growth and the grid's distance to MMP area by fitting the following regression equation:

$$\text{NDVI growth}_i = \sum_j \beta_j \cdot 1(\text{distance bin} = j)_i + \varepsilon_i \qquad (3)$$

where NDVI growth_i is rate of NDVI growth for grid i over either the 2001-2011 or 2012-2020 period, $1(\text{distance bin} = j)_i$ is a series of dummies indicating bins of distance-to-MMP-areas in 200-meter increment spanning -1km to 2km omitting the furthest bin 1.8km to 2km, and ε_i is the error term. We are interested in the β_j estimates, which shows the difference in NDVI growth rate of grids in a given distance bin j relative to grids in the reference bin 1.8km to 2km. We report confidence intervals calculated from 1-km gridded cluster bootstrap standard errors.[8] Note that equation (3) is simply a non-parametric expression of NDVI growth and distance-to-MMP.

The city of Beijing comprises the plain area in the southeast where the vast majority of the people live, and the mountainous area in the northwest. Appendix Figures A.2 and A.3 provide illustration. Given our objective is to measure the impact of the policy on citizens' experiences, our analysis will focus on the plain area, which we define as grids with elevation lower than 100 meters. This area contains 94 percent of the city's population. We report robustness checks using alternative elevation cutoffs, such as 50 meters or 200 meters, in Appendix Figure A.4.

Figure 3, panel (a) shows the results. The solid line, showing data from the post-MMP period, shows a sharp gradient of NDVI growth with respect to distance. This suggests the spatial correspondence between NDVI growth and MMP polygon that one can eyeball from Figure 2 before indeed holds true systematically in the data. The satellite measure detects differential change (that is, faster NDVI growth relative to the reference bin of 1.8-2km) outside of the MMP boundary up to 400 meters, which can either reflect errors evolved in government-provided MMP boundary map or in our digitization. However, it is reassuring that overall the MMP boundary corresponds to NDVI growth very well.

[8] That is, we project all individual grids i's onto a 1-km gridded map of the city. We call each 1-km grid a "cluster", and calculate standard errors using a cluster bootstrap. We go with this approach because other commonly-used methods, such as Conley standard errors, took infeasibly long to compute given the size of our data (over 120,000 grids).

One potential concern when interpreting the distance gradient as the causal effect of the MMP policy is whether the MMP areas are more likely to be located in places that would have experienced faster NDVI growth anyway.[9] To speak to this concern, panel (a) of Figure 3 also shows the NDVI growth-distance relationship for the *pre*-MMP period, where no sharp change in NDVI growth is observed around MMP boundary. In Appendix Figure A.2, panel (b), we show that land use categorization at baseline (year 2010) also change smoothly around MMP boundary. Together, the evidence suggests that the sharp, nonlinear increase in NDVI growth around MMP boundary, and only for the post-MMP period, is likely a consequence of the MMP policy.

4.3 Change in Air Quality near MMP Planting Areas

The estimation framework of equation (3) gives us an opportunity to take an initial look into the effect of vegetation growth on air quality. We repeat equation (3) but replace the outcome variable with the grid's growth rate of aerosol pollution – in fact, it is the rate of *decline* as most of the grids saw decline of pollution during the 2012-2020 period (Greenstone et al., 2021).

Figure 3, panel (b) shows a mirror image of panel (a), where the rate of decline of pollution is much faster within MMP areas. The fact that the gradient of pollution change starts almost exactly where the gradient of NDVI change starts adds to the confidence of the causal effect of MMP. Once again, no similar pattern is shown for the pre-MMP period of 2001-2011.

Three points worth noting. First, we have made two causal claims. The first causal claim is that plantation under MMP causes NDVI growth to accelerate after 2012. The second causal claim is that increased vegetation growth causes the reduction in air pollution near the planting area. As we discussed above, these claims are supported by the fact that, spatially, NDVI and pollution gradient changes exactly where MMP areas start, and, temporally, such relationships are only observed after MMP was in place.

One remaining confounder in making the latter causal claim (vegetation causes pollution decline) which we have not addressed is the fact that both the NDVI and aerosol pollution data are drawn from satellite sources, which are ultimately derived from shared underlying data. So suppose for some reason it is easier to detect pollution change in areas with faster NDVI change, then panel (b) of Figure 3 could reflect a mechanical relationship due to natures of satellite detection, rather than genuine changes in pollution. We are not aware of a clear pathway that can lead to such mechanical relationship, though we

[9] For example, suppose MMP areas are disproportionately located in high-vegetation regions, and high-vegetation regions tend to feature higher speed of vegetation growth, then the observed relationship may not purely reflect the effect of the MMP policy but rather just site selection.

are not able to fully rule it out either. However, our analysis in Section 5 using ground level pollution monitoring data will not be subject to this concern.

Second, the graphical pattern of panel (b) of Figure 3 appears to suggest a lack of effect on NDVI or pollution beyond the MMP boundary. But notice that equation (3) can only speak to changes in the outcome *relative to* the baseline group. In other words, while it is possible that MMP vegetation has spillover effect on air quality beyond the MMP boundary – for example, through wind transport – equation (3) is not suited to estimate that "general equilibrium" effect. We will directly estimate spillover effect in Section 5 next.

Third, in terms of effect sizes, panel (a) of Figure 3 suggests MMP grids grew about 8.1 percentage points faster than the growth rate of grids outside MMP; panel (b) suggests the rate of pollution reduction is 1.0 percentage point faster for MMP grids than non-MMP grids. This implies a long-term growth rate elasticity between NDVI and pollution of -0.125. We will show later that our short-term estimate of NDVI-pollution elasticity at the daily level using ground-based monitor data has larger but same order-of-magnitude elasticity (about -0.40).

5. Environmental Effects

Section 4 documents the substantial greening up of the city and demonstrates the role of the MMP policy. A limitation of the analysis, as we briefly discussed, is that the cross-sectional approach focused on fine-grained comparison of environmental conditions in areas in- and outside of MMP planting areas, and is not suited to identify the potential spillover effect of vegetation growth on the broader population, which might have missed the bulk of the policy. We estimate that about 1.3 million people live within MMP planting areas, while the city's population is 20 million people. The goal of this section is to estimate the causal effect of the vegetation growth on environmental conditions felt by the "typical resident", who lives in city population centers.

5.1 Estimation

The key insight of the empirical analysis of this section is this: to estimate environmental changes experienced by the typical resident, we will simply use air quality monitor as a unit of analysis, and estimate how changes in vegetation near the monitor has caused a response in air quality values as captured by the monitor. There are two main rationales behind this decision.[10] First, monitors are installed by the government to record population exposure to environmental conditions, and therefore most monitors are sited in areas with high population density. In panel (a) of Figure 4, we overlay the location of air pollution (or pollen) monitors on a map of grid-level population estimates. Evidently, monitors are placed near most densely populated areas. Conducting analysis at the monitor level therefore will give rise to estimates that reflect conditions that are representative to population exposure. Second, monitors capture environmental conditions that happen at the ground level and therefore more accurately reflect what people experience. This is in contrast with satellite data of pollution, which measures pollution concentration for the entire ground-to-top-of-atmospheric column of air. Using data directly collected by the monitors also helps avoids potential mechanical link between our satellite-based vegetation measure and pollution measure, which we discussed earlier in Section 4.

Recall that the objective of our analysis is to speak to how MMP-led vegetation growth may cause changes in air quality outside of the immediate planting area. In other words, for a given monitor, we are interested in not only how vegetation condition at that exact location matters for air quality recording of the monitor, but also potentially the impact of vegetation condition farther away. The particular mechanism we have in mind for this to render is through wind transport. To see this, let $\mathrm{P_{it}}$ denote the air quality condition (pollution or pollen) captured by monitor i on date t. A monitor can capture air quality because there is air movement, and therefore $\mathrm{P_{it}}$ is a measure of the (unobservable) true air quality condition upwind of the monitor, which we denote $\mathrm{P_{it}^U}$. Hence, one can write:

$$\mathrm{P_{it}} = \gamma^{\mathrm{U}} \cdot \mathrm{P_{it}^U} + \varepsilon_{\mathrm{it}} \quad (4)$$

where γ^{U} is a positive coefficient, and $\varepsilon_{\mathrm{it}}$ is the error term. Now consider the role of (observable) vegetation in the upwind area, denoted $\mathrm{Green_{it}^U}$, which we assume interact with pollution in the area following:

$$\mathrm{P_{it}^U} = \tau \cdot \mathrm{Green_{it}^U} + \epsilon_{\mathrm{it}} \quad (5)$$

where τ is an emission factor. For pollution, the emission factor of greenery is negative (vegetation absorbs or filters out pollution); for pollen, the emission factor is positive (vegetation emits pollen). ϵ_{it} captures remaining randomness in the relationships. Substitute (5) into (4), one gets an equation that relates air quality to vegetation condition in the upwind area:

[10] For the sake of brevity, we will henceforth use "air quality" to refer to either air pollution concentration or pollen counts. We will be more specific which one we refer to when necessary.

$$P_{it} = \gamma^{U}\tau \cdot \mathrm{Green}_{it}^{U} + e_{it} \quad (6)$$

This is the equation we take to the data. Specifically, we estimate:

$$\mathrm{Log\ Air\ Quality}_{it} = \beta^{U} \cdot \mathrm{Log\ NDVI}_{it}^{U} + \alpha_i + \alpha_t + e_{it} \quad (7)$$

In this equation, $\mathrm{Log\ Air\ Quality}_{it}$ is logged air quality measured at monitor i on date t. $\mathrm{Log\ NDVI}_{it}^{U}$ is logged NDVI at the "upwind area," defined as a 135-degree cone of a 10km-radius area in the upwind direction of the monitor. See panel (b) of Figure 4 for an illustration. In practice, because NDVI is available at the monthly frequency, the upwind NDVI variable is constructed using daily wind direction data but monthly NDVI value. For simplicity, in our primary specification, we calculate upwind NDVI using all grids that fall within the upwind cone. In the appendix, we report an alternative version where we restrict only to MMP grids within the upwind cone. Our results turn out to be more precise if we use MMP grids only. However, we stick with all grids because we do not see a strong reason to assume that local air quality changes are driven only by vegetation growth in the MMP boundary. α_i and α_t are unit and time fixed effects. Our baseline model uses monitor, year, month, and day-of-week fixed effects. e_{it} is an error term, and we cluster standard errors two-way at both the monitor and the date levels.

The coefficient of interest is β^{U}. Following our previous discussion, we expect this coefficient to be negative for air pollution, and positive for pollen.

We augment the baseline estimation equation (7) in the following ways to assess its validity and robustness. First, we estimate a version of equation (7) adding 20 leads and 20 lags of $\mathrm{Log\ NDVI}_{it}^{U}$, giving rise to an event study representation of the impact of vegetation shocks on air quality outcomes. To be precise, we estimate:

$$\mathrm{Log\ Air\ Quality}_{it} = \textstyle\sum_{k\in[-20,20]} \beta_k^{U} \cdot \mathrm{Log\ NDVI}_{i(t+k)}^{U} + \alpha_i + \alpha_t + e_{it} \quad (8)$$

and we plot the $\{\beta_k^{U}\}_{k\in[-20,20]}$ coefficient estimates. Second, we conduct a placebo test, replacing the upwind shock variable to measured vegetation at the *downwind* direction of the monitor, which is expected to have no impact on air quality readings capture at the monitor's location.[11] See panel (b) of Figure 4 for an illustration. Third, we report sensitivity checks along three dimensions where we had to make some

[11] This is, of course, subject to potential autocorrelation of NDVI at the upwind and downwind direction. In practice, our results are similar if we include both upwind and downwind NDVI in the same specification, that is:

$$\mathrm{Log\ Air\ Quality}_{it} = \beta^{U} \cdot \mathrm{Log\ NDVI}_{it}^{U} + \beta^{D} \cdot \mathrm{Log\ NDVI}_{it}^{D} + \alpha_i + \alpha_t + e_{it} \quad (9)$$

where $\mathrm{Log\ NDVI}_{it}^{D}$ is logged NDVI at the downwind area. We report these results in Appendix Table A.3.

arbitrary specification decisions: the radian of the cone that defines upwind and downwind area, the radius around the monitor where we calculate vegetation exposure, and fixed effects choices. We will give more details to these sensitivity checks when we come to discuss those results.

5.2 Results

Figure 5 summarizes the main results using equation (8). To recap, the structure of the underlying data is a daily panel dataset of pollution or pollen monitors. On the horizontal axis, we order coefficients so that negative event days are coefficients of the lead terms of Log NDVI_{it}^{U}, positive event days are coefficients of the lag terms, and event day zero represents the day-of term – hence the axis title "Days since shock." We plot the coefficient estimates of $\{\beta_k^U\}$ and the associated 95% confidence intervals. On the same chart, we overlay the placebo test where we run the exact same regression, but using *downwind* NDVI as the right-hand-side variables instead.

Panel (a) shows results for fine particulate matter ($PM_{2.5}$), which shows a decline of pollution of about 0.5 log units for two days upon a log increase in the upwind NDVI shock. Panel (b) shows results for pollen counts, showing an increase of about 1 log unit per log increase in upwind NDVI. The pollen effect shows no lagged effect, concentrating on the day of the shock. For both the pollution and the pollen outcome, we find no effects for the placebo, downwind NDVI shocks.

In Appendix Figure A.5, we further report the event study analysis for five other air pollutants we have data on: coarse particulate matter (PM_{10}), ozone (O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO). We find clear evidence of a pollution reduction effect across these pollutants except for O_3, which shows no response to vegetation shocks. As discussed in Section 2, ozone is a secondary pollutant that forms due to reactions of precursor pollutants, including nitrogen oxides (NOx) and volatile organic compounds (VOCs); while urban forests absorb precursors of ozone such as NO_2, they generate VOCs, which may cause the net effect on ozone ambiguous.

The Online Appendix reports a series of sensitivity checks. Appendix Figure A.6 and Appendix Tables A.2 through A.4 show results when we alter the definition of upwind cone using alternative radian degrees (0.25π, 0.75π, or 0.875π), alternative fixed effects (the inclusion of more flexible, station-by-month fixed effects and/or year-by-month fixed effects, alternative radius area around the monitor (5 km, 10 km, or 15 km), and whether we calculate NDVI using all grids or grids fall within the MMP planting areas. Our results are robust to these arguably significant specification changes. One exception is panel II of Appendix Table A.2, where we lose power in the pollution reduction effect estimation when we confine the definition of upwind NDVI to areas within a 5km radius of the pollution monitor.

It is worth noting that this pattern does not necessarily conflict with our cross-sectional findings in Section 4, where we find that satellite-measured pollution has declined faster within the MMP area, where more NDVI growth is observed. The estimation here diverges from the cross-sectional exercise in that it links the day-to-day fluctuations in pollution *within a location* to NDVI from a specific wind direction on that day. Our results suggest that the impact of urban forestry on pollution reduction requires a fairly expansive area to become apparent. In other words, the reduction in pollution isn't solely dependent on the area immediately surrounding the point of observation, but instead is determined by a larger area of forests upwind. The cross-sectional exercise in Section 4 cannot tease out this relationship.

Circling back to Figure 5. To the best of our knowledge, this is the first incident where researchers were able to estimate – both qualitatively and quantitatively – the impact of vegetation shock on air pollution reduction and pollen increase within the same empirical framework. Taken at the face value – that is, if we just judge the results from a pollution-vegetation and pollen-vegetation elasticity perspective – the pollen response appears large and potentially important. The natural next step is to quantify what are the magnitudes of pollution and pollen shocks in terms of health costs, and how do they compare with each other.

