Measuring the Long-Term Health Impact of Environmental Externalities

Steffen Andersen* Lorenzo Bruno† Luigi Guiso‡

Federico Marciano§ Paola Sapienza§ Luigi Zingales†

March 2025

Preliminary and incomplete

Abstract

We exploit a forty-year longitudinal Danish dataset with detailed health outcomes to estimate the long-term health effects of human exposure to pollution. By exploiting openings and closings of waste incinerators, our study makes a novel contribution by disentangling two key mechanisms: cumulative exposure (how long one is exposed) and latency (the delay between exposure cessation and health improvement). Using a difference-in-differences approach, we find that people living within 4 km of an incinerator experience a 24% increase in lung cancer risk that only becomes apparent two decades after initial exposure. Similarly, individuals whose exposure ends due to incinerator closures show no reduction in cancer rates until at least 20 years after closure, revealing the substantial latency period of pollution's health effects. These delayed effects are economically significant: the increased mortality risk is twice the short-term effect found by Deryugina et al. (2019), 6 times greater than the risk of dying in a car accident and 14 times higher than work-related death risk. Using the value of statistical life, we estimate that the health costs of these emissions are nearly three times the incinerators' operating costs, suggesting that conventional cost-benefit analyses substantially underestimate the true impact of pollution regulation.

Keywords: Pollution, Lung Cancer, Long Term Externalities.
JEL codes: I18, Q52, Q53

*Copenhagen Business School.
†University of Chicago Booth School of Business.
‡Einaudi Institute for Economics and Finance.
§Stanford University.

Introduction

It is well-known that the impact of smoking on health is delayed, often by 15-20 years (Lubin et al. (2007)). Yet, when it comes to the impact of pollution on health, most studies focus on short-term effects (e.g., Chay and Greenstone (2003b); Deryugina et al. (2019)) due to the unavailability of long-term data on exposure and health outcomes at the individual level. As a result, we tend to underestimate the cost of pollution by an unknown amount.

This paper exploits access to a forty-year longitudinal Danish dataset with detailed health outcomes to estimate the long-term effects of human exposure to pollution. By exploiting both openings and closings of waste incinerators in Denmark, we are able to disentangle two key aspects of pollution's health impacts that previous research has been unable to separate: the role of cumulative exposure (how long one is exposed) and latency (the delay between exposure and disease manifestation). This innovation is particularly important for diseases such as lung cancer, for which biological mechanisms suggest a complex relationship between exposure and disease manifestation, with distinct stages of initiation, promotion, and progression that unfold over many years.

Our analysis offers several advantages for studying these long-term effects. First, we can track individuals over four decades, providing the longitudinal depth needed to measure cumulative exposure and observe delayed health effects. Second, we can link this data to complete health histories, including lung cancer diagnoses and mortality. Third, while we control for background pollution, we focus on one specific source of pollution (waste incinerators) for which we have information on the actual location, dates of operation, type of pollutants, and emissions. Moreover, crucially for our identification strategy, our sample period contains numerous incinerator openings and closings, which create sudden and persistent changes in pollution exposure for large segments of the population. These events provide the exogenous variation needed to establish causal relationships between long-term exposure and health outcomes.

We begin by establishing a robust correlation between proximity to incinerators and lung cancer risk. Living within 2 km of an incinerator is associated with a 7.3 basis point higher risk of lung cancer diagnosis compared to individuals living farther away - an increase of the same order of magnitude as the average population risk. This correlation diminishes both

economically and statistically beyond 4 km.

To move beyond correlations, we introduce our main innovation to the literature and measure cumulative exposure for everyone in our sample — the number of years each individual has lived within 4 km of an incinerator. We study how this variable affects the probability of developing lung cancer and dying from it. To address potential endogeneity concerns due to residential sorting, we instrument the exposure time using incinerator openings and closures.

Our most novel contribution comes from exploiting openings and closings in a difference-in-differences setup where our treated group comprises individuals living in locations affected by an opening (or closing) of an incinerator, and the control group consists of individuals living in locations unaffected by the treatment. The results show openings cause a 3.28 basis point increase in cancer diagnoses (which corresponds to a 24% increase with respect to the baseline probability) with a two-decade delay, suggesting the importance of taking into account the length of exposure when studying pollution effects. Most importantly, our analysis of closings enables us to separate cumulative exposure from latency effects — a distinction critical for both scientific understanding and policy evaluation. By comparing individuals whose exposure ended due to a closure with a control group that remained exposed, we find no difference in cancer rates for the first 20 years after closure. Only after two decades did treated individuals experience a statistically and economically significant drop in lung cancer rates. This reveals that pollution's latency effect can extend as long as 20 years after exposure ceases.

This finding has profound implications for regulatory policy and cost-benefit analysis. Conventional approaches that focus on immediate or short-term outcomes after pollution reduction interventions may dramatically underestimate the true benefits of such policies, as health improvements might not become apparent until decades later. Without accounting for these delayed effects, policymakers might incorrectly conclude that pollution reduction measures are less effective than they truly are.

The effects we estimate are substantial. Lung cancer mortality due to long-term exposure is economically relevant. The cumulative probability of dying from lung cancer as a consequence of exposure is twice the short-term effect found by Deryugina et al. (2019), 6 times higher than that of dying in a car accident and 14 times higher than that of a work-related death.

Economically, the ”value of statistical life” of the deaths is almost three times the total operating costs of the incinerators.

We are neither the only nor the first study of the long-term effects of pollution. Barreca et al. (2021) study the effect of pollution reduction caused by the Acid Rain Program over a 10-year horizon. They rely, however, on county-level data and do not have data on individual cumulative exposure. Anderson (2020) studies the effect of highway pollution on mortality on a 3-year horizon, a period too short for the latency effects to manifest. **?** study the long-term economic impact of an exogenous increase in the use of coal in East Germany, but they do not have health outcomes and use economic outcomes as proxies for health. Finally, Deryugina and Julian (2023) developed a dynamic health production model to estimate long-term consequences from short-term exposure effects, extending the short-term results obtained in Deryugina et al. (2019). However, these estimates do not take into account the cumulative exposure of individuals over time.

1 Available Evidence

Understanding how air pollution affects human health requires careful consideration of both empirical evidence and biological mechanisms. This section first reviews the evolution of empirical approaches to studying pollution's health effects, highlighting methodological advances and remaining challenges in establishing causal relationships. We then examine the biological pathways through which pollution exposure leads to cancer development, focusing on the temporal dynamics that make the study of pollution's carcinogenic effects particularly complex.

1.1 Health Effects of Air Pollution

The literature examining the health effects of air pollution has evolved substantially over time, progressing from observational studies to quasi-experimental designs. However, significant methodological challenges persist, particularly in establishing causal relationships between long-term exposure and chronic health conditions such as cancer. Early research relied heavily

on observational studies to establish correlations between air pollution and adverse health outcomes. The seminal Six Cities Study (Dockery et al., 1993) documented a significant association between fine particulate matter (PM2.5) and mortality by tracking individuals across six U.S. cities with varying pollution levels. This approach was subsequently expanded to encompass more than 500,000 adults across 51 metropolitan areas, revealing correlations between particulate air pollution and various mortality causes, including cancer, even after controlling for individual risk factors such as smoking and education. While these studies made important contributions by identifying potential health risks, their observational nature limited causal inference. A fundamental challenge in these observational studies stems from self-selection bias. Individuals who are more susceptible to pollution-related health problems might actively avoid highly polluted areas or adopt protective measures, such as using air purifiers or spending less time outdoors. These behavioral responses, which are typically unobserved in available datasets, introduce measurement errors and potential biases that complicate causal interpretation. To address these methodological limitations, economists have increasingly turned to quasi-experimental approaches that leverage exogenous variations in pollution levels. For instance, Chay and Greenstone (2003b) exploited the natural experiment created by the 1981-82 recession, which caused sharp geographic variations in pollution reduction. Similarly, the Clean Air Act of 1970 provided another opportunity to study pollution's health effects by inducing substantial county-level variation in pollution changes (Chay et al., 2003; Chay and Greenstone, 2003a). More recent work by Deryugina et al. (2019) used changes in wind direction as an instrumental variable to study the acute effects of PM2.5 exposure on elderly mortality and healthcare costs. These quasi-experimental studies have made groundbreaking contributions to our understanding of pollution's health effects, particularly in contexts where exposure creates permanent and irreversible impacts. The methodology has been especially valuable in advancing our knowledge of how air pollution affects children, where even short-term exposure can have lasting consequences[1]. However, when applied to studying health impacts more broadly, the predominant focus on short-term outcomes, typically measured over brief windows ranging from days to months, creates a crucial gap in our understanding of pollution's full health effects.

[1] Currie and Neidell (2005); Currie and Schmieder (2009); Currie and Walker (2011)

Many serious health conditions, particularly cancer and cardiovascular diseases, develop through cumulative exposure over extended periods, sometimes decades. The biological mechanisms underlying carcinogenesis, for instance, suggest that short-term effects may not accurately predict long-term outcomes. Some researchers have attempted to bridge this gap by extending short-term analyses to longer horizons. Isen et al. (2017) tracked infants affected by pollution using discontinuities generated by the Clean Air Act to study long-term consequences on labor market outcomes. Barreca et al. (2021) study the long-term health beneficial effects of the Acid Rain program, relying on county-level information. Deryugina and Julian (2023) developed a dynamic health production model to estimate long-term consequences from short-term exposure effects. However, these approaches still face significant data limitations. Without information on cumulative exposure over decades and individual behavioral responses such as residential mobility, these studies must rely on strong assumptions about the persistence of exposure-response relationships. This methodological constraint is particularly problematic for diseases with long latency periods, such as lung cancer, where the relationship between exposure and disease development cannot be adequately captured by short-term data. The biological mechanism underlying cancer development makes continuous exposure especially critical: when exposure to pollutants is interrupted, the human body has an opportunity to eliminate existing cancer cells whose replication is no longer being facilitated by pollutants. Therefore, studies that cannot account for gaps or interruptions in exposure may miss crucial variations in cancer risk. Consequently, existing research may substantially underestimate the true health costs of pollution by failing to account for these long-term effects. Our study addresses this critical gap by utilizing a longitudinal panel that allows us to measure cumulative exposure for each individual while accounting for changes in location and other individual characteristics. This unique dataset enables us to make more precise causal inferences about the long-term health consequences of pollution exposure, particularly for conditions that develop over extended periods.

1.2 Cancer Development and Exposure Dynamics

Given our interest in estimating the causal effect of pollution on cancer, our specifications need to be guided by knowledge of the medical literature. The medical consensus (e.g., Pitot, 1993) has it that cancer development consists of three main stages. The first, cancer initiation, is triggered by a DNA mutation that transforms a normal cell into a cancer cell. The second, cancer promotion, consists of the replication of the initial cancer cell. The third, cancer progression, is the continuation of cancer promotion and involves the transformation of a benign tumor into a neoplasm and into malignancy[2]. The completion of these three stages can take a long time, depending on the cancer site. In addition, the human body has its own mechanisms to repair DNA damage and to react to the replication of the first cancer cell, slowing down the process. For these reasons, decade-long latency periods can separate the first mutation in the DNA of the initial cancer cell from the detection of the disease by means of a diagnosis. For example, Nadler and Zurbenko (2014) document that 35 out of 44 cancer sites are associated with more than 10 years of latency. In particular, lung cancer has a latency of 13.6 years.

Latency has important implications for our empirical analysis. In our framework, specifications that rely on a contemporaneous relationship between treatment and outcome are bound to be imprecise and unreliable. Instead, we need to resort to specifications that allow for decade-long delays between the time of first exposure to the treatment and the manifestation of the outcome.

The medical literature sheds light on another relevant mechanism. Carcinogenic substances can be divided into three categories: cancer initiators, cancer promoters, and strong carcinogens. The first group of substances contributes to cancer initiation by determining the initial DNA mutation that gives rise to cancer development. The second group of substances contributes to cancer promotion by facilitating and accelerating the replication of pre-existing cancer cells. The third group of substances is able to contribute to both stages. To give some examples, tobacco smoke is a strong carcinogen, but its cancer-promotion properties are thought to prevail over the cancer-initiation effects (Weisburger, 1990). Whether generic air pollution is a cancer promoter or a cancer initiator is still an open question (Santibáñez-Andrade et al., 2019). The

[2] https://www.cancerquest.org/cancer-biology/cancer-development

pollution mix emitted by incinerators contains several toxic substances (heavy metals, dioxins, SO2, PM2.5), which all contribute to generic air pollution.

This classification of carcinogenic substances has important consequences on how we should model the effect of cumulative exposure on cancer risk. If the toxic substances emitted from incinerators are only cancer initiators, cumulative exposure should play a minor role to the extent to which cancer initiation was triggered at the time of first exposure; in particular, if the cancer development process has already been initiated, the end of exposure cannot reduce the risk of a diagnosis. In contrast, if the pollution mix emitted from incinerators acts as a cancer promoter, cumulative exposure is essential to the acceleration of a latent cancer development process. In addition, if the exposure to the cancer promoter ends (e.g., an incinerator is shut down), the excess risk of cancer should vanish some years after the end of exposure, when the human body has time to destroy the existing cancer cells, whose replication is not facilitated by the action of an external agent anymore. Since we do not know in advance whether incinerators contribute to cancer initiation or cancer promotion, in our specification it is important to control for cumulative exposure and to observe treated individuals also after they stop being exposed to incinerators.

2 Data

Our data consists of several longitudinal datasets: individual-level data on residency, health, and economic status; data on air quality from incinerators and air pollution models; and sales data from real estate trades.

2.1 Individual-Level Data

Our core source for individual-level information is a unique dataset covering the Danish population from 1980 to 2018, the Danish Civil Registration System (CPR Registry). By law, each individual in Denmark is given a unique identification number at birth in The Danish Civil Registration System, the CPR number. The CPR Registry allows us to observe demographic variables such as geocoded residential address, age, sex, marital status, and the identity (CPR) of

spouses, children, and parents. The CPR allows us to track individuals, family, and generation links over time. The 1980 addresses are also associated with the date the individual of interest moved to the relevant address, back to 1900. From the CPR, we obtain 208,338,560 person-year observations of individuals between 0 years and 122 years.

Given that we are interested in proxying for overall socioeconomic status as they are strong predictors of health outcomes, we obtain financial data from the Danish Tax Authority (mainly income and wealth information) expressed in 2015-DKK. In all our specifications, we lag income and wealth by one year, given that there is a mechanical relationship between the manifestation of events such as cancer diagnosis and a decline in economic status Chetty et al. (2016). As a result, we lose one year of data for every individual, and the sample for which we have the complete set of variables consists of the Danish population aged at least one year of age when observed between 1981 and 2018, lowering our sample to (199,715,678 person-years). We additionally drop immigrants and individuals with observed gaps (i.e. leaving and returning expats), removing around 20 million person-years for our final population sample of (175,188,237 person-years).

We collate information on causes of death and disease treatment and observe the health status of individuals in the population from several different health registers. We obtain the deceased's cause of death from the Cause of Death Registry, where a medical doctor attests to the cause of death following WHO IDC10 codes. We obtain any treatment patients receive at any hospital in Denmark from the National Patient Registry and the Cancer Registry, also following IDC10 codes. For observed health status, in particular of interest smoking behavior, we retrieve records from all Danish pregnant women between 1991 and 2018 (1,712,734 pregnant-woman years). From this data source, we observe a dummy variable (`Smoker`) that takes the value of 1 if a pregnant woman is classified as a current smoker.

To enable us to assign each individual a pollution exposure measure, we use a grid of geographical units to match individual-level data with environmental data on incinerator exposure from an air dispersion model. Given that our core CPR data source allows us to easily match addresses to parishes, constituting both a small administrative as well as a small and compact geographical unit in Denmark, we chose to use Danish parishes as the relevant grid unit. In

2018, there were 2,165 parishes, with an average of 7.839 people per parish, covering on average (median) 16.57 (11.66) square kilometers.

Since our analysis focuses on the cumulative amount of exposure to incinerators, we are forced to drop all individuals without information on residential address for at least one year. This reduces the sample size to 172,107,549 person-exposure years.

2.2 Data on Air Quality

We collected data on the 75 incinerators in Denmark covering the period for active incinerators we could identify between 1980 and 2018. The collected information contains the location, opening and closing (if any) year, and height of the chimney; our sources are Kleis and Dalager (2007) and the local historical archives of the municipalities associated with the incinerators. We are unable to identify the opening and closing years for five incinerators, located in Helsingor, Svejstrup, Skibby, Struer, and Ringsted. Consequently, we drop these five incinerators. We are also unable to verify the usage of three small incinerators. Removing these three from the sample leaves us with 67 unique incinerators, with 1,402 incinerator-years observed between 1980 and 2018.