6. Health

This section aims to quantify the health parameters in equation (2). We begin by providing new estimates on the health effects of pollen exposure, linking day-to-day variation in pollen counts to emergency room visits and spending. We next discuss estimates on the health effects of $PM_{2.5}$ pollution from the existing literature, and a direct estimation using our own data.

6.1 Estimation

We do not observe the residential location of the patient from the medical claims data, but rather the hospital at which the healthcare was delivered. Similarly, as we discussed above, the city monitors pollen and pollution at 12 and 35 monitoring locations, respectively. To streamline these measurements, in our analysis we use "district" as the cross-sectional unit. There are 16 districts in Beijing, and Panel (a) of Figure 4 shows the district delineation. We make this choice because every district has at least one hospital, one pollen monitor, and one pollution monitor, and therefore it is intuitive to aggregate the medical claims data, the pollen data, and the pollution data to the district-daily level.[12]

We estimate the health effects of pollen using a fixed effects panel estimation model:

$$\text{Log Emergency Room Uses}_{it}^{g} = \beta \cdot \text{Log Pollen Counts}_{it} + \alpha_i + \alpha_t + X_{it}\gamma + e_{it} \quad (10)$$

where $\text{Log Emergency Room Visits}_{it}^{g}$ denotes the log of total emergency room (ER) utilization – total number of visits or total medical spending associated with the visits – for district i and date t. In the main specification, we measure ER outcomes using a three-day lookahead window, counting ER uses for day t, $t+1$, and $t+2$. This is intended to capture the effect of a transient pollen shock on health outcomes that might take some days to manifest. In Appendix Tables A.5 and A.6, we repeat the estimation using same-day outcome (no lookahead) or seven-day outcomes (even longer lookahead) as robustness checks.

In heterogeneity analysis, we will look at ER outcomes for different subgroups, and hence the superscript g. The subgroup categorization we are most interested in is the underlying medical conditions that led to the ER. We will report coefficients when the outcomes are ER for all causes, respiratory causes, sensory causes, and other non-respiratory and non-sensory causes. We will show that the effects of pollen exposure are driven by respiratory and/or sensory-related ER, which helps provide a sense of plausibility of the results. In other heterogeneity analyses we will also look at subgroups by, for example, patient's age.

Regarding the rest of the equation (10): $\text{Log Pollen Counts}_{it}$ is the log of pollen counts at district i on date t. We include district-by-month, year, month, day-of-week, and holiday fixed effects, which we denote α_i and α_t for simplicity. In some robustness specifications we further control for factors that can potentially correlate with both pollen and influence health outcomes (such as particulate pollution or air temperature), which is captured by the matrix X_{it}. Standard errors are two-way clustered at the district and the date levels.

Equation (10) does not have a causal research design, and so whether β reflects the causal impact of pollen depends on how much we believe the day-to-day variation in pollen, conditional on the fixed effects controls, is as good as random. We first inspect the nature of pollen variation. Panel (a) of Figure 6 shows the daily time series of city-average pollen counts. The city has two pollen seasons, one around

[12] Implicitly, we assume here that a monitor's measurement is a good proxy for the average population exposure in the district, and that people seek healthcare in hospitals in their home district. The former assumption is reasonable because monitors are placed in densely-populated areas. District-based healthcare seeking is a reasonable assumption to the first order, especially in the context of emergency room visits – our focal outcome measure of healthcare use – where people are unlikely to travel long distances to seek emergency care.

March and the other around September. This seasonality holds consistently across the four years (2013-2016) of our data, and therefore any strategy that relies on within-year, cross-season comparison will miss out confounders that are correlated with such seasonality. Instead, we note that there seems to be abundant pollen variation *within* a given season, but *across* different years, which is more likely to be driven by idiosyncratic variation. This underlies our choice to include the *district-by-month* fixed effects in equation (10), which allows us to compare conditions at a given district *and* in a given month-of-year, but across different years when pollen counts are high versus low.

For the rest of the discussion, we will simply refer to β as the *effect* of pollen exposure on ER uses.

6.2 Results

Panel (b) of Figure 6 provides an initial look of the main results. It shows decile bin scatterplots of fixed effects-residualized ER visits and pollen counts, done separately for respiratory/sensory-related ER visits and all other ER visits. The two dashed lines are superimposed linear fit, whose slopes correspond to the β coefficient from equation (10). The results reveal that the effect of pollen concentrates mainly for respiratory-sensory causes, but not others. In a way, this finding provides a degree of ex-post confidence that our estimates do capture the health effects of pollen – in lieu of a spurious correlation – because a priori clinical knowledge suggests the potential effect of pollen should operate through respiratory and sensory channels (Section 2).

Table 2 provides more details of the estimation results. Each cell in this table represents the β coefficient estimate from a separate regression. Given the log-log specification, all numbers in the table represents elasticity. Columns 1-4 look at number of ER visits as the outcome variable, and columns 5-8 look at ER spending as the outcome variable. For both ER visits and spending, we examine all-cause ERs, those due to respiratory or sensory causes, and all other non-respiratory/sensory ERs.

Panel A reports the baseline results. We find an ER-pollen elasticity of 0.021 (SE = 0.004) for respiratory ER visits and 0.042 (SE = 0.006) for sensory visits (columns 2 and 3). We found similar effect sizes for ER spending (columns 6 and 7). The effects on non-respiratory/sensory ER uses are an order of magnitude smaller and, in the case of spending, statistically insignificant (column 4 and 8).

Panels B, C, and D of Table 2 documents three dimensions of heterogeneity. In panel B, we stratify ERs by whether the patient ended up being hospitalized or not. We find that the effect comes through almost entirely by ERs that do not need further inpatient care.

In panel C, we stratify ERs by age of the patient. We find that the effect sizes (in elasticity) are statistically indistinguishable for elderly (aged over 60) and non-elderly. If anything, some of the effect sizes – for example, sensory-related ER visits – are larger for the non-elderly.

These heterogeneity results in panels B and C are worth noting because they suggest the characteristics of pollen's health effects differ from what we know about the health effects of industrial pollution. For example, $PM_{2.5}$ pollution is well known to disproportionately affect the elderly population and can cause severe conditions including hospitalization and even death (Deschenes, Greenstone, and Shapiro, 2017; Deryugina et al., 2019). This is something to keep in mind when we compare the health costs of pollen versus $PM_{2.5}$: while we argue that pollen exposure imposes non-neglectable health risks, they are unlikely to *exceed* the health benefits of industrial pollution reduction. We will come back to this point in Section 7.

In panel D, we further document that pollen's effects are larger for patients with prior respiratory-sensory conditions, defined as those who had respiratory and sensory visits in the past month. From a health management perspective, this result suggests the health effects of pollen might be more manageable than those of pollution, as one may target the protection of easily-identifiable vulnerable group during pollen seasons, which are also relatively fixed and predictable than industrial pollution episodes.

In Table 3, to cast light on the type of pollen species that are particularly damaging, we replace the right hand side variable $\text{Log Pollen Counts}_{it}$ of equation (10) to be the log of pollen counts of a specific species. We find that both tree (*Cupressaceae*, *Salicaceae*, *Moraceae*) and weed pollen (*Artemisia*, *Chenopodiaceae*) seem to trigger increased ER visits. Once again, for all of these species, respiratory and sensory ERs are driving the effects. The only exception where we observe no significant ER effect is *Pinaceae* pollen that is mostly associated with pine trees. This is consistent with clinical evidence that pine tree pollen-caused allergy is uncommon (AAAI, 2018). The takeaway of Table 3 is that, except for pine tree pollen, the other five species appear almost equally bad as ER triggers. If we view ER spending as a measure of severity, then it seems reasonable to conclude that *Cupressaceae* (Cypress) and *Moraceae* (Mulberry) have the largest impacts.

6.3 Effect Sizes

The analysis so far finds a statistically significant increase in ERs due to pollen exposure. For example, panel (a) of Figure 6 suggests an ER visits-pollen elasticity of 0.0061 (SE = 0.0013) and an ER spending-pollen elasticity of 0.0037 (SE = 0.0012). How important are these effects? A simple way to answer this question is to estimate the same sets of ER-elasticities, but for industrial pollution such as

$PM_{2.5}$. The prior literature on the health effects of $PM_{2.5}$ also has established estimates that we can benchmark against, which we discuss towards the end of this section.

To estimate the effect of $PM_{2.5}$ exposure on ER, one replaces the log pollen variable on the right hand side of equation (10) with Log PM2.5_{it}, i.e., log $PM_{2.5}$ concentration at district i on date t. Before doing that, it worth considering whether one actually has independent variation between pollen and $PM_{2.5}$ concentration to pick up effect separately for these two environmental agents. We discussed in Section 2 that, in theory, most pollen grain diameters range from about 10 to 100 micrometers, which are too large to be picked up by $PM_{2.5}$ monitors which are calibrated to measure particles of 2.5 micrometers in diameter or smaller. To assess this fact empirically, in panel (a) of Figure 7, we provide scatterplots of daily $PM_{2.5}$ against pollen counts at the district level, separately for each month of the pollen monitoring season (March to September). We find that in most month of the year, $PM_{2.5}$ and pollen exhibit small and statistically insignificant correlation, with the direction of correlation showing no particularly consistent pattern (positive in March and May, and negative for other months). The exception is the month of August where the two variables exhibit a strongly *negative* correlation (elasticity = -0.240, SE = 0.045). Given the overall abundant independent variation between the two variables, it seems reasonable to proceed with estimating the separate impacts of the two environmental agents on health outcomes.

Panel (b) of Figure 7 reports the estimates. The solid lines repeat Table 2a, columns (1) and (5) on the effects of pollen on ER visits and ER spending. The dashed lines show the effects of $PM_{2.5}$ using the exact same specifications. We find that a log increase in $PM_{2.5}$ leads to an increase of three-day ER visits by 0.007 log points (SE = 0.0023), which is similar to the pollen-ER visits elasticity estimate. The $PM_{2.5}$ elasticity for ER spending is estimated to be 0.0095 (SE = 0.0027), over two-fold of the pollen-ER spending elasticity. This is consistent with our prior observation that pollen exposure tends to cause less severe conditions, and thus the effect of $PM_{2.5}$ on ER spending is larger, even if the effect on visitation rate is similar.

Are these reasonable estimates of the effect size of $PM_{2.5}$? We compare our estimates to published estimates from the literature that are similar in approach. Deryugina et al. (2019) links daily variation of $PM_{2.5}$ driven by changes in wind directions to health outcomes in the U.S. Medicare population. In their primary specification, the effects of $PM_{2.5}$ on three-day ER visits and spending converts to an elasticity of 0.0071 and 0.0126, respectively. A potential concern with comparing to is that the background $PM_{2.5}$ level in Deryugina et al. (11 ug/m3) is much lower than that Beijing during our study period (67 ug/m3), and so the elasticities may not be comparable. Barwick et al. (2022) links daily $PM_{2.5}$ in China to frequency of bank card transactions in healthcare categories. Their primary specification converts to a daily $PM_{2.5}$-transaction elasticity of 0.0364. Xia et al. (2019) studies the effect of $PM_{2.5}$ on medical spending using the same source of medical claims data from Beijing. Their most precise estimate linking daily $PM_{2.5}$ to three-day medical spending converts to an elasticity of 0.0132. Our pollution estimates are smaller but generally in line with prior evidence.[13] One source of difference is that we hold fixed the estimation strategy for $PM_{2.5}$ and pollen, both using simple OLS regressions, whereas most prior $PM_{2.5}$ studies adopt some form of instrumental variables approach, which tends to produce larger effect sizes.

It is worth emphasizing that the point of this subsection is to put our pollen-ER elasticity estimates in perspective by comparing them with $PM_{2.5}$-ER elasticities, both estimated using our own data and from the literature. Our headline conclusion – that the healthcare externality of pollen allergy is nonnegligible – is based on our finding that pollen exposure is qualitatively bad enough to trigger ER visits, and quantitatively similar with ER effects of pollution. However, this is not to say that the *overall* health costs of pollen are necessarily on par with those from industrial pollution exposure. In particular, various studies have documented the mortality effect of $PM_{2.5}$, which is believed to be an overwhelming component of the overall health costs of pollution when compared to morbidity and healthcare costs (e.g., Landrigan et al., 2018). Though our data do not allow us to observe mortality rate and so we cannot directly estimate the pollen-mortality coefficient, we expect the effect to be small, if any, given that most pollen-related ER does not lead to hospitalization, let alone more severe, life-threatening complications.

7. Discussion and Conclusion

7.1 Health Implications of the Million *Mu* Project

We are now ready to plug in the marginal parameter estimates from Sections 4, 5, and 6 to the environment and health equations of Section 3. We have three key sets of parameters: First, the MMP increases planting site NDVI by 0.032 units or about 10.5 percent over the 2012-2020 period. Second, for the typical resident, a log unit increase in upwind NDVI in MMP areas causes -0.4 log point reduction in $PM_{2.5}$ and +0.7 log point increase in pollen. Third, a log unit decrease in a district's $PM_{2.5}$ leads to 0.007 log points decrease in ER visits; a log unit increase in pollen leads to 0.0061 log points increase.

Together, we estimate that, by 2020, the MMP policy reduces the average population $PM_{2.5}$ exposure in the city by 4.2 percent (about 2.9 ug/m^3 from 2012 baseline).[14] This effect is significant: the

[13] There is a large literature on the mortality effects of air pollution, but the availability of morbidity and healthcare costs estimates is more limited. We will discuss the mortality literature in more detail in Section 7.

[14] To get baseline 2012 $PM_{2.5}$ value, we use annual average 69.5 ug/m^3 as reported by U.S. Embassy in Beijing.

city of Beijing has achieved a 40 percent reduction in pollution since China's War on Pollution campaign. Our estimates suggest a sizable share of that reduction is contributed by urban afforestation.

We can calculate the health value of such air quality improvement. There are many established estimates we can choose from the literature. Here we focus on quasi-experimental, economics studies that are conducted using health data from China, so that the study contexts are more compatible. Starting with healthcare costs. Xia et al. (2019) uses medical claims data from Beijing, and finds that a 10 ug/m3 increase in $PM_{2.5}$ concentration increases city-wide medical spending by 791 million CNY (112 million USD) per year. Xia et al. (2019) focuses on short-term impact of $PM_{2.5}$ shock on spending over a three-day window. Barwick et al. (2022) introduces a finite B-splines method to capture lagged effects of pollution for up to three-months. Using national data on credit and debit card transactions took place in healthcare institutions, the study finds that the cumulative effect of pollution at the medium run (three-month) is about four times larger than the effect of a contemporaneous, one-day shock. If we scale up the Xia et al. (2019) estimate accordingly, it implies an annualized cost of 3.2 billion CNY (448 million USD) per 10 ug/m3 increase in $PM_{2.5}$. Therefore, the annual healthcare benefits of a 2.9 ug/m^3 reduction in $PM_{2.5}$ due to the MMP project is estimated to be between **229 to 916 million CNY** (32 to 88 million USD). This amounts to 0.1 to 0.4 percent of Beijing's reported annual total health spending of 219 billion CNY.