For each incinerator year, we collect information on annual emissions of the toxic pollutants relevant to our analysis to predict lung cancer (SO_2, $PM_{2.5}$, dioxins, and heavy metals). Our data sources for annual emissions are the Green Accounts, the Danish Energy Agency, and the CORINAIR Program of the European Environment Agency.

To supplement and complement the incompleteness of the "actual incinerator emissions" as reported in the Green Accounts, we compute "imputed emissions" as the product of the CORINAIR emissions factors and the annual amount of waste burnt. In addition, we download annual average concentrations of SO_2 registered by the network of measuring stations managed by the University of Aarhus as part of the Air Quality Monitoring Program of the Danish EPA.

2.3 Real Estate Transactions

We obtain sale prices of housing from real estate transactions sold between 1992 and 2013. As each property is also geolocated, as well as the data containing essential characteristics of

the land and buildings, we can examine the interaction between air quality and local house price developments. We can uniquely identify 1,525,584 house transactions. Each sold property is, on average, sold 4.96 times in the sample period.

3 Methodology and Identification

Our investigation begins with a descriptive analysis examining how health outcomes vary with distance from pollution sources. Specifically, we estimate hazard ratios using a Cox Proportional Hazards Model that measures the relationship between health outcomes and proximity to incinerators. We divide the area around each incinerator into five concentric rings: 0-2km, 2-4km, 4-6km, 6-8km, 8-10km and 10+km from the source. The model incorporates controls for both individual characteristics and municipality-level factors.

Our next steps employ multiple alternative methodological techniques to address the challenges of establishing a causal relationship between pollution and health outcomes. First, Our study examines long-term health effects by focusing on cumulative exposure to pollution over time, rather than just measuring proximity at a single point. We innovate by developing a measure of cumulative exposure, defined as the total number of years an individual resided within 4 kilometers of an incinerator.[3] This approach allows us to capture the duration of potential exposure, providing a more complete picture of how extended proximity to pollution sources may affect health outcomes. Using this variable, we first estimate both a Cox regression and an OLS model. However, these initial approaches face significant endogeneity concerns, as residential choices may be correlated with unobserved characteristics that also influence health outcomes.

The location of incinerators and residential choices around them are not random, creating potential selection bias in our analysis. In the United States, this selection occurs through two key mechanisms. First, companies strategically place polluting facilities in lower-income neighborhoods to minimize political resistance from wealthy, politically connected residents (Banzhaf et al. (2019)). Second, the presence of these facilities tends to depress local housing

[3]We choose 4Km as our descriptive analysis suggests that further away the effect is not significant, but in robustness analysis we experiment with different thresholds.

prices, attracting lower-income residents who face budget constraints in their housing choices (Currie et al. (2015)). This dual selection process means that populations near polluting facilities often differ systematically from control groups in ways that affect health outcomes independently of pollution exposure, such as pre-existing health conditions and socioeconomic factors (Graff Zivin and Neidell (2013)). The Danish context offers important advantages for addressing these selection concerns. Unlike in the United States, Danish incinerators are distributed throughout the country, with their locations primarily determined by proximity to economic activity centers to minimize waste transportation costs. This pattern is exemplified by Copenhagen, where two major facilities operate, including the Amager Bakke plant, which has been integrated into urban life as a recreational facility in the city center. Furthermore, our analysis benefits from high-quality administrative data that allows us to control for precisely measured individual-level characteristics including income, wealth, and marital status, all sourced from government records.

Despite these advantages of the Danish setting and our rich controls, we must still address potential unobserved differences between exposed and unexposed individuals that could bias our results. To tackle this challenge, we employ an instrumental variables strategy. For individuals who experience the opening of a new incinerator during our study period we instrument their cumulative exposure with a binary variable indicating whether an incinerator opened in their initial residential location. This approach leverages the fact that while current residential location may be endogenous, historical location choices were made before individuals could have anticipated the future placement of incinerators.

While our instrumental variables approach addresses some endogeneity concerns, it cannot fully rule out the possibility that unobserved characteristics correlated with incinerator placement might influence health outcomes. For instance, areas selected for new incinerators might systematically differ in ways that affect cancer risk independently of pollution exposure. To strengthen our identification strategy, we employ two difference-in-differences analyses examining both incinerator openings and closings. This approach allows us to control for time-invariant differences between treated and untreated populations by focusing on how health outcomes evolve differently over time in response to changes in incinerator operations.

For our analysis of openings, we define treatment as living within 4 kilometers of a newly opened incinerator. Our control group consists of individuals who have never been exposed to an incinerator throughout the study period. In our analysis of closings, we construct the treatment group similarly, including individuals who either lived within 4 kilometers of an incinerator in its closing year or who died in a location where an incinerator would eventually close. The control group for the closing analysis comprises individuals exposed to incinerators that remained operational, with exposure measured relative to a matched fictitious closing year. To further refine our identification strategy, we develop a more nuanced measure of pollution exposure that accounts for multiple factors affecting pollution dispersion. Simple distance-based measures may miss important variation in exposure levels that depend on facility characteristics and environmental conditions. We address this limitation by employing an air dispersion model that incorporates multiple factors: emission source locations, pollution quantities, chimney heights, and local weather patterns. This model generates estimates of pollutant concentrations across different geographical areas, providing a more precise measure of individual exposure. We provide a detailed discussion of this model's specification, implementation, and validation in Section 4.

4 An Alternative Measure of Exposure to Incinerators

4.1 Air Dispersion Modeling

Our primary objective is to assess the impact on lung cancer of pollution from incinerators. Thus, we need to able to distinguish between individuals that are exposed to incinerators and individuals that are not. The most natural measure of exposure is proximity: one can rely on a threshold in terms of distance between the residential address of an individual and the closest pollution source (in our case, the closest incinerator) to tell part the exposed from the unexposed. This is indeed the approach that we follow in our main specifications, in line with the environmental economics literature (e.g., Banzhaf and Walsh, 2008 and Currie et al., 2015). However, measuring exposure in terms of distance from the closest incinerator involves overlooking some important determinants of the intensity of exposure. First, heterogeneity

in size and emissions from different incinerators. Second, the height of the chimney (which is important to assess how far pollution can go). Third, wind direction, which can make the level of exposure of upwind and downwind locations very different, controlling for distance from a point source of interest. Fourth, the impact from multiple point sources that affect the same location jointly. As pointed out by Hernandez-Cortes and Meng (2020), this can result in exposure misclassification, which is a form of measurement error.

To make sure that our results are not driven by the way we measure exposure, in addition to distance we also consider an air dispersion model. In a nutshell, an air dispersion model takes as inputs the coordinates of multiple point sources, their emissions, their characteristics (mainly the height of the chimney), and weather conditions (mainly wind direction); the output is the spatial distribution of the concentration of the pollutants of interest. Since the 2000s, air dispersion models have become the standard approach to measure air pollution consequences in the environmental and medical science literature (Næss et al., 2007, Pennington et al., 2018, and Filigrana et al., 2022). So far, only a few economists have taken advantage of this tool to measure exposure (important references include Sullivan, 2017, Grainger and Ruangmas, 2018, Hernandez-Cortes and Meng, 2020, and Christensen et al., 2022).

4.2 HYSPLIT

The two models of air dispersion most commonly used in the literature are AERMOD[4] and HYSPLIT[5], developed respectively by the US EPA and by the National Oceanographic and Atmospheric Administration. Between AERMOD and HYSPLIT, we decide to use HYSPLIT to model the air dispersion of annual emissions from incinerators. AERMOD provides a more rigorous modelling of the dispersion of pollutants in the atmosphere, taking into account chemical reactions between different pollutants and the relationship between each substance and its precursors; nevertheless, this model is not computationally agile and is typically used by regulatory agencies to perform short-range analyses (at most 50Km) based on one or a few pollution sources. In contrast, HYSPLIT, thanks to its ability to balance computational tractability and scientific realism, is more suitable for our large-scale analysis that involves

[4] https://www.epa.gov/scram/aermod-modeling-system-development
[5] https://www.ready.noaa.gov/HYSPLIT.php

several point sources every year (Hernandez-Cortes and Meng, 2020).

We run HYSPLIT by adapting the procedure proposed in Henneman et al. (2019) to our framework[6]. This procedure involves simulating the release of 100 fictitious particles of a pollutant of interest (in our case, SO2) from each incinerator-year between 1980 and 2018, 4 times a day (12am, 6am, 12pm, and 6pm). Each release is tracked for a number of hours that depends on the pollutant under investigation (according to Hernandez-Cortes and Meng, 2020, 13 hours are needed for SO2), taking into consideration the height of the chimney of the relevant incinerator as well as hourly weather conditions provided by the National Oceanographic and Atmospheric Administration reanalysis database. The results obtained from daily simulations yield the annual concentration of the particles emitted from each incinerator with reference to some geographical units, parishes in our case. Then, for each parish, the weighted sum of the concentrations associated with different incinerators is computed, considering annual emissions of the pollutant of interest as weights. The resulting unitless measure is called HyADS (short for HYSPLIT Average Dispersion) and ranks parish-years based on exposure to incinerators. Depending on the parish where a person i lives in year t, we can quantify i's exposure to pollution from incinerators. While HyADS is not a cardinal measure, it certainly provides an objective and quantitative way to rank the intensity of exposure for individuals living in different locations and also at different points in time; in addition, contrary to distance, this measure considers heterogeneity across incinerators, wind direction, and the joint effect of multiple point sources on the same location.

4.3 SO2 as a Proxy for Pollution from Incinerators

A key choice in our computation of HyADS is to select the pollutants of interest (or the best proxies for such pollutants) and to measure annual emissions of such pollutants from incinerators, which represent the main input used to run the dispersion model. We decide to use SO2 as a proxy for the entire toxic pollution mix (dioxins, NOx, PM2.5, and heavy metals) emitted by incinerators. We justify this choice by resorting to two reasons. First, SO2 is the

[6]To run HYSPLIT, we make use of the R package called DisperseR. In addition, we exploit the resources provided by the Mercury high-performance computing cluster at The University of Chicago Booth School of Business, which is supported by the Office of the Dean.

substance for which we have the most complete data on actual emissions from the available Green Accounts. Second, SO2 is strongly correlated with other air pollutants, performing very well as a general pollution tracker and being used since the 1970s as a criterion pollutant by regulatory agencies all over the world, including the US EPA.

4.4 Assigning SO2 Emissions to Incinerator-Years without Green Accounts

While the Green Accounts are the best source of data on annual SO2 emissions, they are available from 1996 (and at the moment we have access only to reports published in the periods 1996-2001 and 2007-2018), with some incinerator-years missing due to the confidentiality of their reports or to size thresholds established by law. In total, out of the 1,435 incinerator-years included in our dataset spanning 1980-2018, we observe SO2 emissions from the Green Accounts for only 516 of them. To supplement the data on SO2 for the incinerator-years without Green Accounts, we use one additional source: imputed emissions computed as the product between the annual sector-specific SO2 emissions factor and the annual amount of waste burnt using CORINAIR data.

For the incinerator-years associated with Green Accounts (516), we consider SO2 emissions as reported in these documents. For the other incinerator-years (919), we exploit CORINAIR SO2, computed as the product between waste burnt and CORINAIR emissions factors. Prior to 1994, we discount 1994 waste by real GDP growth. Prior to 1990, we assume the 1990 emissions factors to be constant. Figure 18 shows the pattern followed by total SO2 emissions from incinerators (in tonnes) according to our strategy, distinguishing between the sum of emissions from incinerator-years with Green Accounts (red area) and the sum of emissions from incinerator-years for which we use CORINAIR SO2 (blue area).

4.5 Computing HyADS Based on SO2 Emissions

Using emissions of SO2 from incinerators (either reported in the Green Accounts or computed with the CORINAIR emissions factors), we run HYSPLIT and compute a value of HyADS

for each parish-year between 1980 and 2018. Figure 19 shows the map of Denmark in 1980, in 1997, and in 2018. HyADS is characterized by a strong variation, both in the time series (the peak is 1.636 in 1980 and 0.266 in 2018) and in the cross-section. Over time, new sources of pollution arise; for example, in 1997, the Naestved plant turns out to be the most polluting incinerator, emitting 380 tonnes of SO2 according to its Green Account.

To show that our measure is a good proxy for exposure to pollution from incinerators, we regress the concentrations of SO2 registered by the monitors of the University of Aarhus on HyADS based on actual and fitted emissions of SO2 and computed with reference to the parishes associated with the monitors. According to Table 23, HyADS alone explains more than half of the variance of the dependent variable. The coefficient on HyADS remains positive and statistically significant after controlling for year and monitor fixed effects. Considering that HyADS is only based on pollution from one source (incinerators) and that monitors are not necessarily located in the vicinity of the plants of interest, these results suggest that HyADS is a good proxy for exposure to incinerators over time and across different areas.

5 Results

In this section, we present the estimation results using a variety of strategies. All estimates include a comprehensive set of controls to account for potential confounders.

First, we control for individual demographics, including gender, marital status, birth cohort, and age group dummies, to capture systematic differences in lung cancer risk across population groups and life-cycle stages. Second, we include economic controls — individual income and family wealth—both lagged by one year to mitigate mechanical correlations between health, income generation, and medical expense financing (via wealth decumulation).

Additionally, we incorporate municipality- and parish-level controls to account for regional heterogeneity in healthcare access and population characteristics. These include the number of hospitals in the municipality, the share of immigrants, the share of individuals with higher education, and the (log) population size.

As noted, these controls play a crucial role. Although our dataset is structured as a panel,

the non-repeated nature of our primary health outcome variable prevents us from relying on fixed effects to control for unobserved heterogeneity. However, given that our control variables are numerous and, importantly, free of measurement error, this approach serves as a reasonable alternative to individual fixed effects.

5.1 Correlations with distance from incinerators

We start discussing Cox regression estimates of the correlation between lung cancer diagnosis and distance from the incinerator run on the whole population data. Each estimate includes the controls just discussed and estimates are stratified by gender and birth cohort to allow for a baseline hazard unique to each stratum. Table 4 shows the results of the estimates. Distance is captured non-parametrically by inserting a set of dummies, with the closest distance dummy including all those leaving within a radius of two kilometers from the incinerator. The other equally-spaced distance dummies are 2-4 km, 4-6 km, 6-8 km and 8-10 km. The excluded group are individuals leaving more than 10 kilometers away from the incinerator. Health outcomes effects are thus measured relatively to this group.

We present results for lung cancer diagnosis (first column) and lung cancer mortality (second column), based on a full population sample of over 168 million observations. The large sample size is necessary to obtain reliable estimates of lung cancer risk, which in the overall population is 7.8 basis points. Our findings indicate that the relationship between lung cancer risk and distance from the emissions source weakens as distance increases. Living within 2 kilometers of an incinerator is associated with a 7.3 basis point higher risk of lung cancer diagnosis (estimated with precision) compared to individuals living more than 10 kilometers away. This increase is of the same order of magnitude as the average population risk. The risk declines to 4.3 basis points for those living between 2 and 4 kilometers, and beyond this range, the correlation diminishes in magnitude and loses statistical significance. Results for lung cancer mortality (second column) closely mirror those for diagnosis, suggesting that the choice of health outcome does not meaningfully affect our conclusions. Additionally, the average population distance from the nearest incinerator is 8.2 kilometers, and the population is relatively evenly distributed across distance groups. This suggests that our correlation estimates are unlikely to be biased by

large differences in sample sizes across groups.

Although we do not attach causal interpretation to the regressions in Table 4, we can gain some insight about the likely direction of the correlation by exploiting two types of events that had a direct effect on exposure to emissions, but no obvious direct effect on lung cancer risk (or health outcomes in general), following them. The first is regulation of emission episodes. We distinguish three regulation subperiods: pre-1991, when regulation on emissions was mild and at its infancy. From 1991 to 2002 - the interval between the first and the second main regulatory reforms. And finally, the post-2003 years when regulation of emissions from incinerators became much tighter and refined. The second type of event is the opening year of the incinerators. As shown in Figure 3a already in 1980 (right before the beginning of our sample) several incinerators were active, implying that they opened earlier on. Indeed, the first opening was in 1950 and several incinerators opened in the 1960s and 1970s. Various incinerators were opened after 1981, partly replacing those that closed. Older incinerators, which opened earlier, tend to be more polluting than newer ones. Consequently, one would expect that living closer to an incinerator before regulatory measures were in place—or near an older, high-emission facility—would be more strongly associated with cancer risk. This is precisely what Figure 6, panel A and Figure 6 panel B illustrate. Living within 2 km of an incinerator that opened before 1982—or within 2-4 km of one—carries a higher risk of lung cancer than living the same distance from a newer incinerator. Similarly, proximity to an incinerator before the 1991 regulations were enacted is associated with a higher cancer risk than exposure at the same distance after the first regulatory measures were introduced. The risk declines even further following the stricter regulations implemented after 2003.