Another important component of the health benefit of pollution reduction is through a mortality effect. He, Fan, and Zhou (2016) links monthly mortality rate to the reduction in PM10 in northern Chinese cities (including Beijing) driven by the 2008 Beijing Olympic Games. The study finds that a 10 ug/m3 reduction in PM10 (about 6 ug/m3 in $PM_{2.5}$) causes an 8 percent reduction in all-cause mortality. Ebenstein et al. (2017) exploits quasi-experimental differences in long-term PM10 exposure across cities in China. The study finds that a 10 ug/m3 increase in PM_{10} causes an 8 percent increase in the cardiovascular mortality. Combining these estimates with an average mortality rate of 4.3 per 1,000 residents (per Beijing city government) and a population of 20 million people, we estimate that a 2.9 ug/m^3 reduction in $PM_{2.5}$ due to the MMP project contributes to an annual reduction of 3,325 deaths. Following Barwick et al. (2022), we use a VSL of 1.5 million CNY and this gives us an annual mortality benefits of **5 billion CNY** (710 million USD). Alternatively, we use the Air Quality Life Index (Greenstone and Fan, 2018) and calculate that the potential gain in life expectancy is **0.28 years** for the average city resident if the reduced pollution effect of the MMP is to sustain in the long run.

Recall that in Section 2.1, we mentioned that the total cost of the MMP project is **75 billion CNY**. Our health calculation thus leads to a conclusion that these costs will likely be recouped via health gains over the next decade.

Our judgement is that the health benefit from pollution reduction likely overwhelms the cost of pollen increases. This is for two reasons. First, we find no evidence that pollen increases the need for inpatient care (Table 2), whereas an inpatient visit costs 40 more than the average hospital visit, and total inpatient spending accounts for 70 percent of overall medical spending in Beijing. Combining this fact with our estimate in Figure 7 that the cost elasticity of pollen is about a third of that of pollution exposure, we estimate the magnitude of the healthcare costs of pollen increased related to MMP project to be about one ninth of the magnitude of the pollution benefit, or **25 to 102 million CNY** (3.5 to 14 million USD). Second, the lack of a hospitalization effect, combined with the fact that we observe a major share of the effect comes through the non-elderly population, makes it reasonable to think that pollen is unlikely to lead to more severe, life-threatening complications that cause deaths. Therefore, if we were to take into account the potential life years saving of pollution reduction, the relative health costs of the pollen increase may be even lower.

We reiterate that it is not the intention of this calculation to downplay the health cost of pollen. Quite the opposite, our empirical analysis points out that pollen allergy is important enough to cause significant increases in emergency room usage. Our findings indicate that pollen triggers emergency room visits to a similar extent as $PM_{2.5}$ does. The smaller health cost numbers are driven by the observation that pollen-related emergency room visits are generally less severe, with a lower likelihood of requiring expensive inpatient care or resulting in fatalities, compared to pollution-related visits. However, a comprehensive understanding of the full range of effects of pollen exposure on cognitive and physical health, as well as other aspects of well-being, requires further research and investigation.

7.2 Conclusion

Urban forests have a ubiquitous presence in cities worldwide, but few economic research studies them. This paper examines the effect of urban forests on air quality and health, leveraging the greening up of the city of Beijing and its policy experiment over the past decade. We begin by documenting a substantial greening up of a mega city and examine the contribution by a government-led mass afforestation policy. We then quantify the impact of urban forests on downwind air quality improvement using a quasi-experimental research design. Our paper also investigates pollen changes as a source of negative health externality. We conclude that Beijing's afforestation brings enormous health benefits: by our calculation, the policy is expected to cause a 0.25 percent reduction in the city's overall healthcare spending due to its air quality improvement effect. The health costs of pollen increase are about an order of magnitude smaller, though we argue it is still significant and worth paying attention to. We provide insights for managing urban forests' health risks, including evidence on the most harmful pollen species and the population subgroups most susceptible to these effects.

References

Abhijith, K. V., Prashant Kumar, John Gallagher, Aonghus McNabola, Richard Baldauf, Francesco Pilla, Brian Broderick, Silvana Di Sabatino, and Beatrice Pulvirenti. "Air pollution abatement performances of green infrastructure in open road and built-up street canyon environments–A review." Atmospheric Environment 162 (2017): 71-86.

Akesaka, Mika, and Hitoshi Shigeoka. "“Invisible Killer”: Seasonal Allergy and Accidents." Working paper (2022).

American College of Allergy, Asthma & Immunology. "Pine Tree Allergy." (2018). Accessed May 21, 2023. https://acaai.org/allergies/allergic-conditions/pine-tree-allergy/.

Araujo, Rafael. "When clouds go dry: an integrated model of deforestation, rainfall, and agriculture." Working paper (2023).

Baldauf, Richard. "Roadside vegetation design characteristics that can improve local, near-road air quality." Transportation Research Part D: Transport and Environment 52 (2017): 354-361.

Bensnes, Simon Søbstad. "You sneeze, you lose: The impact of pollen exposure on cognitive performance during high-stakes high school exams." Journal of Health Economics 49 (2016): 1-13.

Biedermann, T., L. Winther, S. J. Till, P. Panzner, A. Knulst, and E. Valovirta. "Birch pollen allergy in Europe." Allergy 74, no. 7 (2019): 1237-1248.

Browne, Oliver, Ludovica Gazze, Michael Greenstone, and Olga Rostapshova. Man vs. Machine: Technological Promise and Political Limits of Automated Regulation Enforcement. No. w30816. National Bureau of Economic Research, 2023.

Brunekreef, Bert, Gerard Hoek, Paul Fischer, and Frits Th M. Spieksma. "Relation between airborne pollen concentrations and daily cardiovascular and respiratory-disease mortality." The Lancet 355, no. 9214 (2000): 1517-1518.

Chalfin, Aaron, Shooshan Danagoulian, and Monica Deza. "More sneezing, less crime? Health shocks and the market for offenses." Journal of health economics 68 (2019): 102230.

Charpin, D., Michel Calleja, C. Lahoz, Christian Pichot, and Y. Waisel. "Allergy to cypress pollen." Allergy 60, no. 3 (2005): 293-301.

Currie, Janet, and Reed Walker. "Traffic congestion and infant health: Evidence from E-ZPass." American Economic Journal: Applied Economics 3, no. 1 (2011): 65-90.

D'Amato, Gennaro, Stephen T. Holgate, Ruby Pawankar, Dennis K. Ledford, Lorenzo Cecchi, Mona Al-Ahmad, Fatma Al-Enezi et al. "Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization." World Allergy Organization Journal 8, no. 1 (2015): 1-52.

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.

Deschenes, 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): 2958-2989.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A. Turner. "Subways and urban air pollution." American Economic Journal: Applied Economics 14, no. 1 (2022): 164-96.

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.

Giles, Cynthia. *Next generation compliance: environmental regulation for the modern era*. Oxford University Press. (2022).

Glaeser, Edward L., and Matthew E. Kahn. "The greenness of cities: Carbon dioxide emissions and urban development." Journal of Urban Economics 67, no. 3 (2010): 404-418.

Gray, Wayne B., and Jay P. Shimshack. "The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence." Review of Environmental Economics and Policy (2011).

Greenstone, Michael, and Claire Qing Fan. "Introducing the air quality life index." AQLI Annual Report (2018).

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. "Can technology solve the principal-agent problem? Evidence from China's war on air pollution." American Economic Review: Insights 4, no. 1 (2022): 54-70.

Greenstone, Michael, Guojun He, Shanjun Li, and Eric Yongchen Zou. "China's war on pollution: Evidence from the first 5 years. " Review of Environmental Economics and Policy 15, no.2 (2021): 281-299.

Greiner, Alexander N., Peter W. Hellings, Guiseppina Rotiroti, and Glenis K. Scadding. "Allergic rhinitis." The Lancet 378, no. 9809 (2011): 2112-2122.

Grosset, Florian and Papp, Anna and Taylor, Charles, Rain Follows the Forest: Land Use Policy, Climate Change, and Adaptation (May 5, 2023). Available at SSRN: https://ssrn.com/abstract=4333147 or http://dx.doi.org/10.2139/ssrn.4333147

Han, Lu, Stephan Heblich, Christopher Timmins, and Yanos Zylberberg. "Cool Cities: The Value of Green Infrastructure." Working paper. (2021).

Harris, Tanner B., and William J. Manning. "Nitrogen dioxide and ozone levels in urban tree canopies." Environmental Pollution 158, no. 7 (2010): 2384-2386.

Huang, Jialin, Jianwei Xing, and Eric Yongchen Zou. "(Re)scheduling pollution exposure: The case of surgery schedules." Journal of Public Economics 219 (2023): 104825.

Jáuregui, Ignacio, Joaquin Mullol, Ignacio Dávila, Marta Ferrer, J. Bartra, A. Del Cuvillo, J. Montoro, J. Sastre, and A. Valero. "Allergic rhinitis and school performance." J Investig Allergol Clin Immunol 19, no. Suppl 1 (2009): 32-39.

Jones, Benjamin A., and Andrew L. Goodkind. "Urban afforestation and infant health: Evidence from MillionTreesNYC." Journal of Environmental Economics and Management 95 (2019): 26-44.

Kahn, Matthew E., and Randall Walsh. "Cities and the Environment." Handbook of Regional and Urban Economics 5 (2015): 405-465.

Kumar, Prashant, et al. "The nexus between air pollution, green infrastructure and human health." Environment International 133 (2019): 105181.

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, Liqing. "Environmental goods provision and gentrification: Evidence from MillionTreesNYC." Journal of Environmental Economics and Management (2023): 102828.

Lyapustin, Alexei, Yujie Wang, Sergey Korkin, and Dong Huang. "MODIS collection 6 MAIAC algorithm." Atmospheric Measurement Techniques 11, no. 10 (2018): 5741-5765.

Marcotte, Dave E. "Allergy test: Seasonal allergens and performance in school." Journal of Health Economics 40 (2015): 132-140.

Mastini, Riccardo, Giorgos Kallis, and Jason Hickel. "A green new deal without growth?" Ecological Economics 179 (2021): 106832.

McDonald, A. G., W. J. Bealey, D. Fowler, U. Dragosits, U. Skiba, R. I. Smith, R. G. Donovan, H. E. Brett, C. N. Hewitt, and E. Nemitz. "Quantifying the effect of urban tree planting on concentrations and depositions of PM10 in two UK conurbations." Atmospheric Environment 41, no. 38 (2007): 8455-8467.

Meltzer, Eli O., Michael S. Blaiss, Robert M. Naclerio, Stuart W. Stoloff, M. Jennifer Derebery, Harold S. Nelson, John M. Boyle, and Mark A. Wingertzahn. "Burden of allergic rhinitis: allergies in America, Latin America, and Asia-Pacific adult surveys." In Allergy and Asthma Proceedings, vol. 33, no. 5, pp. S113-S141. OceanSide Publications, Inc, 2012.

Meng, Kyle C. "Using a free permit rule to forecast the marginal abatement cost of proposed climate policy." American Economic Review 107, no. 3 (2017): 748-784.

Nowak, David J. "Air pollution removal by Chicago's urban forest." Chicago's urban forest ecosystem: Results of the Chicago urban forest climate project (1994): 63-81.

Nowak, David J., Daniel E. Crane, and Jack C. Stevens. "Air pollution removal by urban trees and shrubs in the United States." Urban Forestry & Urban Greening 4, no. 3-4 (2006): 115-123.

Nowak, David J., and Eric J. Greenfield. "US urban forest statistics, values, and projections." Journal of Forestry 116, no. 2 (2018): 164-177.

Nowak, David J., and Eric J. Greenfield. "The increase of impervious cover and decrease of tree cover within urban areas globally (2012–2017)." Urban Forestry & Urban Greening 49 (2020): 126638.

Pettorelli, Nathalie, Jon O. Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J. Tucker, and Nils C. Stenseth. "Using the satellite-derived NDVI to assess ecological responses to environmental change." Trends in Ecology & Evolution 20, no.9 (2005): 503-510.

Rojas-Rueda, David, Mark J. Nieuwenhuijsen, Mireia Gascon, Daniela Perez-Leon, and Pierpaolo Mudu. "Green spaces and mortality: a systematic review and meta-analysis of cohort studies." The Lancet Planetary Health 3, no. 11 (2019): e469-e477.

Sävje, Fredrik, Peter Aronow, and Michael Hudgens. "Average treatment effects in the presence of unknown interference." Annals of Statistics 49, no. 2 (2021): 673.

Sierra-Heredia, Cecilia, Michelle North, Jeff Brook, Christina Daly, Anne K. Ellis, Dave Henderson, Sarah B. Henderson, Éric Lavigne, and Tim K. Takaro. "Aeroallergens in Canada: distribution, public health impacts, and opportunities for prevention." International Journal of Environmental Research and Public Health 15, no. 8 (2018): 1577.

Sun, Xuezheng, Anna Waller, Karin B. Yeatts, and Lauren Thie. "Pollen concentration and asthma exacerbations in Wake County, North Carolina, 2006–2012." Science of the Total Environment 544 (2016): 185-191.

Thacker, Scott, Daniel Adshead, Marianne Fay, Stéphane Hallegatte, Mark Harvey, Hendrik Meller, Nicholas O'Regan, Julie Rozenberg, Graham Watkins, and Jim W. Hall. "Infrastructure for sustainable development." Nature Sustainability 2, no. 4 (2019): 324-331.

Xia, Fan, Jianwei Xing, Jintao Xu, and Xiaochuan Pan. "The short-term impact of air pollution on medical expenditures: Evidence from Beijing." Journal of Environmental Economics and Management 114 (2022): 102680.

Yang, Jie, and Xin Huang. "The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. " Earth System Science Data 13, no. 8 (2021): 3907-3925.

Yin, Shan, Zhemin Shen, Pisheng Zhou, Xiaodong Zou, Shengquan Che, and Wenhua Wang. "Quantifying air pollution attenuation within urban parks: An experimental approach in Shanghai, China." Environmental Pollution 159, no. 8-9 (2011): 2155-2163.

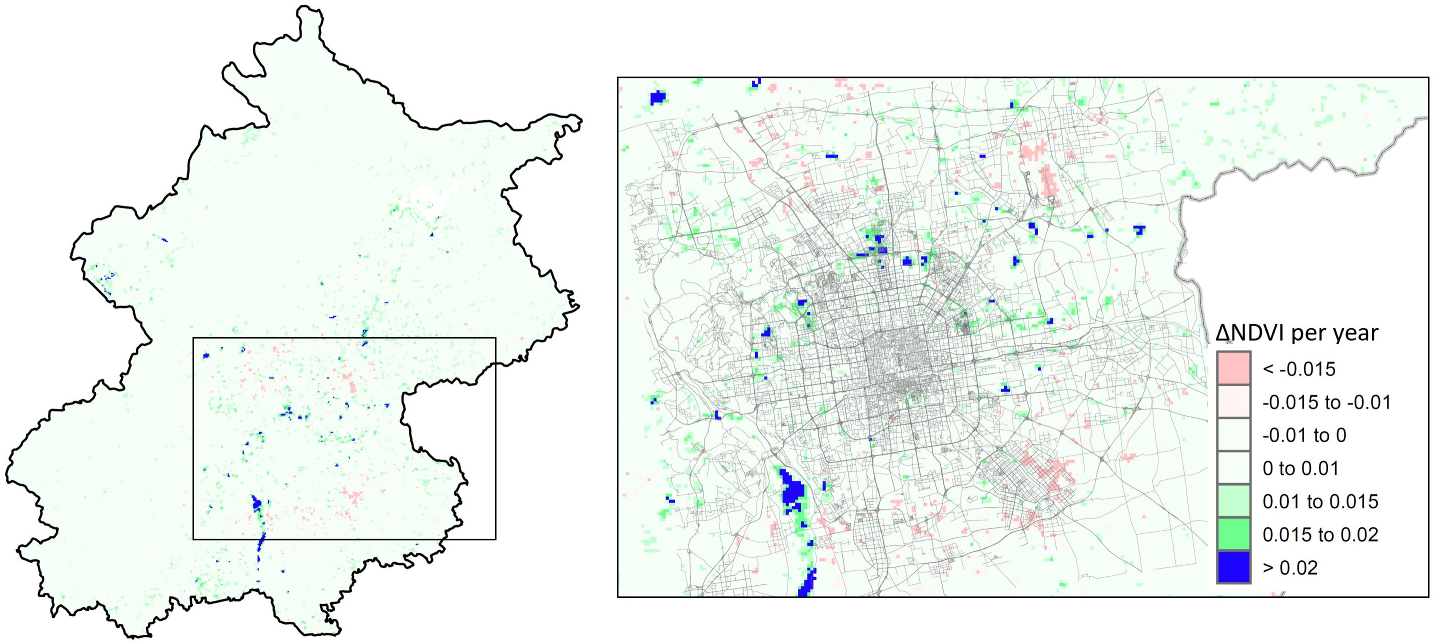
Ziska, Lewis H., László Makra, Susan K. Harry, Nicolas Bruffaerts, Marijke Hendrickx, Frances Coates, Annika Saarto et al. "Temperature-related changes in airborne allergenic pollen abundance and seasonality across the northern hemisphere: a retrospective data analysis." The Lancet Planetary Health 3, no. 3 (2019): e124-e131.