In a similar vein, as a final exercise, we estimate Cox regressions that allow for heterogeneous correlations between distance and lung cancer risk across age groups. The null hypothesis is that proximity to an incinerator poses a greater health risk for older individuals than for younger ones. Indeed, Figure 7 lends support to the null.

5.1.1 Years of Exposure and Instrumental Variable Estimates

The Cox regression estimates face several challenges. First, self-selection in residential location raises an endogeneity issue, particularly if individuals move away from incinerators due to health concerns or anticipated exposure risks. Since relocation is costly, our income and wealth controls help mitigate these concerns.

Second, while distance captures proximity to the emissions source, it does not differentiate between individuals with long-term and short-term exposure. Moreover, since current and past exposures are correlated, it is unclear whether distance reflects recent or cumulative exposure. As a result, the observed correlation between cancer risk and distance likely reflects both latency and cumulative effects.

As a first step in addressing these issues, we adopt an instrumental variable approach. Specifically, we measure cumulative exposure based on the number of years an individual has lived within a given distance of an incinerator.

To implement this strategy, we select a subsample of individuals who, in their first observed year in the dataset, were not exposed to any active incinerator. The Cox regression results in Table 4 suggest that the most relevant exposure zone is within 4 kilometers of an incinerator. Accordingly, we define cumulative exposure as the total number of years an individual in our sample has lived within 4 km of an active incinerator up to time t. We then instrument the length of exposure using the timing of incinerator openings in an individual's municipality. Specifically, we define an indicator variable (*PostOpening*), which equals 1 for all years following an incinerator's opening and 0 otherwise. Identification relies on variation in whether and when an incinerator was introduced in different municipalities. Our sample includes only individuals who were initially unexposed. During the sample period, 21 incinerators were opened in previously unexposed municipalities, with the earliest in 1982 and the latest in 2003. Overall, 7.41% of the sample observations are affected by an incinerator opening.

Table 5 shows the estimation results. Regressions include the usual battery of individual controls (demographics, income and wealth controls) and a battery of municipality controls, namely municipality fixed effects both at the time they were first observed and where they currently live, and year and age fixed effects. In addition we control for distance from the nearest

active *Large Point Source* that is not an incinerator – i.e. large polluting plants - using three dummies corresponding to distance ranges: 0-4 km, 4-10 km and 10-20 km.

The first column reports the Cox estimates, the second the OLS estimates, the third the first stage and the fourth the IV estimates. Length of exposure significantly predicts higher cancer rates in the OLS estimation with an increase in the probability of incurring lung cancer of 0.066 basis point for each additional year of exposure. The IV estimates yield the same result, with a marginally larger coefficient (0.07 compared to 0.066). The first stage estimates show that openings strongly cause the number of year of exposure to increase, implying that there is no concern about the power of the instrument (implied by the extremely high values of the Cragg-Donald and Kleibergen-Paap F Statistic). On average, following an opening the num er of year of exposure increases by 7.3 years – almost twice as large as the overall sample mean. Economically, the estimates imply that a 10 years in increase in exposure cause an increase in the chances of a lung cancer diagnosis of 0.7 basis points, or 10% of the share of cancer diagnosis in the population.

5.2 Openings and closings

The IV approach of the previous Section allows to address the distance endogeneity concern and the fact that distance is an imperfect proxy for cumulative exposure. But it does not distinguish between short and long run effects, and thus does not help separate length of exposure from latency. The latter is critical because heavily back-loaded health effects of current environmental externalities pose serious challenges to the design of regulation.

To speak on these issues, we exploit the unique feature of our data where we can observe the birth and the death of the source of the externality – i.e. the opening and closing of the incinerators. Openings, besides serving as an instrument for the length of exposure can be used to test for the presence of long delayed effects of exposure. This can be done by monitoring the time evolution of lung cancer diagnosis of unexposed individuals who become exposed at some point in time, compared to a control group of people that are never exposed. Closing can be used to document latency. Starting from a sample of individuals exposed to an incinerator and a comparable group of never exposed, latency implies that when the first group experiences

a closure and emissions stop, their cancer risk should converge to that of the never exposed only late in time. Below we test these intuitive implications starring from openings.

5.2.1 Openings

To study openings we rely on an event study framework, where the opening of an incinerator is the event and we are interested in the cancer diagnosis rate at distance d from the event compared to that of a well-constructed control group, following a difference in difference design.

Specifically, we define the treatment as the opening of an incinerator within 4km from the location of the individual. We use the 4km distance because, as already remarked, the results in Table 4 show that longer distance bins do not appear to correlate with lung cancer risk. Since our panel starts in 1980 and we lose the first year because we need to lag income and wealth, we can exploit the openings that occurred from 1982 onward. Indeed, that's the first year for which we can have people observed both before and after the treatment.

Table 2, summarizes all the openings that occurred after 1981. We drop the 2003 opening because it leaves too few observations after its occurrence to to estimate the long run effect. The final sample contains 20 openings spread between 1982 and 1996. The location of these plants is shown in Figure 8. The figure also shows plants that opened before 1981:individuals exposed to these plans are dropped from the sample.

Our long panel and to the availability of the address of the individuals allows us to construct accurate treatment and control groups for our difference-in-differences estimates. However, since the treatment is defined at the geographical level and (some) individuals move, an individual is not always associated to one single location. This implies that one may be within 4 km away from the incinerator in the first years after its opening (and thus be treated according to our definition), and at a longer distance in the years after (and thus be non-treated). This implies we need to be very careful in defining the treatment. First, we cannot simply define the treatment as a dummy equal to one if one has ever been exposed to an incinerator that opened after 1981. This is because, this would require that to be exposed individuals must survive at least until the opening, which would impose a zero-mortality rate before the opening in the treatment group. Since lung cancer diagnosis is very highly correlated with mortality (Figure 1), this would also

mechanically result in a close to zero lung cancer risk in the pre-treatment period among the treated.

To avoid these problems we adopt a synthetic difference-in-differences approach ((**?**) and define our sample as follows:

- At Risk Group. Everyone who was living within 4km in the opening year of a plant or died in a place where, in the future, a plant would open. Therefore, if the plant opens in 1990 and the individual lived within 4km from it in that year, the individual is considered treated. However, if the plant opens in 1990, the individual was initially living somewhere else and than moves to an address that is within 4km from the plant in 1992, he/she is dropped from the sample. Analogously, an individual living in a place where an incinerator will open in the future and dies in that place, is considered treated. Instead, if he/she initially lives in that place but moves to an dies in an unexposed place is dropped from the sample.

- Control Group. Includes all individuals who have never lived within 4km from any active plant. The only exception is when they live within 4km from a plant that will open in the future and they die before the opening. In this case, they are added to the At risk group.

- Always Treated. Anyone who has ever lived within 4km from an active plant that opened before 1982 is dropped from the sample.

- Problematic Plants. A few plants have been used as incinerators for just a few years. It would not be correct to consider those who ever lived close to those plants untreated but it's also not correct to put them in the at risk group because those are not "clean" openings. Therefore, those who have ever lived within 4km from those plants are dropped from the sample.

- Matching. Once the at risk and control groups are formed, we perform a propensity score matching with the following procedure

    - First, we keep all individuals starting the first year they are observed; we drop individuals observed for the first time after the opening of the incinerator.

- For each combination of birth year and gender, we use individual characteristics, municipality characteristics and distance dummies from non-incinerator LPS to calculate the propensity score of being in the *at risk* group. Then, we match each at risk individual to an untreated individual.
- Each untreated individual is assigned a fictitious treatment year corresponding to the actual treatment year of their at risk counterpart.

• Post 1998. Finally, we drop anyone who is treated for the first time from 1998 onwards. The main reason is that to pin down the long run effect of exposure we need enough years ahead following an opening, i.e. at least 20 years. Our sample ends in 2018, thus our sample selection choice.