Zhao, Dejingyu, Caihua Ye, Yufei Wang, and Yifeng Yao. "京津冀地区气传花粉数据分析." 植物学报 56, no. 6 (2021): 751.

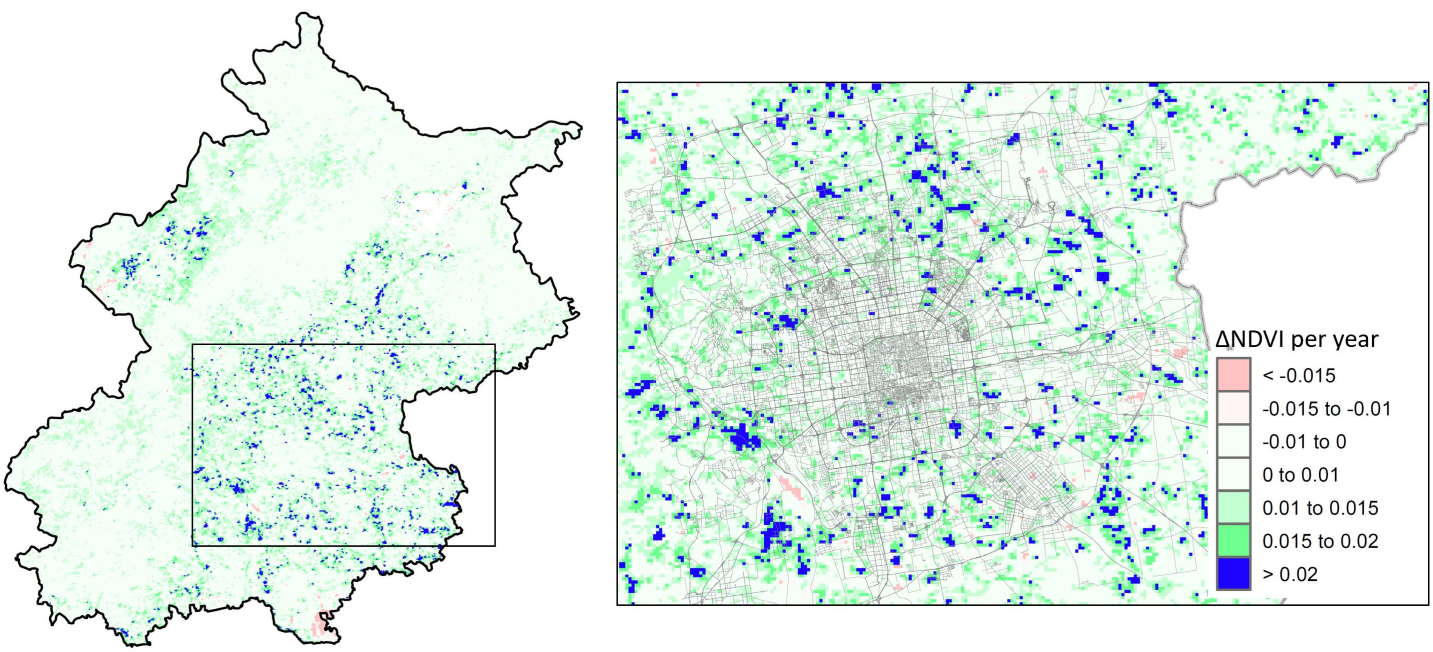
Zheng, Siqi, and Matthew E. Kahn. "Understanding China's urban pollution dynamics." Journal of Economic Literature 51, no. 3 (2013): 731-772.

Figure 1. The Greening Up of Beijing

(a) Rate of vegetation growth, 2001-2011



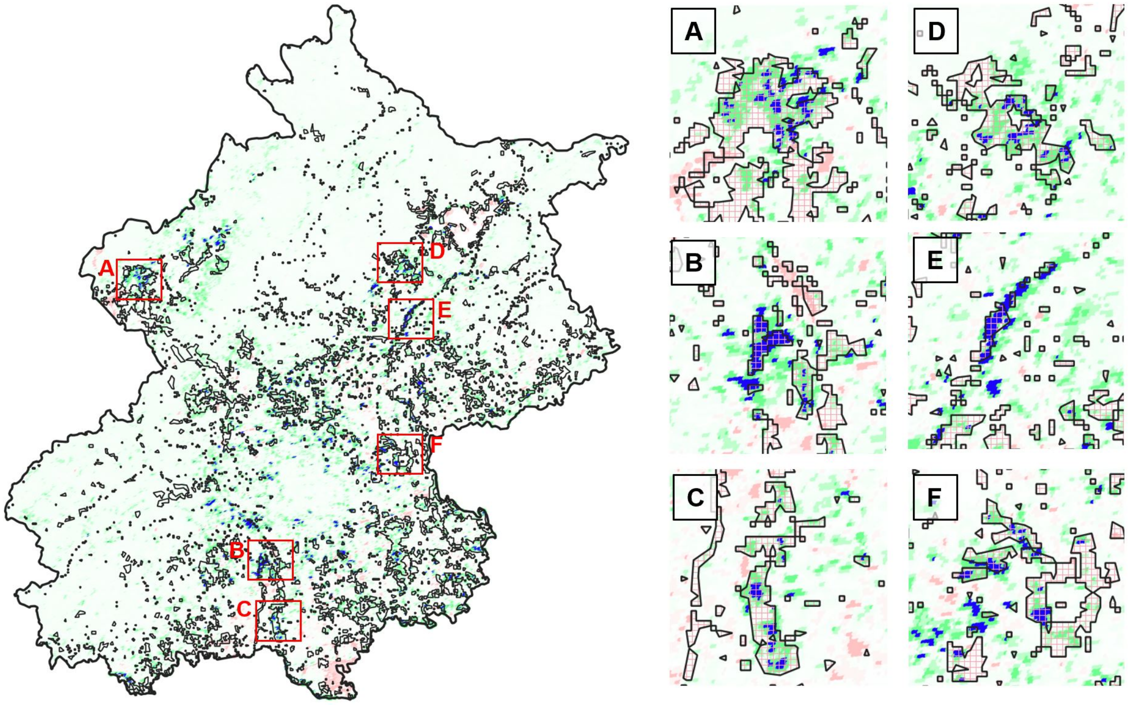
(b) Rate of vegetation growth, 2012-2020



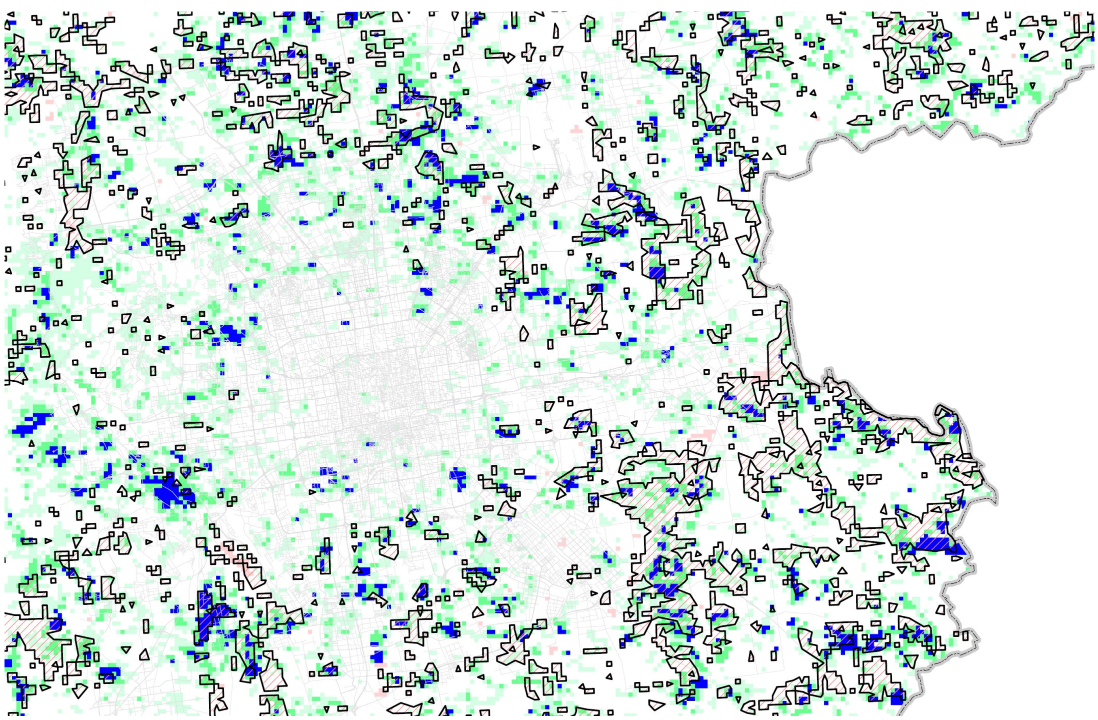
Notes: This figure shows the grid-level annual rate of NDVI growth for the 2001-2011 period (panel a) and the 2012-2020 period (panel b). The left side of each panel shows the conditions for the entire prefecture city of Beijing. A zoomed-in view of the city center is provided on the right side of the panel. Gray lines within the zoomed-in view represent the road networks.

Figure 2. The Million *Mu* Project (MMP) and Vegetation Growth

(a) Location of MMP planting sites



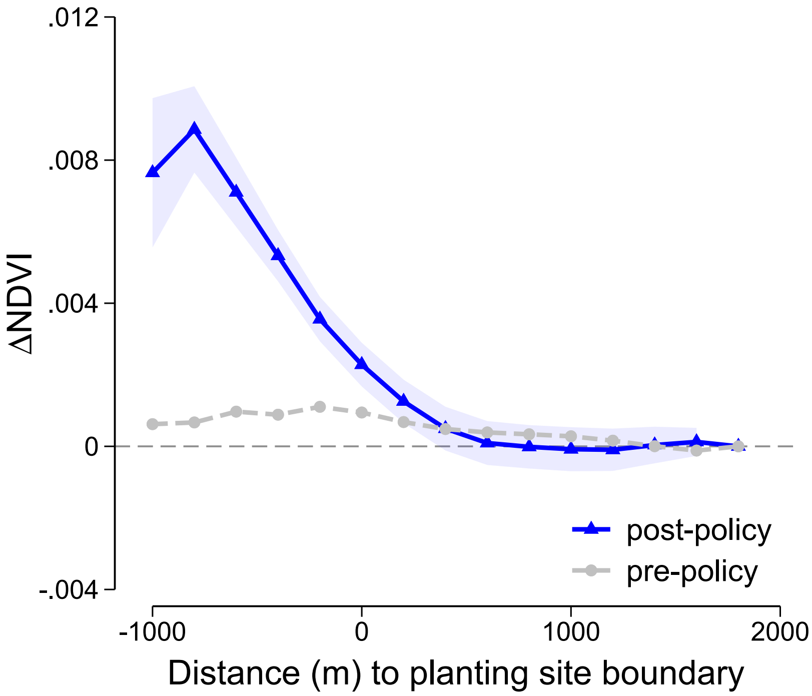
(b) Location of MMP planting sites (city center)



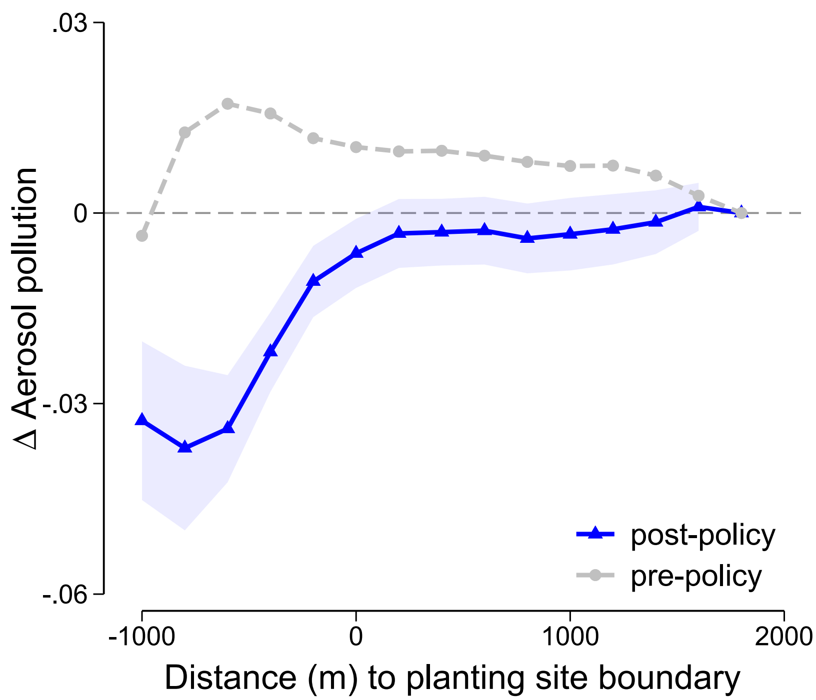
Notes: Color indicates the annual rate of growth of NDVI between 2012-2020, with blue indicating strong growth. Polygons represent MMP planting sites. Panel (a) shows the condition for the entire prefecture city of Beijing, and six example zoom-in areas. Panel (b) provides a zoomed-in view of the city center.