Table 7 shows summary statistics for the treatment and control group, standardized differences and associated t-tests (last two columns). As can be seen, standardized differences tend to be economically small but normally statistically different, reflecting the large sample size involved. Overall, the two samples seems quite balanced: age, gender and marital status are essentially the same. Treated individuals have somewhat lower income and wealth. They tend to live in municipalities with comparable levels of income per capita and somewhat lower wealth, same share of immigrants and a higher share of hospital. Interestingly, before the treatment cancer risk is of similar size in the treatment and control group and not statistically different.

On this sample we estimate the following difference in difference model [Insert equation form the Tex text in overleaf] where we interact the treatment dummy AtRisk we a set of pre and post-event distance dummies. The later are defined grouping together intervals of 5 years since the opening to have enough observations to measure the effect of the treatment on cancer risk with sufficient precisions. The model includes controls for individual economic (one year lagged income and wealth quintile dummies) and demographic variables (age range dummies - younger than 35, between 35 and 59 and older than 60 - education level dummies, marriage status, dummies, average municipality income). We also control for municipality level variables (average municipality wealth, share of immigrants, municipality, log municipality population, number of hospitals, distance from the nearest hospital and distance from the nearest LPS dummies). θ_g, λ_t and δ_m are group, year and municipality fixed effects, respectively. Group

fixed effects are defined as the interaction between event year-either actual or fictitious-and $AtRisk_i$.

Figure 10 shows the event study coefficients together with their 95% confidence interval, in deviation from the first-time interval prior to the opening. There is no evidence of pre-opening difference in cancer risk between treated and control groups, i.e., we can rule our pre-trends. In the first five years since the opening cancer diagnosis risk is somewhat higher among the treated but effect is not statistically significant. It is somewhat higher in size (and statistically significant at the 5%) 5-to-9 years after the opening , while the point estimate becomes lower 10-14 and 15-20 years from opening , losing statistical significance. But it than rebounds two decades or more after the opening, reaching its highest level and easily rejecting the null that the effect is zero. This is consistent with serious health effects of emissions externalities manifest with decades-long delays since the start of exposure, an effect so far unproved in the literature. Economically, after 20 years since opening cancer risk increases by about 2 basis points, or 21% of the sample average, a non-negligible effect.

To gain confidence on the effects in Table 8 we allow for heterogenous effects across people of different characteristics at time of opening. We show in columns 2 and 3 that exposure to an additional source of pollution should have stronger effects on people that, at the time of opening, are older. One reason for this is that, among the individuals exposed to the emissions from the opening, there may be some who already had some carcinogenic cells and this is more likely among the elderly. For these people the opening acts as a cancer promoter, i.e. it speeds up a process that had already begun. Results strongly support this idea. We distinguish two age groups at time of opening (20-39 and 40 or more). Interestingly exposure to a new incinerator has no health – neither current nor delayed at any distance from the event - for the first youngest group aged 20-39 at opening. The effect is magnified for those aged 40 or more at opening who mark a higher risk of cancer diagnosis of 5.9 basis points – almost a third of the estimation sample mean - compared to non-treated people belonging to the same age bracket and characteristics. This is quite remarkable. Furthermore, for this group exposure to the opening of an incinerator causes an increase in lung cancer risk already at the shorter delays of 0-10 years and 10-20 years. Interestingly, the effect is stronger for the first group. One rationale is that

this "shorter" term effects of pollution externalities hit relatively fragile individuals which not surprisingly are those that at opening are relatively aged.

Overall our evidence implies that exposure are markedly delayed effects on cancer risk, with length of the delay being very long for those still young at time of imitating exposure and significantly shorter for those aged. Average effects, measured in Figure 10 are a combination of these heterogeneous effects.

We have also run our diff-in-diff model distinguishing between early and late openings to capture differences in the vintage of the incinerator and thus in annual volume of their emissions. Consistent with this logic results in Table 8, columns 4 and 5 show that early openings (those before 1991, the median opening year) tend to have more front loaded effects on lung cancer risk.

As a final exercise, we distinguish between individuals with low and high wealth prior to the opening. Using each individual's first observation, we estimate age-year-specific wealth quintiles and compare outcomes for those in the top and bottom quintiles. The results, presented in columns 6 and 7 of Table 8, reveal a striking disparity. The effect is effectively zero for the wealthiest 20% of the population, whereas for the poorest 20%, it amounts to nearly half of the baseline mean.

5.2.2 Closings

The long-delayed effects of lung cancer caused by the opening of an incinerator can reflect both latency as well as the fact that what causes cancer is cumulative exposure to emissions and the later piles up slowly. To separate latency from cumulative exposure we rely on plant closures. At closure exposure stops cumulating. Delayed effects post-event can thus identify latency.

To study closures we follow a similar procedure, with adaptation, as for openings in the design of the sample. Namely, in the group at risk we include all people who either lived within 4km from an incinerator in the closing year or died in a place where an incinerator would eventually close. The construction of the control group is a bit different. Since we want to compare individuals who, before the treatment, are similar, those who have never been exposed to an incinerator are no longer a valid control group. A proper control would be made

of (otherwise equal) individuals who are exposed to non-closing plants, ideally even after the fictitious closing year they will be assigned to with the matching procedure. Therefore, in defining this group, we need to make sure that they are consistently exposed to a plant and, at the same time, make sure that the control group status is not correlated with the probability of surviving. If we simply defined it as a dummy equal to one if an individual is exposed to a plant for at least one year, this dummy is more likely to turn on the longer one survives. What we did, instead, is to define membership in the control group as a dummy equal to one for each individual that meets both of the following conditions:

- He/she has been exposed to active plants for at least 50% of the time (i.e. in at least 50% of the years in which you he/she is observed)
- He/she has lived close to an active plant in the last year is observed. This second condition makes sure that individuals in the control group are exposed not only in the early years but also later.

As with the openings, in order to have enough years to the detect the long run effects of a closing we keep only the closing before 1998. Table 3 lists all the closing occurring before 1998. The final sample includes 18 closings. Figure 9 highlights on the map the closing plants (blue), the control plants (green) and the plants dropped from this exercise. Control plants tend to be geographically close to the closing plants.

Characteristics of the treated and control are shown in Table 9 while Table 10 shows covariate balance tests. Before closing cancer rate is higher among the treated (8.4 versus 5.5 basis points). Given that the two groups are exposed to different plants, it is reasonable to observe some difference in the health outcomes before the closing. The difference is also consistent with the fact that the closed plants tend to be the oldest, least efficient and thus most harmful. Indeed the plant average opening year in the treatment group is 1972 while it is 1979 in plants in the control group. Covariates instead are balanced. Individuals demographics such as age, gender and marital status , as well as income and wealth levels are quite similar, as a coarse look to their level and to their standardized difference reveals. The same hold true for municipality and Parish characteristics (except for distance from an hospital which is higher among the treated

than the control. But this if anything may play down the beneficial effects of the closing). Where the two samples differ is in the distribution of the distance from a LPS, which appears shifted to the left (shorter distance) in the treated group.

Figure 11 shows the main event study result using the same intervals of distance from the event as in the opening estimates. The evidence is quite clear. Up to 10 years after the closing there is no difference in mortality risk between people who "benefit" from the closing and the control group of individuals who continue to accumulate exposure to emissions from an incinerator. Cancer risk starts to decrease among the treated after 10 to 19 years, though they are marginally statistically significant. After 20 or more years from closing we see a precisely estimated drop in mortality relatively to the control, revealing substantial latency. Economically, the drop is sizeable, falling by 2.82 basis points, almost 1/3 of the average lung cancer rate among the treated before closure.

As with openings, we further strengthen the evidence in Table 11 by examining responses to closings across different groups at the time of closure. The greatest benefits are observed among the oldest individuals (columns 2 and 3), those with the highest cumulative exposure (columns 4 and 5), and the poorest individuals (columns 6 and 7). Notably, these benefits emerge more than a decade after closure, further reinforcing the evidence of substantial latency.

6 Robustness

In this section, we conduct three main robustness tests to verify that the diff-in-diff results are not driven by spurious correlations.

6.1 Smokers

Since the main cause of lung cancer is smoking, we want to make sure that our results are not driven by our lack of information about this important variable. In particular, we want to ensure that the opening and closing results are not driven by smokers. This would be the case if the opening or closing of an incinerator was associated with an increase or decrease in the number of smokers living near the site where an incinerator was opened or closed. This hypothesis is

not so remote: Smokers might be people who care less about their health, and so they are more willing to stay next to an incinerator or to move near one.

The only data on smokers we have concern women who became pregnant after 1991. Pregnant women are asked about their smoking habits during routine visits, and this information is recorded.

We used these data to test whether the probability that a woman living near an incinerator's site smokes increases after the incinerator's opening and decreases after its closing.

6.1.1 Openings

Table 12 shows the characteristics of the population living near a site of an incinerator that was opened during the period 1993-2018. There were 7 openings during this period. In panel A, we report the share of smokers. In panel B, we report the characteristics of the mother. In panels C-D we report the usual set of municipality and parish characteristics, while in panel E we report the distance from other LPS.

Table 14 reports the covariate balance test for the opening sample for the period prior to the opening. We do observe a higher incidence of women smokers in the treatment group (women living near an incinerator site) than in the control group (25 versus 23 percent). There is also a difference in income, with people living near an incinerator site prior to an opening being poorer and less educated. While this pattern suggests that incinerators are opened in poor areas, it does not imply that an opening leads to an increase in the percentage of residents who are smokers.

To address this question, we need to resort to the following difference in difference estimation:

$$Smoker_{it} = \sum_{j=-5, j\neq -1}^{5} \gamma_j Treated_i \times TimeSince_{gt}^{j} + \sum_{j=-5, j\neq -1}^{5} \eta_j TimeSince_{gt}^{j} + \mathbf{X}_{it}\beta + \lambda_t + \delta_g + u_{it}. \quad (1)$$

$Smoker_{it}$ is a dummy equal to one if individual i is a smoker in year t. $Treated_i$ is a dummy equal to one for being in the group *at risk*, i.e. the group of those who live in places that eventually experience an opening. $TimeSince_{gt}^{j}$ is a dummy equal to one in the j-th year since

the opening. For the *at risk* group, this is equal to the actual opening year, while for the control group, this is equal to the fictitious opening year they were assigned to through the PS-matching procedure. This variable does not only depend on calendar year t, but also on the opening year g. The diff-in-diff coefficients are, therefore, given by γ_j. $\mathbf{X}_{it}$ is a vector of individual, municipality and parish characteristics. λ_t and δ_g are year and group (opening year) fixed effects, respectively.

Figure 12 reports the key estimates of this regression in the form of an event study. It reports the coefficient of the treatment interacted with the year dummy, where the omitted dummy is the event year minus 1. If anything, the percentage of women smokers decreases after the opening, albeit the effect is not statistically significant. What we can reject is that the percentage of women smokers increased after the opening.

6.1.2 Closings

During the sample period, there were 16 incinerator closures. Tables 14 and 15 report the sample statistics and the covariate balance test, respectively.

For these closings we use the same difference-in-difference specification as for the openings. Figure 13 reports the estimates of this regression in the event-study format. The closing of an incinerator is associated with a slight decrease in the percentage of women smokers, but this effect is not statistically different from zero.

6.2 Movers

The above analysis shows that the openings and closings of incinerators are not associated with an increase or decrease of the smoking population. Yet, there could be other reasons why our results were spurious. For example, individuals who are more concerned about their health may react to the information that an incinerator is about to open by moving to cleaner areas. If the people more concerned about their health are more healthy in general and less likely to develop lung cancer, then our results could be spurious.

One way to test this hypothesis is to check whether the opening/closing of an incinerator affects the probability of moving in/out of a certain location. To conduct this test, we construct a synthetic diff-in-diff analysis of the openings at the parish-level, analogous to the one conducted

at the individual-level in the main analysis.

We define treated and untreated parishes according to the distance between their centroid and the incinerator. Then, we match each treated parish to an untreated one based on their propensity score, estimated using some parish characteristics such as the average wealth, average income, share of immigrants, and share of people with higher education.

Table 20 shows the characteristics of the sample. In panel A, we report the share of in-movers and out-movers as a fraction of the parish population. There are more people moving in and more people moving out in the treatment group than in the control. Yet, this difference does not seem to be driven by income. Poorer people (from the bottom quartile of the income distribution) are more likely to move out of a parish close to a newly opened incinerator than those who are richer (top quartile). In contrast, both rich and poor seems to move in at the same rate.

In panel B, we show some characteristics of the parishes and, in the last panel, we show the distance dummies between the parish centroid and non-incinerator LPS plants. In Table 21, instead, we show the covariate balance test.

Figure 17 reports the estimates of a diff-in-diff regression, presented as an event study. We want to test whether people are more likely to leave a parish when an incinerator opens nearby. In fact, we do not see this effect. If anything, there seems to be a spike in out-movers two years before an opening. This could signify an anticipation effect, people move when an incinerator is announced, not when it is opened. As Figure 17 shows, this effect is not statistically significant.

6.3 House Prices

Another way to check whether the opening/closing affects the appeal of a given location is by looking at house prices. In fact, lower house prices after the opening or higher house prices after the closing may indicate that the incinerator is perceived negatively by those who live around it.

For this specification, we construct a diff-in-diff sample using the same procedure as before. However, this time, the final dataset will be at the transaction level. Therefore, we compare the house prices around the incinerators before and after the opening/closing. Since the housing data is only available from 1992, we will focus on openings/closings that occurred after that year.

6.3.1 Openings

In Table 16, we show the characteristics of the housing openings sample. In panel A, we show the average selling price in thousands of DKK and its logarithm. In panel B, we show the characteristics of the houses sold, which include the number of floors, the number of rooms, the building year, and the current age. In Table 17, we show the covariate balance test.

In Figure 14, we show the main event study for the openings. It seems there is neither an anticipation effect nor an immediate effect. However, two years after an opening, we see the beginning of a downward trend, which becomes statistically significant in year 4. The effect is not large (about 5 percent) and is consistent with what has been found in the literature (CITE). This effect would be problematic if it led to a sorting of the population starting in year 4. In fact, we do not see any evidence of people moving out of the area during that period (see Figure 17).

6.3.2 Closings

In Tables 18 and 19, we replicate Tables 16 and 17 using the closings sample. Just like in the main individual-level analysis, we do not expect the two groups to be perfectly balanced because, even though they are both initially exposed to an incinerator, such incinerators have different characteristics.

Figure 15 shows the diff-in-diff estimates in the form of an event study. We do not see any statistically significant change in house prices until year four, where we see a *decrease* of about 7 percent. Note that we would have expected an increase, since an area becomes more appealing when an incinerator closes. This gives us confidence that there is no sorting effect of the population around closing.

References

Anderson, M. L. (2020). As the Wind Blows: The Effects of Long-term Exposure to Air Pollution on Mortality. *Journal of the European Economic Association*, 18(4):1886–1927.

Banzhaf, H. S. and Walsh, R. P. (2008). Do People Vote with Their Feet? An Empirical Test of Tiebout. *American Economic Review*, 98(3):843–63.

Banzhaf, S., Ma, L., and Timmins, C. (2019). Environmental Justice: The Economics of Race, Place, and Pollution. *Journal of Economic Perspectives*, 33(1):185–208.

Barreca, A. I., Neidell, M., and Sanders, N. J. (2021). Long-run Pollution Exposure and Mortality: Evidence from the Acid Rain Program. *Journal of Public Economics*, 200:104440.

Chay, K., Dobkin, C., and Greenstone, M. (2003). The Clean Air Act of 1970 and adult mortality. *Journal of risk and uncertainty*, 27:279–300.

Chay, K. and Greenstone, M. (2003a). Air Quality, Infant Mortality, and the Clean Air Act of 1970. *NBER Working Paper*.

Chay, K. Y. and Greenstone, M. (2003b). The Impact of Air pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, 118(3):1121–1167.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, 315(16):1750–1766.

Christensen, P., Sarmiento-Barbieri, I., and Timmins, C. (2022). Housing Discrimination and the Toxics Exposure Gap in the United States: Evidence from the Rental Market. *Review of Economics and Statistics*, 104(4):807–818.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant openings and Closings. *American Economic Review*, 105(2):678–709.

Currie, J. and Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn from California's Recent Experience? *Quarterly Journal of Economics*, 120(3):1003–1030.

Currie, J. and Schmieder, J. F. (2009). Fetal Exposures to Toxic Releases and Infant Health. *American Economic Review*, 99(2):177–183.

Currie, J. and Walker, R. (2011). Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, 3(1):65–90.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review*, 109(12):4178–4219.

Deryugina, T. and Julian, R. (November 2023). The Long-run Effect of Air Pollution on Survival. *NBER Working Paper*.