Figure 3. Changes in Vegetation and Air Quality Near Planting Sites

(a). Vegetation index (satellite)



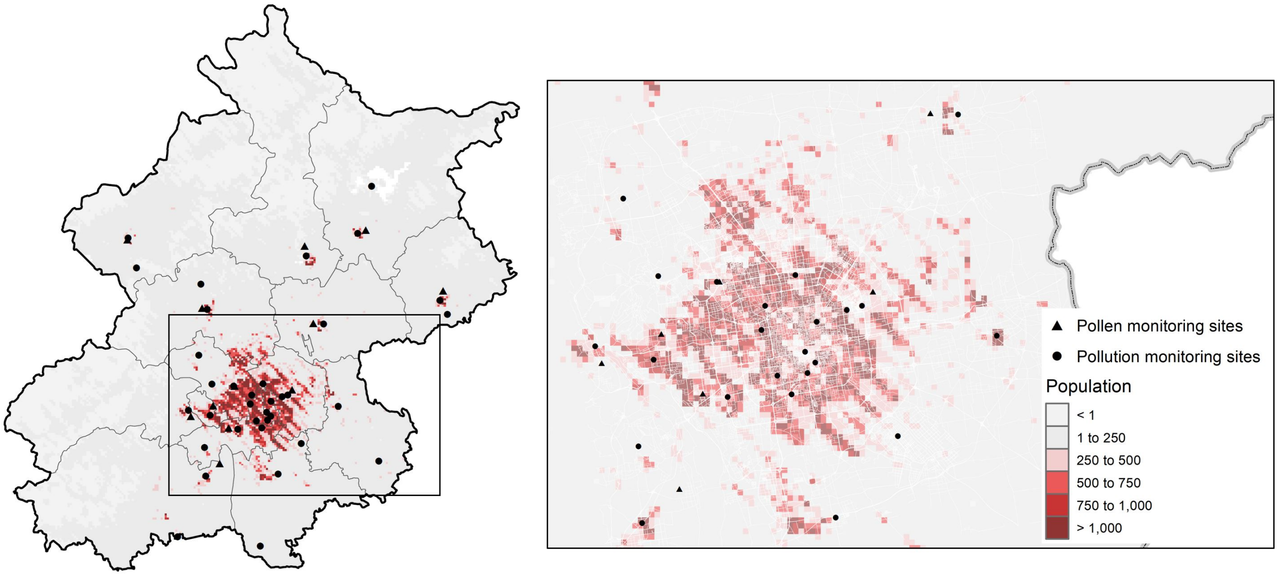
(b). Aerosol pollution (satellite)



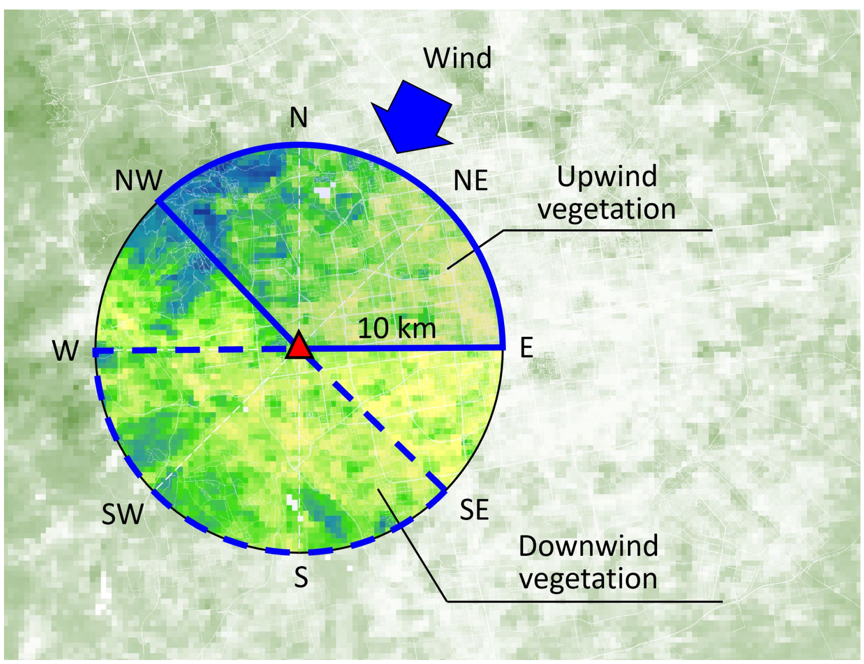
Notes: This figure shows the annual rate of change in NDVI (panel a) and AOD (panel b) as a function of the distance to the nearest MMP planting site boundary. The dashed line represents the pre-MMP period of 2001-2011, and the solid line corresponds to the post-MMP period of 2012-2020. These estimates are derived from grid-level, cross-sectional regressions of the annual rate of change on a series of dummy variables indicating distance bins (200-meter increments), with the 1800-2000 meter bin as the omitted category. The estimation is done separately for the pre- and post-MMP periods. For the post-MMP estimates, range plots display the 95% confidence interval, constructed using 1-km grid cluster bootstrap standard errors.

Figure 4. Urban Center Effects: Illustration of Empirical Strategy

(a) Population representation of pollution and pollen monitors



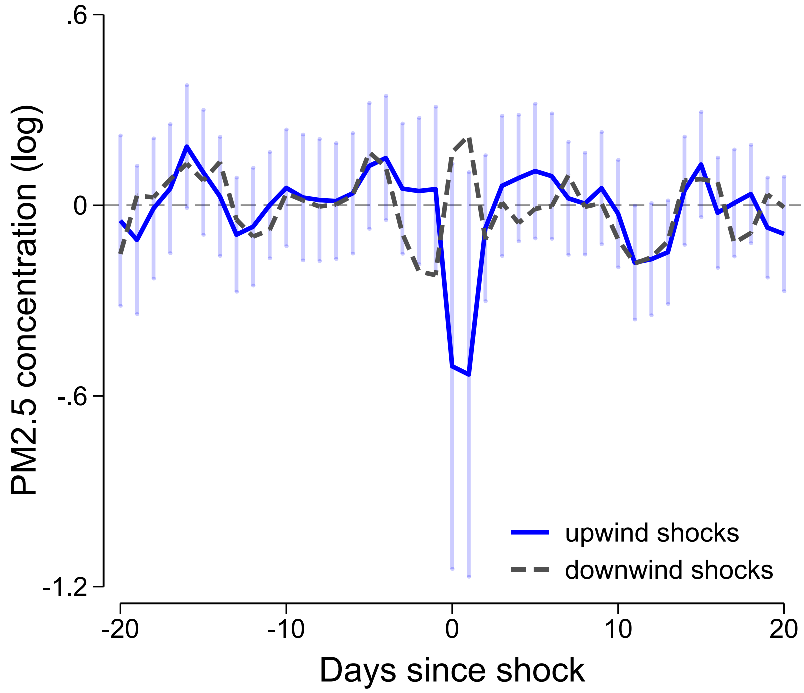
(b) Illustration of Empirical Strategy



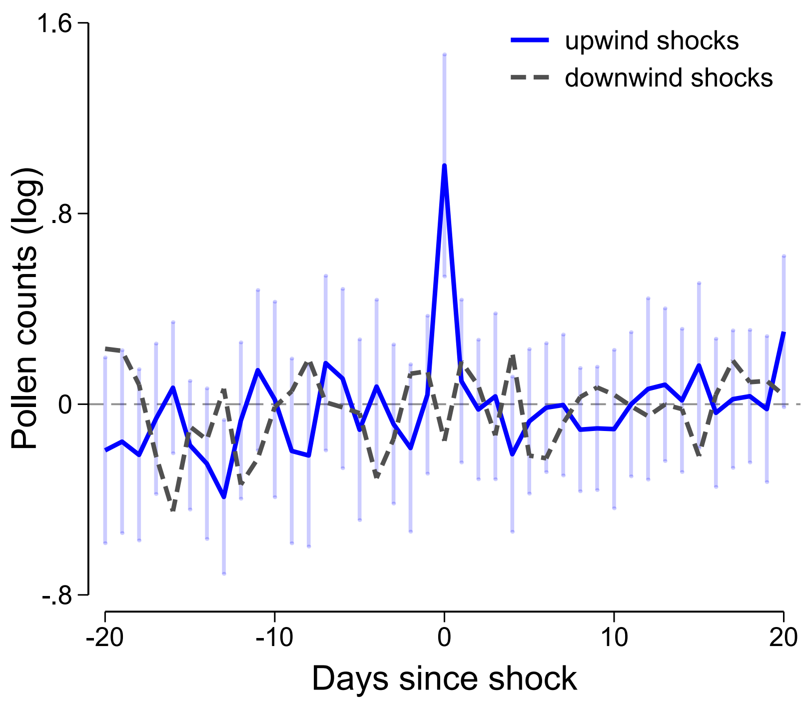
Notes: Panel (a) shows the location of in-situ air pollution monitors (dots) and pollen monitors (triangles), overlaid with grid-level estimates of the population as of the year 2010. Panel (b) provides an illustration of the identification strategy used to estimate the causal effect of vegetation on pollution and pollen. Consider a specific location of a pollution or pollen monitor, represented by the red triangle. The blue arrow indicates the prevailing wind direction for a given day. "Upwind vegetation" is defined as the average NDVI index across all grids that fall within a 135-degree cone in the upwind direction of the monitor on that day. "Downwind vegetation" is defined similarly but using the downwind direction.

Figure 5. Urban Center Effects: Upwind Shocks, Particulates Pollution, and Pollen Counts

(a). Particulate matter (monitors)



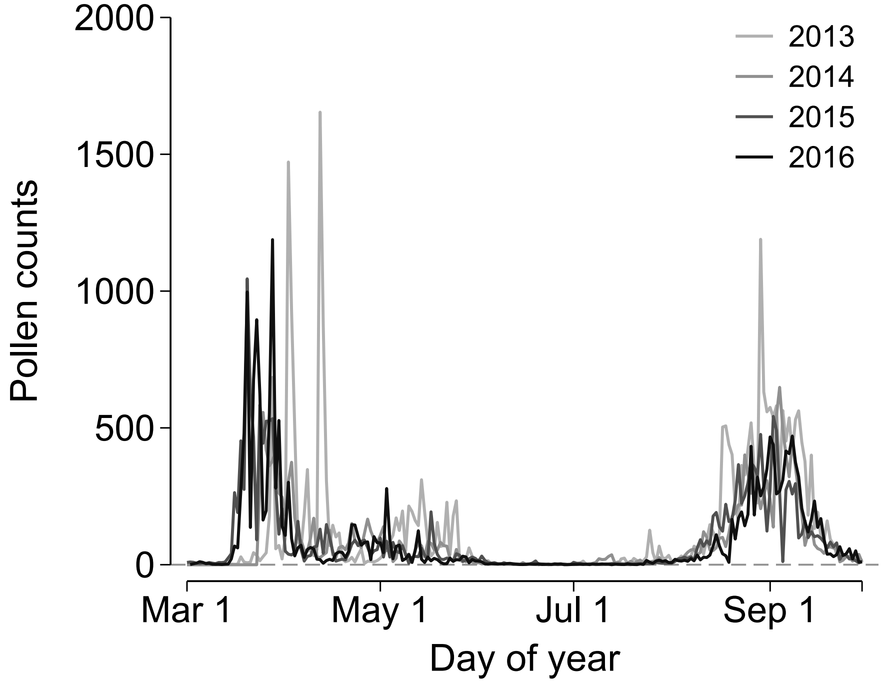
(b). Pollen counts (monitors)



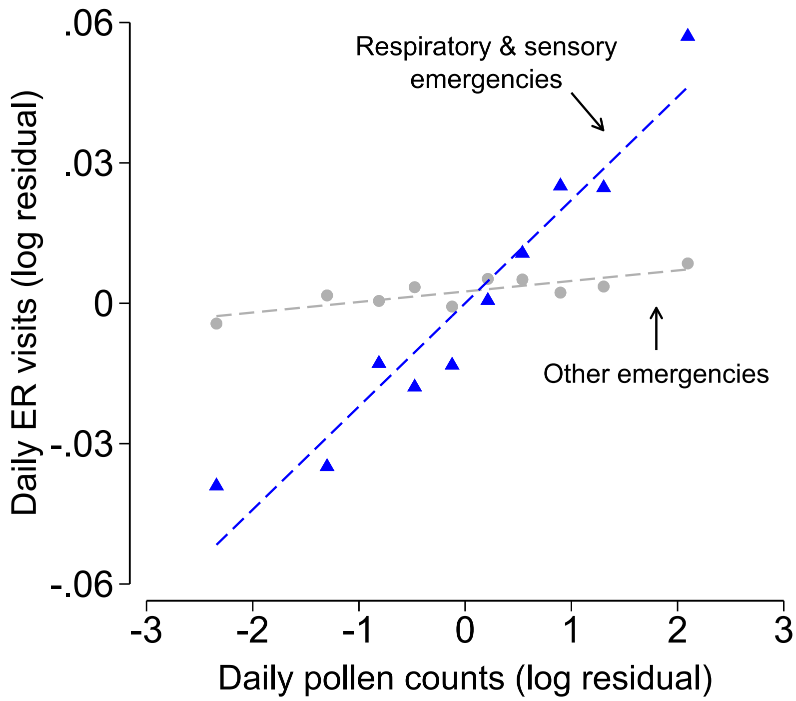
Notes: This figure shows regression coefficients of air pollution (panel a) and pollen counts (panel b) on 20 lead, contemporaneous, and 20 lag terms of upwind vegetation shocks. We focus on fine particulate matter pollution ($PM_{2.5}$) in panel (a), and we report results for other criteria air pollutants (PM_{10}, O_3, NO_2, SO_2, and CO) in the Online Appendix. In each panel, a set of placebo coefficients is also displayed, obtained by running the same regression but replacing upwind vegetation shocks with downwind shocks. Range bars show 95% confidence intervals constructed using standard errors two-way clustered at both the monitor and the day-of-sample level.

Figure 6. Health Effects: Daily Pollen Exposure and Emergency Room Visits

(a). Day-to-day fluctuations in pollen counts



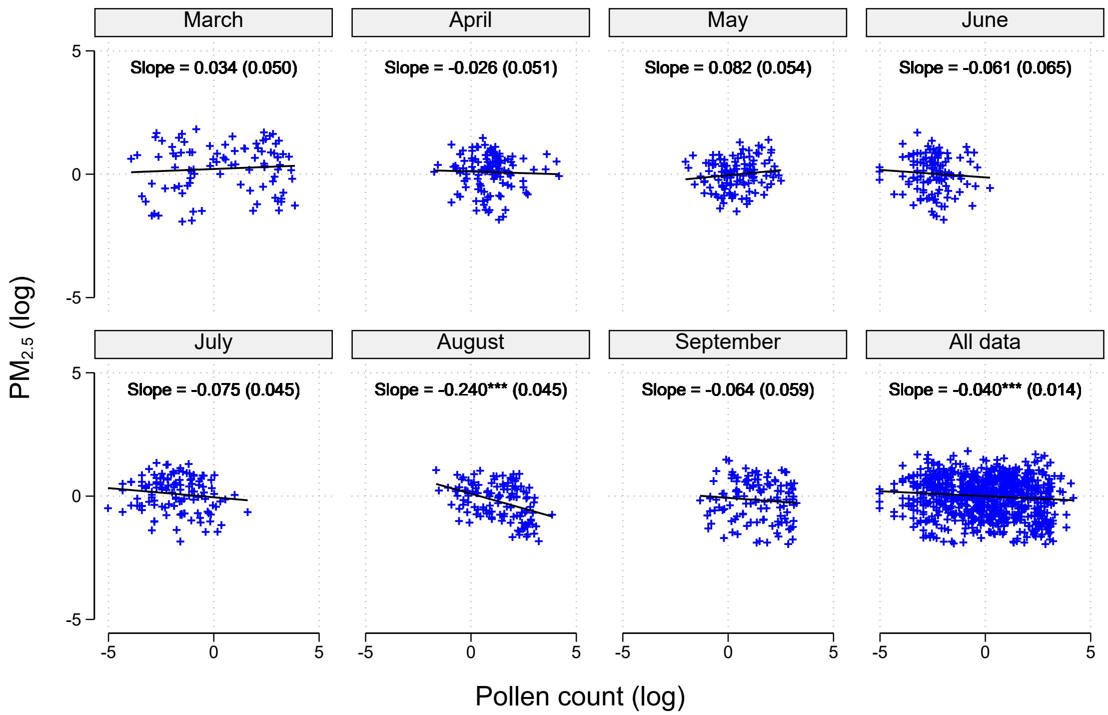
(b). Pollen shocks and emergency room visits



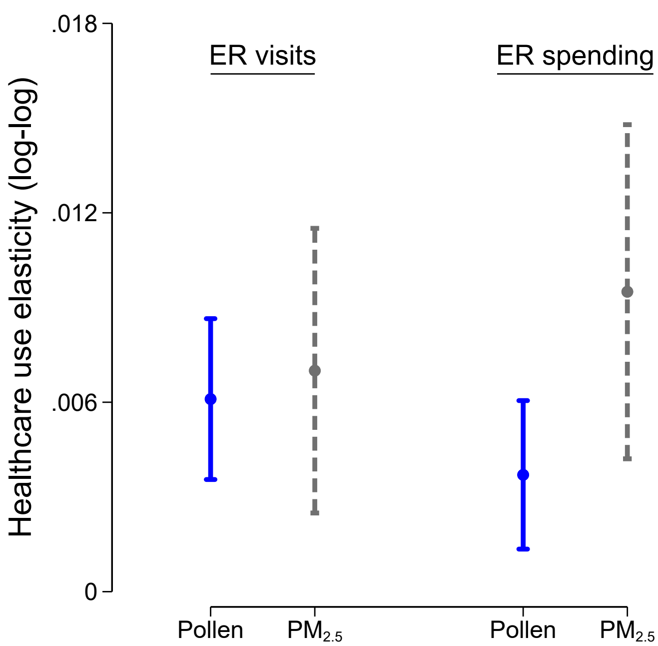
Notes: Panel (a) shows the average pollen counts per day-of-year at the site level for each year of data during the pollen monitoring season (March 1 to October 15). Panel (b) is a bin scatterplot on the relationship between fixed effects-residualized log daily emergency room visits and residualized log daily pollen counts. The patterns are separated for respiratory and sensory emergencies versus all other emergencies.

Figure 7. Comparing Health Effects: Pollen vs. Fine Particulate Matter ($PM_{2.5}$)

(a). Independent variation between pollen and $PM_{2.5}$



(b). ER effects of Pollen vs. $PM_{2.5}$



Notes: Panel (a) shows the correlation between log $PM_{2.5}$ concentration and log pollen count at the district-day level. Each cross represents a district-day observation, and the black line is the superimposed OLS regression line, with slope and standard error estimates reported. Regressions are done separately for each month-of-year and for the pooled sample. Notice that data are only shown for March-September which is the pollen-monitoring season. Panel (b) shows coefficient estimates from four separate regression of an ER outcome (log visits on the left and log spending on the right) on log pollen or log $PM_{2.5}$. All regressions control for district-by-month fixed effects, year-by-month fixed effects, day-of-week fixed effects, and holiday fixed effects. Range plots show 95% confidence intervals constructed used standard errors two-way clustered at the district and day-of-sample levels.

Table 1. Summary of Data Sources

Variable	(1) Data source	(2) Spatial frequency	(3) Temporal frequency	(4) Collection methods	(5) Used in
MMP planting site	City government	N/A	N/A	Survey	Section 4
Land use	RESDC	30 m	2000, 2010, 2020	Remote-sensing	Section 4
Population	WorldPop	100m	2010	Remote-sensing	Section 4
AOD	NASA MODIS	1 km	Daily, 2001-2020	Remote-sensing	Section 4
NDVI	NASA MODIS	250 m	16-day, 2001-2020	Remote-sensing	Sections 4 & 5
Wind	City weather bureau	20 stations	Daily, 2001-2019	In-situ sampling	Section 5
Pollution	City environmental agency	35 stations	Daily, 2014-2019	In-situ sampling	Sections 5 & 6
Pollen counts	City weather bureau	12 station	Daily, 2013-2016	In-situ sampling	Sections 5 & 6
Emergency room visits	City medical bureau	All hospitals	Records, 2013-2016	Administrative	Section 6

Notes: This table tabulates key variables used, source of data (column 1), spatial and temporal resolution of the raw data (columns 2-3), collection methods of the raw data (column 4), and where the variables are mainly used in the paper (column 5).

Table 2. Pollen Exposure and Emergency Room (ER) Utilization

	(1) All causes	(2) Respiratory	(3) Sensory	(4) Others	(5) All causes	(6) Respiratory	(7) Sensory	(8) Others
	ER visits				ER spending			
A. Baseline specification								
All ER	0.0061***	0.0214***	0.0424***	0.0034**	0.0037***	0.0161***	0.0429***	0.0009
	(0.0013)	(0.0038)	(0.0056)	(0.0012)	(0.0012)	(0.0033)	(0.0117)	(0.0013)
B. Effects by severity								
ER → Not hospitalized	0.0059***	0.0214***	0.0424***	0.0031**	0.0036**	0.0164***	0.0432***	0.0008
	(0.0013)	(0.0038)	(0.0056)	(0.0013)	(0.0012)	(0.0036)	(0.0114)	(0.0013)
ER → Hospitalized	0.0019	0.0043	-0.0025	-0.0002	0.0029	0.0207	-0.0314	-0.0006
	(0.0063)	(0.0057)	(0.0031)	(0.0056)	(0.0310)	(0.0470)	(0.0403)	(0.0318)
C. Effects by age								
Age < 60	0.0067***	0.0242***	0.0440***	0.0037**	0.0041*	0.0211***	0.0450***	0.0013
	(0.0015)	(0.0045)	(0.0060)	(0.0013)	(0.0019)	(0.0043)	(0.0118)	(0.0017)
Age ≥ 60	0.0044**	0.0149***	0.0176**	0.0023	0.0038*	0.0141**	0.0594*	0.0006
	(0.0014)	(0.0042)	(0.0059)	(0.0015)	(0.0020)	(0.0056)	(0.0295)	(0.0023)
D. Effects by prior condition								
Prior respiratory-sensory visits	0.0310***	0.0410***	0.0395***	0.0258***	0.0231***	0.0688**	0.1609***	0.0249**
	(0.0042)	(0.0082)	(0.0080)	(0.0048)	(0.0070)	(0.0297)	(0.0317)	(0.0083)
No prior visits	0.0057***	0.0191***	0.0353***	0.0033**	0.0036***	0.0140***	0.0386***	0.0012
	(0.0012)	(0.0036)	(0.0059)	(0.0012)	(0.0011)	(0.0030)	(0.0119)	(0.0013)

Notes: Each cell represents a separate regression using district-day level data. Each column presents ER records corresponding to different diagnoses. All outcomes are measured using a three-day look-ahead window (e.g., total number of ER visits today, tomorrow, and the day after tomorrow). Panel (a) uses all ER records. Panel (b) stratifies by visits that did and did not end up with hospital admissions. Panel (c) stratifies by age of the patient. Panel (d) stratifies by whether the patient had respiratory and sensory visits in the previous 30 days. All regressions control for district-by-month fixed effects, year-by-month fixed effects, day-of-week fixed effects, and holiday fixed effects. Standard errors are two-way clustered at the district and day-of-sample levels. Number of observation for each regression is 8,394. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table 3. Pollen Exposure and Emergency Room (ER) Utilization: Effects by Pollen Species

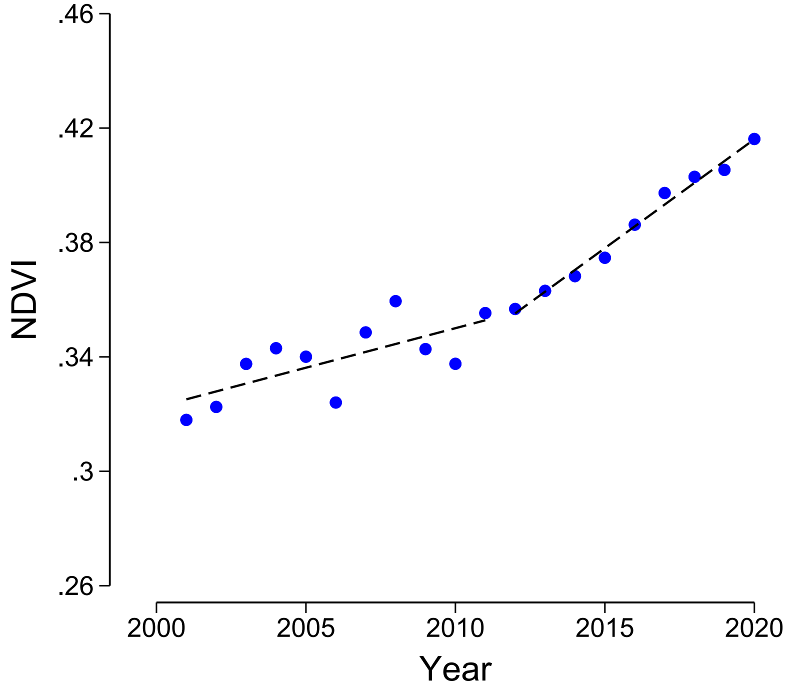
	(1) All causes	(2) Respiratory	(3) Sensory	(4) Others	(5) All causes	(6) Respiratory	(7) Sensory	(8) Others
	ER visits				ER spending			
Cupressaceae (Cypress)	0.0069***	0.0216***	0.0519***	0.0042*	0.0022	0.0134***	0.0552***	-0.0010
	(0.0019)	(0.0034)	(0.0065)	(0.0019)	(0.0018)	(0.0038)	(0.0098)	(0.0015)
Salicaceae (Willow)	0.0061**	0.0179***	0.0330***	0.0039	0.0001	0.0085	0.0310	-0.0018
	(0.0026)	(0.0034)	(0.0090)	(0.0027)	(0.0033)	(0.0065)	(0.0244)	(0.0031)
Pinaceae (Pine)	-0.0017	0.0024	0.0025	-0.0023	-0.0008	-0.0025	0.0044	-0.0005
	(0.0022)	(0.0053)	(0.0071)	(0.0020)	(0.0026)	(0.0060)	(0.0169)	(0.0026)
Moraceae (Mulberry)	0.0059**	0.0225***	0.0549***	0.0029	0.0058**	0.0225**	0.0491*	0.0023
	(0.0024)	(0.0054)	(0.0081)	(0.0024)	(0.0025)	(0.0076)	(0.0238)	(0.0026)
Artemisia (Sagebrush/Wormwood)	0.0041*	0.0252***	0.0422***	0.0006	0.0034*	0.0267***	0.0254	-0.0012
	(0.0019)	(0.0056)	(0.0082)	(0.0017)	(0.0016)	(0.0074)	(0.0217)	(0.0015)
Chenopodiaceae (Goosefoot)	0.0010	0.0216***	0.0488***	-0.0024	-0.0001	0.0165**	0.0172	-0.0033
	(0.0022)	(0.0059)	(0.0088)	(0.0024)	(0.0017)	(0.0070)	(0.0218)	(0.0024)

Notes: Each cell represents a separate regression using district-day level data. Each column presents ER records corresponding to different diagnoses. All outcomes are measured using a three-day look-ahead window (e.g., total number of ER visits today, tomorrow, and the day after tomorrow). Each row corresponds to regression using a different right-hand-side measure of pollen species. All regressions control for district-by-month fixed effects, year-by-month fixed effects, day-of-week fixed effects, and holiday fixed effects. Standard errors are two-way clustered at the district and day-of-sample levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

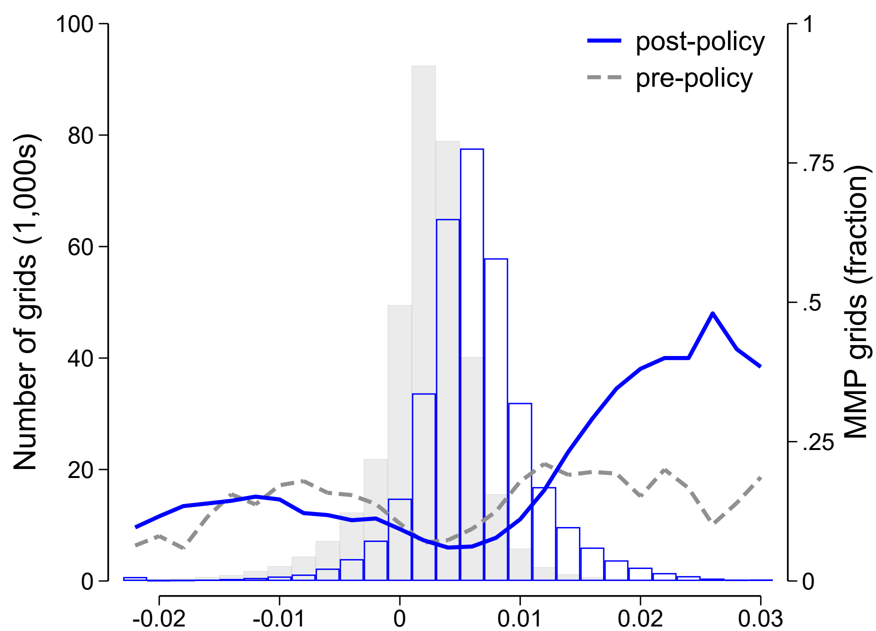
Appendix. Additional Figures and Tables

Figure A.1. Summary Statistics of Beijing's Vegetation Growth

(a). Annual trends



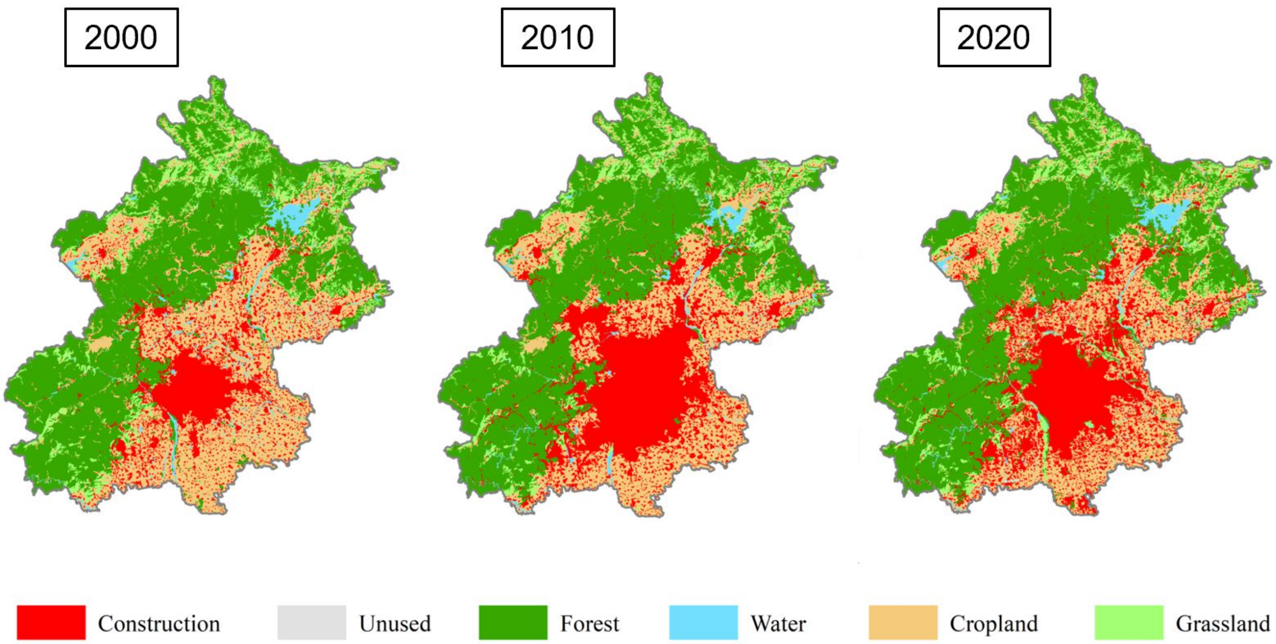
(b). Distribution of grid-level vegetation growth



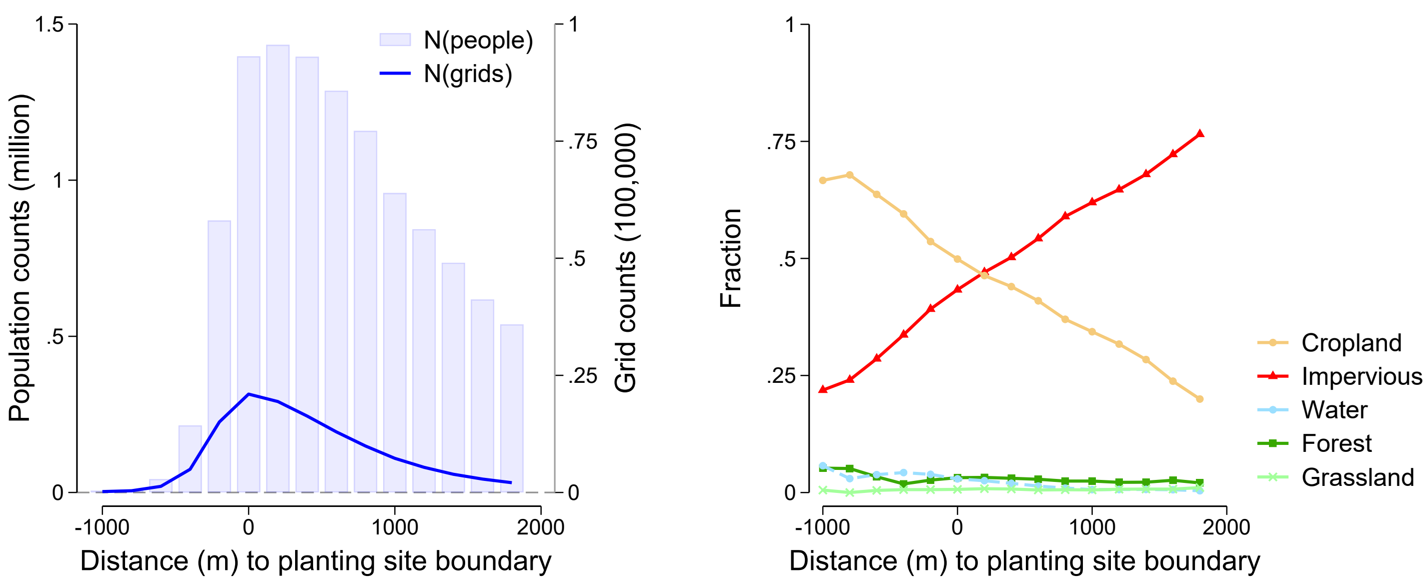
Notes: Panel (a) shows trends in annual average NDVI for the city of Beijing. Panel (b) shows a grid-level distribution of NDVI growth rate in the pre-policy (gray) and post-policy (blue) period. Lines show fraction of grids in the corresponding bins that are MMP during the pre-policy (dashed) and post-policy (solid) period.

Figure A.2. Summary Statistics of Land Use

(a). Land use changes

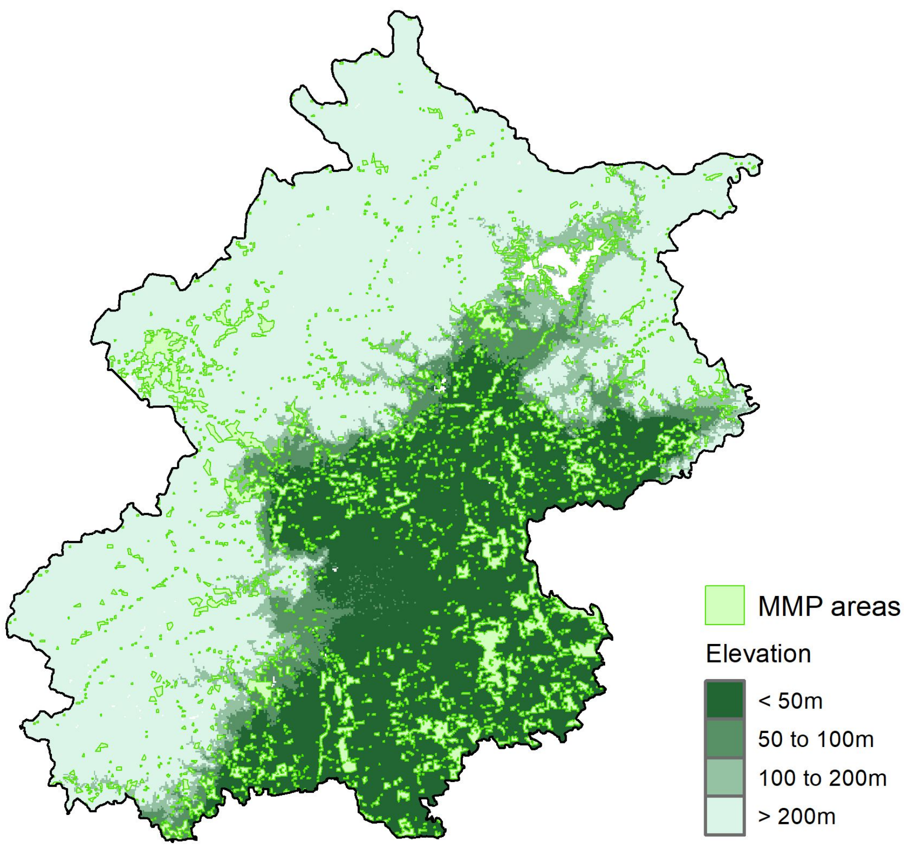


(b). MMP Planting Sites and Baseline (2010) Land Use Categorization



Notes: Panel (a) shows land use categorization in 2000, 2010, and 2020. Panel B shows the distribution of population and grids (left) and the distribution of land use as of year 2010 (right) as a function of distance to the nearest MMP planting site boundary. The graph excludes a very small proportion of the "unused" category.

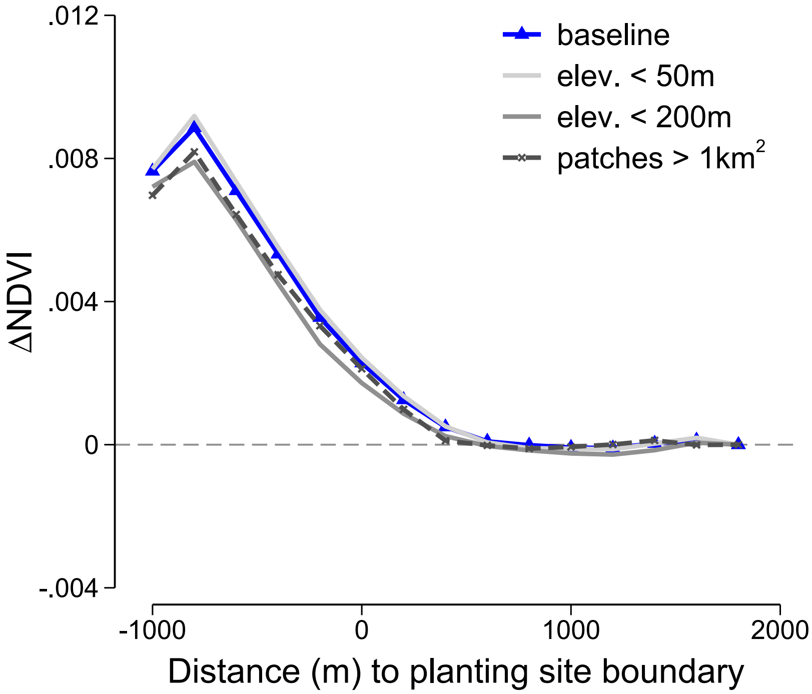
Figure A.3. Beijing Elevation and MMP Planting Sites



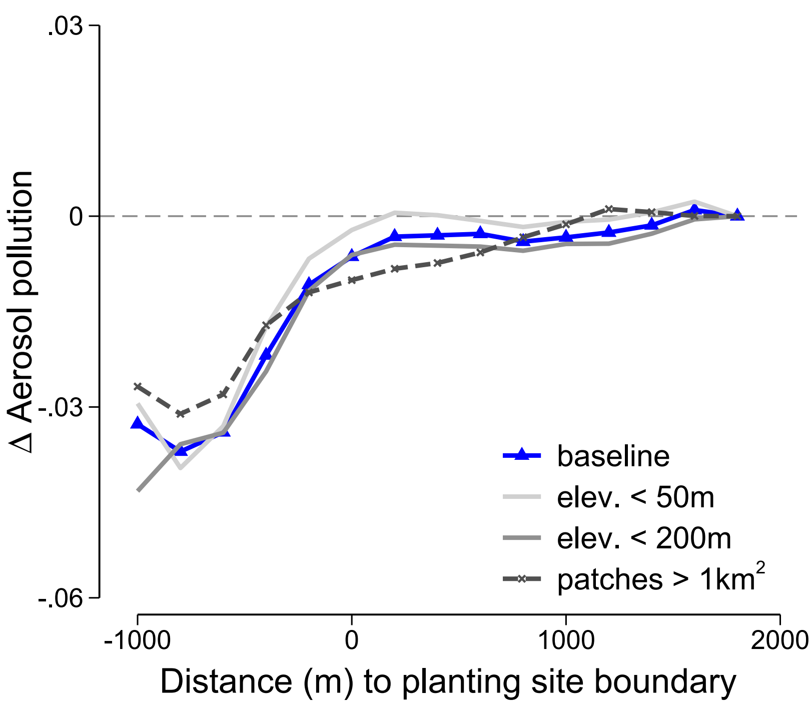
Notes: This map shows elevation of Beijing, overlaid with the location of the Million *Mu* Project (MMP) planting sites.

Figure A.4. Robustness: Changes in Vegetation and Air Quality Near Planting Sites

(a). Vegetation index (satellite)

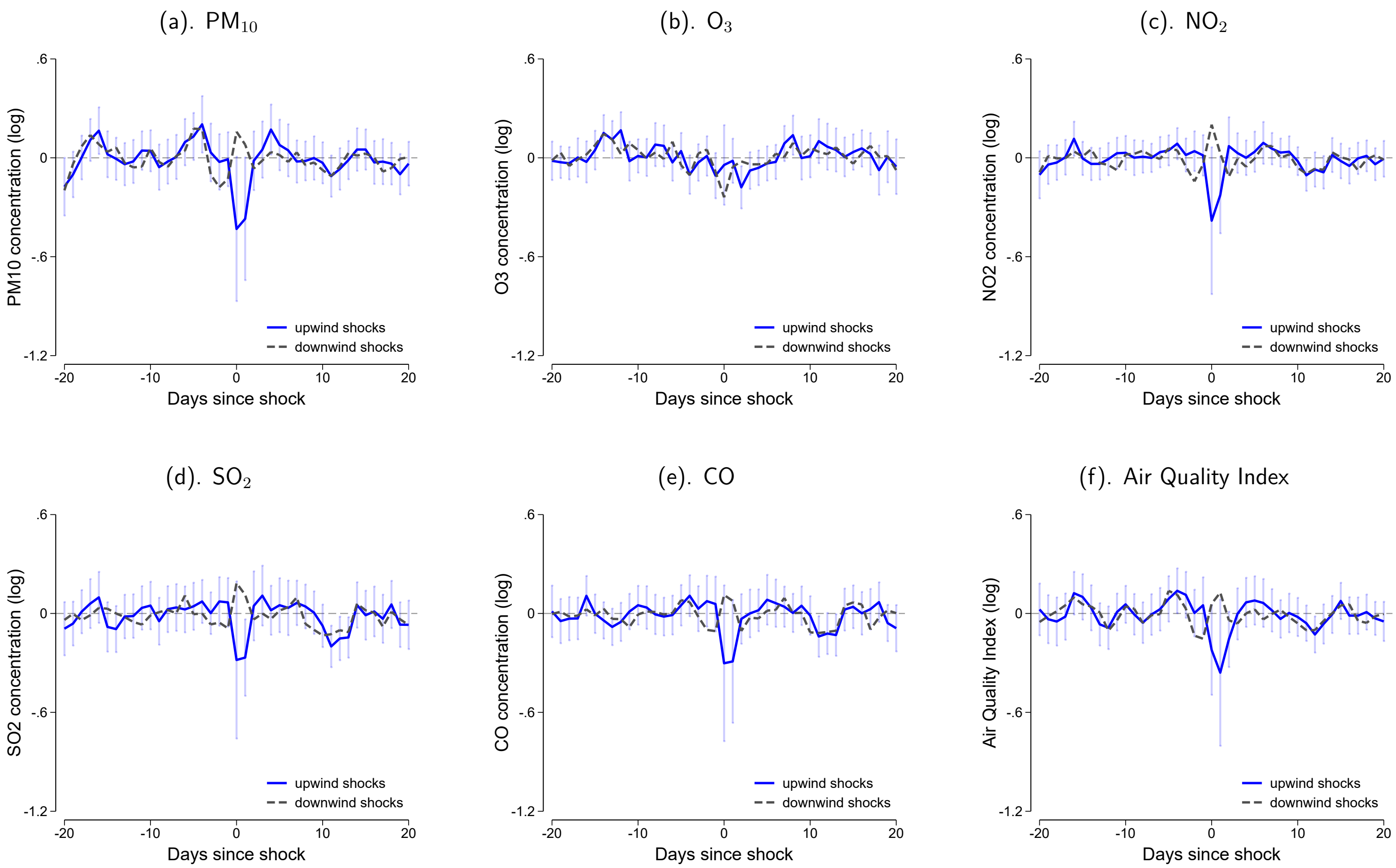


(b). Aerosol pollution (satellite)



Notes: This figure shows the annual rate of change in NDVI (panel a) and AOD (panel b) as a function of the distance to the nearest MMP planting site boundary for the post-MMP period of 2012-2020. Each line is from a separate robustness check. These estimates are derived from grid-level, cross-sectional regressions of the annual rate of change on a series of dummy variables indicating distance bins (200-meter increments), with the 1800-2000 meter bin as the omitted category.

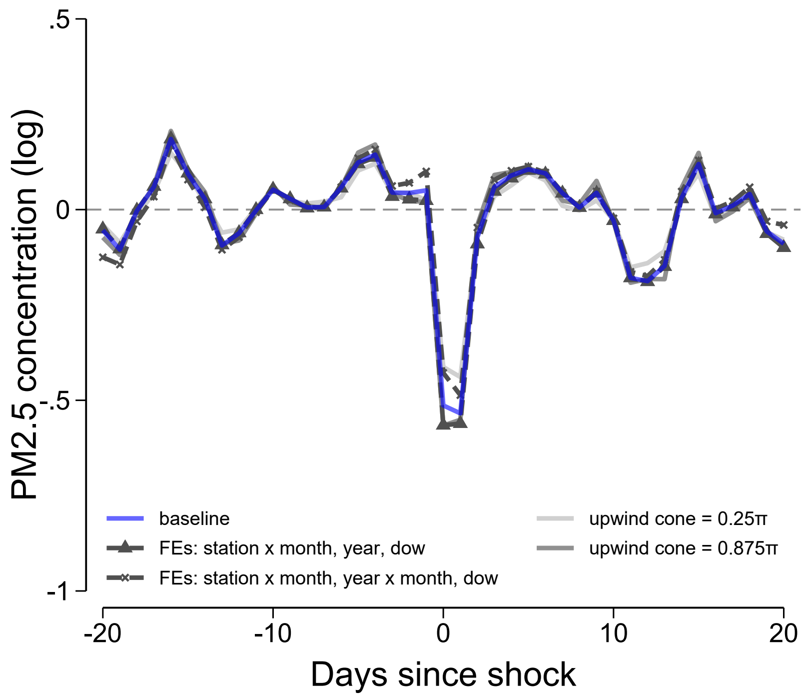
Figure A.5. The Effect of Upwind Shocks on Urban Center Pollution: Other Pollutants



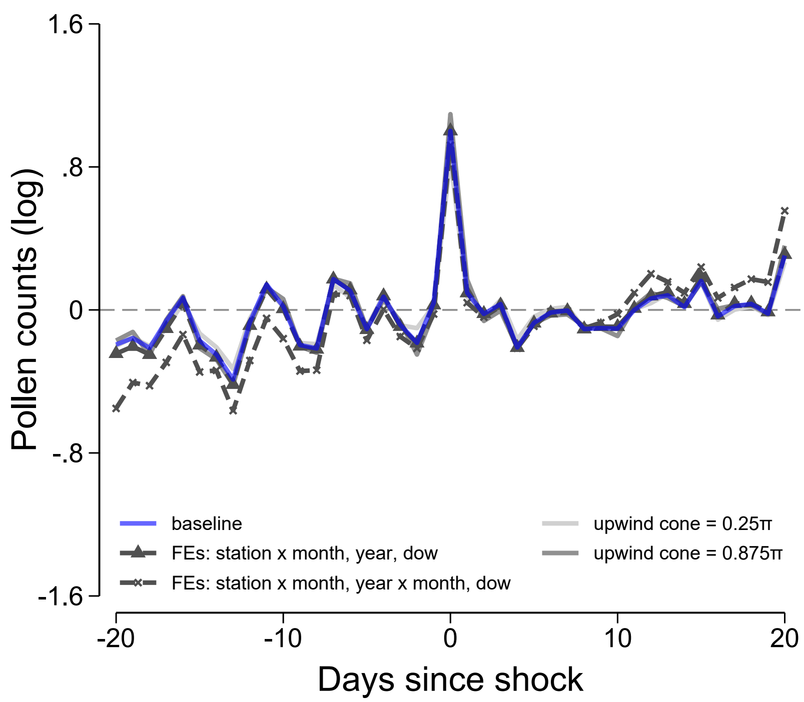
Notes: This figure shows regression coefficients of air pollution on 20 lead, contemporaneous, and 20 lag terms of upwind vegetation shocks. In each panel, a set of placebo coefficients is also displayed, obtained by running the same regression but replacing upwind vegetation shocks with downwind shocks. Range bars show 95% confidence intervals constructed using standard errors two-way clustered at both the monitor and the day-of-sample level.

Figure A.6. Urban Center Effects: Upwind Shocks, Particulates Pollution, and Pollen Counts

(a). Particulate matter (monitors)



(b). Pollen counts (monitors)



Notes: This figure shows regression coefficients of air pollution (panel a) and pollen counts (panel b) on 20 lead, contemporaneous, and 20 lag terms of upwind vegetation shocks. Each line is from a separate robustness check.

Table A.1. Top Ten Medications Most Frequently Prescribed for Respiratory and Sensory ER Visits

	(1)	(2)
Rank	Respiratory	Sensory
1	ambroxol hydrochloride	levofloxacin eye drops
2	promethazine hydrochloride injection	ofloxacin eye ointment
3	ambroxol hydrochloride injection	rb-bFGF
4	acetaminophen tablets (Tylenol)	tobramycin eye drops
5	doxofylline injection	levofloxacin hydrochloride eye gel
6	compound liquorice tablets	pranoprofen eye drops
7	pseudoephedrine hydrochloride tablets (Sudafed)	erythromycin eye ointment
8	diprophylline injection	emedastine difumarate eye drops
9	ambroxol hydrochloride tablets	levofloxacin hydrochloride eye drops
10	budesonide inhaler (Pulmicort)	sodium hyaluronate eye drops

Notes: This table shows the top ten medications that are most commonly prescribed for respiratory emergencies (column 1) and sensory emergencies (column 2) in the emergency room.

Table A.2. Robustness: Wind Shocks, Particulates Pollution, and Pollen Counts

	(1)	(2)	(3)	(4)	(5)	(6)
Source of vegetation shock:	MMP grids (5km rad.)			All grids (5km rad.)		
I. Outcome = Log pollen counts						
Log upwind NDVI	0.640**	0.735**	0.613***	0.688**	0.745**	0.473**
	(0.232)	(0.288)	(0.160)	(0.253)	(0.323)	(0.177)
Log downwind NDVI	0.097	0.082	-0.014	-0.043	-0.067	-0.339
	(0.274)	(0.250)	(0.247)	(0.258)	(0.298)	(0.249)
Observations	6,912	6,912	6,912	8,394	8,394	8,394
II. Outcome = Log $PM_{2.5}$						
Log upwind NDVI	-0.050	-0.005	0.113	-0.491	-0.555	-0.197
	(0.255)	(0.262)	(0.235)	(0.338)	(0.339)	(0.304)
Log downwind NDVI	-0.098	-0.042	0.098	0.064	0.063	0.405
	(0.255)	(0.269)	(0.236)	(0.310)	(0.327)	(0.289)
Observations	47,974	47,974	47,974	74,022	74,022	74,022
FEs: monitor	✓			✓		
FEs: month	✓			✓		
FEs: year	✓	✓		✓	✓	
FEs: monitor×month		✓	✓		✓	✓
FEs: year×month			✓			✓
FEs: day-of-week	✓	✓	✓	✓	✓	✓

Notes: Each panel-column is a separate regression. Each panel looks at a different outcome variable. Each column uses a different set of fixed effects controls. In columns 1-3, upwind and downwind NDVIs are defined using MMP grids within 5km radius of the monitor. In columns 4-6, upwind and downwind NDVIs are defined using all grids within 5km radius of the monitor. Standard errors are two-way clustered at both the monitor and the day-of-sample levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.3. Alternative Specifications: Wind Shocks, Particulates Pollution, and Pollen Counts

	(1)	(2)	(3)	(4)	(5)	(6)
Source of vegetation shock:	MMP grids (10km rad.)			All grids (10km rad.)		
I. Outcome = Log pollen counts						
Log upwind NDVI	0.727***	0.730**	0.601**	0.612**	0.660*	0.374*
	(0.234)	(0.279)	(0.210)	(0.261)	(0.333)	(0.194)
Log downwind NDVI	-0.079	-0.063	-0.164	-0.011	0.007	-0.292
	(0.247)	(0.272)	(0.247)	(0.252)	(0.312)	(0.186)
Observations	8,385	8,385	8,385	8,394	8,394	8,394
II. Outcome = Log $PM_{2.5}$						
Log upwind NDVI	-0.416*	-0.425*	-0.277	-0.561*	-0.649**	-0.299
	(0.230)	(0.234)	(0.221)	(0.294)	(0.288)	(0.253)
Log downwind NDVI	0.030	0.022	0.127	0.020	-0.033	0.298
	(0.222)	(0.238)	(0.215)	(0.262)	(0.294)	(0.271)
Observations	64,854	64,854	64,854	74,022	74,022	74,022
FEs: monitor	✓			✓		
FEs: month	✓			✓		
FEs: year	✓	✓		✓	✓	
FEs: monitor×month		✓	✓		✓	✓
FEs: year×month			✓			✓
FEs: day-of-week	✓	✓	✓	✓	✓	✓

Notes: Each panel-column is a separate regression. Each panel looks at a different outcome variable. Each column uses a different set of fixed effects controls. In columns 1-3, upwind and downwind NDVIs are defined using MMP grids within 10km radius of the monitor. In columns 4-6, upwind and downwind NDVIs are defined using all grids within 10km radius of the monitor. Standard errors are two-way clustered at both the monitor and the day-of-sample levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.4. Alternative Specifications: Wind Shocks, Particulates Pollution, and Pollen Counts

	(1)	(2)	(3)	(4)	(5)	(6)
Source of vegetation shock:	MMP grids only (15km rad.)			All grids (15km rad.)		
I. Outcome = Log pollen counts						
Log upwind NDVI	0.788**	0.863**	0.799**	0.580**	0.599*	0.239
	(0.303)	(0.343)	(0.289)	(0.261)	(0.323)	(0.208)
Log downwind NDVI	-0.215	-0.160	-0.244	0.001	0.020	-0.368*
	(0.274)	(0.311)	(0.238)	(0.262)	(0.315)	(0.183)
Observations	8,394	8,394	8,394	8,394	8,394	8,394
II. Outcome = Log $PM_{2.5}$						
Log upwind NDVI	-0.835***	-0.881***	-0.671**	-0.783***	-0.870***	-0.512**
	(0.292)	(0.291)	(0.261)	(0.274)	(0.260)	(0.235)
Log downwind NDVI	0.384	0.374	0.463*	0.202	0.130	0.435
	(0.275)	(0.296)	(0.263)	(0.237)	(0.289)	(0.287)
Observations	74,022	74,022	74,022	74,022	74,022	74,022
FEs: monitor	✓			✓		
FEs: month	✓			✓		
FEs: year	✓	✓		✓	✓	
FEs: monitor×month		✓	✓		✓	✓
FEs: year×month			✓			✓
FEs: day-of-week	✓	✓	✓	✓	✓	✓

Notes: Each panel-column is a separate regression. Each panel looks at a different outcome variable. Each column uses a different set of fixed effects controls. In columns 1-3, upwind and downwind NDVIs are defined using MMP grids within 15km radius of the monitor. In columns 4-6, upwind and downwind NDVIs are defined using all grids within 15km radius of the monitor. Standard errors are two-way clustered at both the monitor and the day-of-sample levels. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.5. Robustness: Pollen Exposure and Emergency Room (ER) Utilization, Same-Day Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All causes	Respiratory	Sensory	Others	All causes	Respiratory	Sensory	Others
	ER visits				ER spending			
A. Baseline specification								
All ER	0.0079***	0.0197***	0.0411***	0.0055***	0.0048***	0.0141***	0.0727***	0.0028
	(0.0017)	(0.0043)	(0.0066)	(0.0016)	(0.0015)	(0.0040)	(0.0202)	(0.0017)
B. Effects by severity								
ER → Not hospitalized	0.0076***	0.0197***	0.0412***	0.0051***	0.0048***	0.0146***	0.0731***	0.0027
	(0.0017)	(0.0043)	(0.0066)	(0.0016)	(0.0015)	(0.0042)	(0.0199)	(0.0017)
ER → Hospitalized	-0.0044	-0.0012	-0.0015	-0.0036	-0.0164	-0.0078	-0.0211	-0.0020
	(0.0044)	(0.0033)	(0.0017)	(0.0047)	(0.0279)	(0.0402)	(0.0244)	(0.0315)
C. Effects by age								
Age $<$ 60	0.0081***	0.0218***	0.0423***	0.0053***	0.0050**	0.0198***	0.0851***	0.0028
	(0.0019)	(0.0047)	(0.0066)	(0.0017)	(0.0021)	(0.0050)	(0.0194)	(0.0021)
Age $\geq$ 60	0.0069***	0.0137***	0.0121**	0.0054***	0.0055*	0.0211*	0.0466*	0.0042
	(0.0014)	(0.0041)	(0.0043)	(0.0016)	(0.0025)	(0.0100)	(0.0216)	(0.0032)
D. Effects by prior condition								
Prior respiratory-sensory visits	0.0330***	0.0313***	0.0193***	0.0278***	0.0322***	0.0883***	0.1122***	0.0429***
	(0.0038)	(0.0051)	(0.0047)	(0.0047)	(0.0084)	(0.0196)	(0.0274)	(0.0112)
No prior visits	0.0073***	0.0169***	0.0328***	0.0052***	0.0047***	0.0133***	0.0638***	0.0030
	(0.0015)	(0.0037)	(0.0060)	(0.0015)	(0.0013)	(0.0033)	(0.0183)	(0.0017)

Notes: Each cell represents a separate regression using district-day level data. Each column presents ER records corresponding to different diagnoses. Panel (a) uses all ER records. Panel (b) stratifies by visits that did and did not end up with hospital admissions. Panel (c) stratifies by age of the patient. Panel (d) stratifies by whether the patient had respiratory and sensory visits in the previous 30 days. All regressions control for district-by-month fixed effects, year-by-month fixed effects, day-of-week fixed effects, and holiday fixed effects. Standard errors are two-way clustered at the district and day-of-sample levels. Number of observation for each regression is 8,394. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Table A.6. Robustness: Pollen Exposure and Emergency Room (ER) Utilization, Seven-Day Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All causes	Respiratory	Sensory	Others	All causes	Respiratory	Sensory	Others
	ER visits				ER spending			
A. Baseline specification								
All ER	0.0042**	0.0149***	0.0353***	0.0023	0.0030**	0.0134***	0.0288***	0.0006
	(0.0014)	(0.0034)	(0.0052)	(0.0014)	(0.0011)	(0.0023)	(0.0072)	(0.0013)
B. Effects by severity								
ER → Not hospitalized	0.0041**	0.0149***	0.0354***	0.0021	0.0030**	0.0138***	0.0296***	0.0005
	(0.0014)	(0.0034)	(0.0053)	(0.0014)	(0.0011)	(0.0025)	(0.0073)	(0.0014)
ER → Hospitalized	0.0036	0.0042	-0.0054	0.0032	0.0021	-0.0044	-0.0939	0.0074
	(0.0084)	(0.0087)	(0.0047)	(0.0071)	(0.0237)	(0.0550)	(0.0566)	(0.0249)
C. Effects by age								
Age < 60	0.0049***	0.0173***	0.0391***	0.0027*	0.0036*	0.0173***	0.0295**	0.0014
	(0.0015)	(0.0041)	(0.0054)	(0.0014)	(0.0017)	(0.0038)	(0.0095)	(0.0016)
Age ≥ 60	0.0027	0.0095*	0.0161	0.0013	0.0029	0.0115**	0.0423	-0.0000
	(0.0018)	(0.0044)	(0.0092)	(0.0017)	(0.0022)	(0.0044)	(0.0351)	(0.0022)
D. Effects by prior condition								
Prior respiratory-sensory visits	0.0301***	0.0453***	0.0480***	0.0254***	0.0218***	0.0430**	0.1204***	0.0229***
	(0.0049)	(0.0091)	(0.0093)	(0.0051)	(0.0059)	(0.0172)	(0.0332)	(0.0067)
No prior visits	0.0039**	0.0130***	0.0297***	0.0022	0.0030**	0.0124***	0.0197*	0.0007
	(0.0014)	(0.0034)	(0.0068)	(0.0014)	(0.0012)	(0.0024)	(0.0090)	(0.0015)

Notes: Each cell represents a separate regression using district-day level data. Each column presents ER records corresponding to different diagnoses. All outcomes are measured using a seven-day look-ahead window (e.g., total number of ER visits today and the next six days). Panel (a) uses all ER records. Panel (b) stratifies by visits that did and did not end up with hospital admissions. Panel (c) stratifies by age of the patient. Panel (d) stratifies by whether the patient had respiratory and sensory visits in the previous 30 days. All regressions control for district-by-month fixed effects, year-by-month fixed effects, day-of-week fixed effects, and holiday fixed effects. Standard errors are two-way clustered at the district and day-of-sample levels. Number of observation for each regression is 8,394. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.