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris Jr, B. G., and Speizer, F. E. (1993). An Association Between Air Pollution and Mortality in Six US Cities. *New England Journal of Medicine*, 329(24):1753–1759.

Filigrana, P., Milando, C., Batterman, S., Levy, J. I., Mukherjee, B., Pedde, M., Szpiro, A. A., and Adar, S. D. (2022). Exposure to Primary Air Pollutants Generated by Highway Traffic and Daily Mortality Risk in Near-Road Communities: A Case-Crossover Study. *American Journal of Epidemiology*, 191(1):63–74.

Graff Zivin, J. and Neidell, M. (2013). Environment, Health, and Human Capital. *Journal of Economic Literature*, 51(3):689–730.

Grainger, C. and Ruangmas, T. (2018). Who Wins from Emissions Trading? Evidence from California. *Environmental and Resource Economics*, 71(3):703–727.

Henneman, L. R., Choirat, C., and Zigler, C. M. (2019). Accountability Assessment of Health Improvements in the United States Associated with Reduced Coal Emissions between 2005 and 2012. *Epidemiology*, 30(4):477.

Hernandez-Cortes, D. and Meng, K. C. (2020). Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon Market. *NBER Working Paper*.

Isen, A., Rossin-Slater, M., and Walker, W. R. (2017). Every Breath You Take Every Dollar You'll Make: The Long-term Consequences of the Clean Air Act of 1970. *Journal of Political Economy*, 125(3):848–902.

Lubin, J. H., Caporaso, N., Wichmann, H. E., Schaffrath-Rosario, A., and Alavanja, M. C. (2007). Cigarette Smoking and Lung Cancer: Modeling Effect Modification of Total Exposure and Intensity. *Epidemiology*, pages 639–648.

Nadler, D. L. and Zurbenko, I. G. (2014). Estimating Cancer Latency Times Using a Weibull Model. *Advances in Epidemiology*, 2014.

Næss, Ø., Nafstad, P., Aamodt, G., Claussen, B., and Rosland, P. (2007). Relation between Concentration of Air pollution and Cause-Specific Mortality: Four-Year Exposures to Nitrogen Dioxide and Particulate Matter Pollutants in 470 Neighborhoods in Oslo, Norway. *American Journal of Epidemiology*, 165(4):435–443.

Pennington, A. F., Strickland, M. J., Klein, M., Zhai, X., Bates, J. T., Drews-Botsch, C., Hansen, C., Russell, A. G., Tolbert, P. E., and Darrow, L. A. (2018). Exposure to Mobile Source Air Pollution in Early Life and Childhood Asthma Incidence: The Kaiser Air Pollution and Pediatric Asthma Study. *Epidemiology*, 29(1):22.

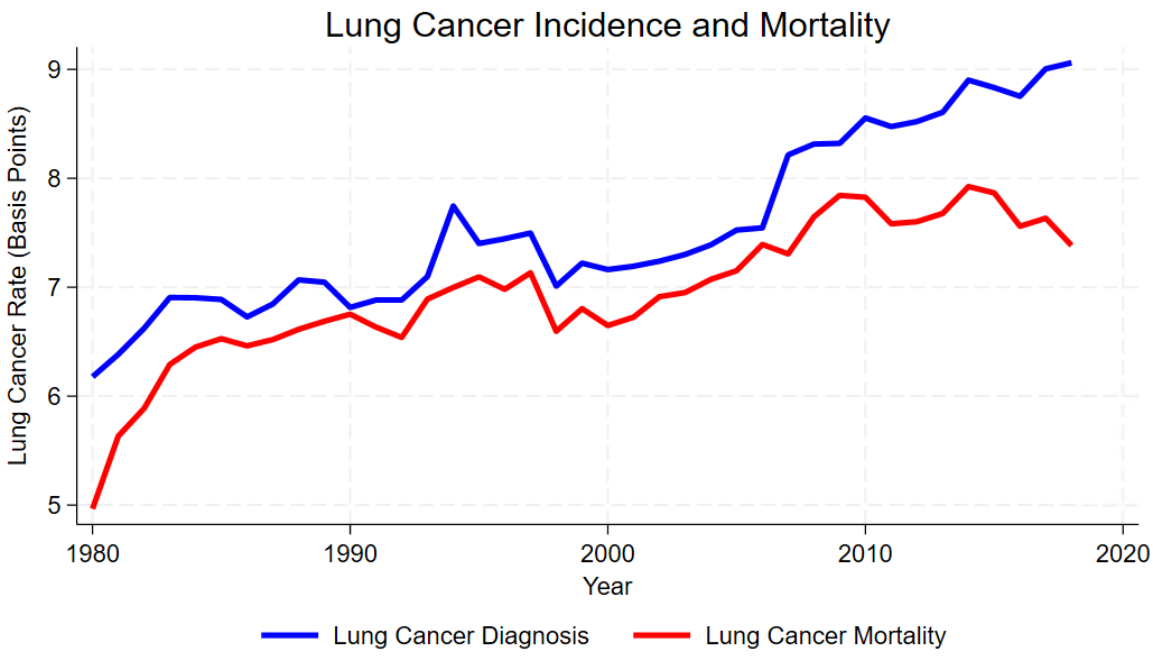
Pitot, H. C. (1993). The Molecular Biology of Carcinogenesis. *Cancer*, 72(S3):962–970.

Santibáñez-Andrade, M., Chirino, Y. I., González-Ramírez, I., Sánchez-Pérez, Y., and García-Cuellar, C. M. (2019). Deciphering the Code Between Air Pollution and Disease: The effect of Particulate Matter on Cancer Hallmarks. *International Journal of Molecular Sciences*, 21(1):136.

Sullivan, D. M. (2017). The True Cost of Air Pollution: Evidence from the Housing Market. *Working Paper*.

Weisburger, J. H. (1990). The Mechanism of Lung Carcinogenesis and Smoking Cessation. *Epidemiology*, 1(4):314–317.

Figure 1: Lung Cancer Diagnosis and Mortality Rates



Figures

Figure 2: Lung Cancer Diagnosis Incidence by Age Group

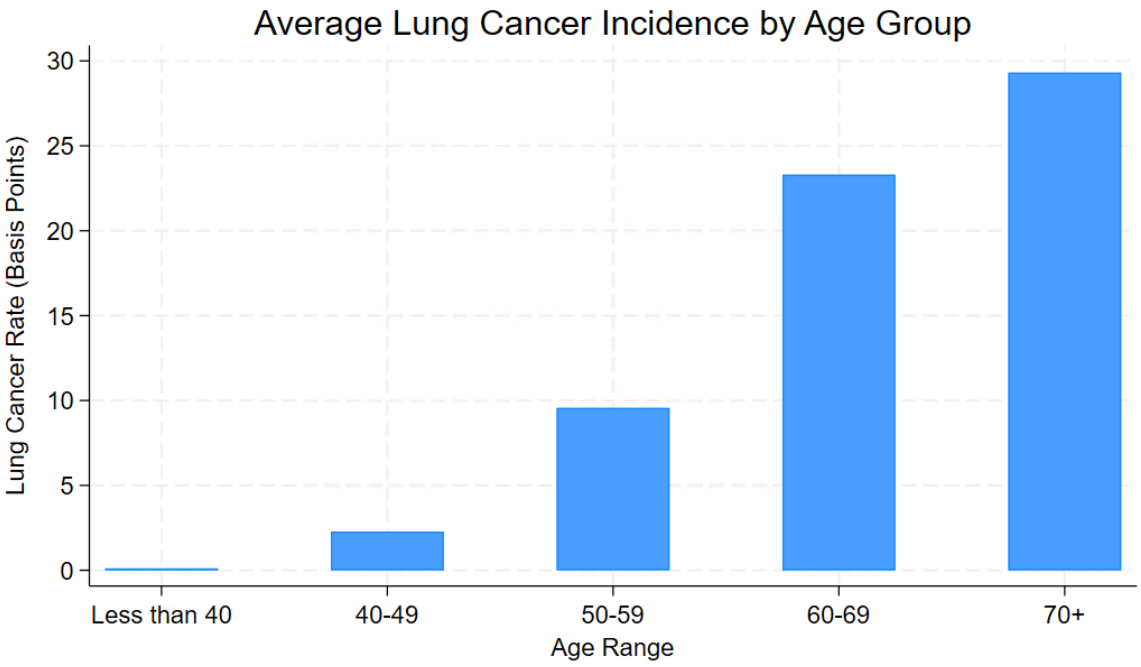
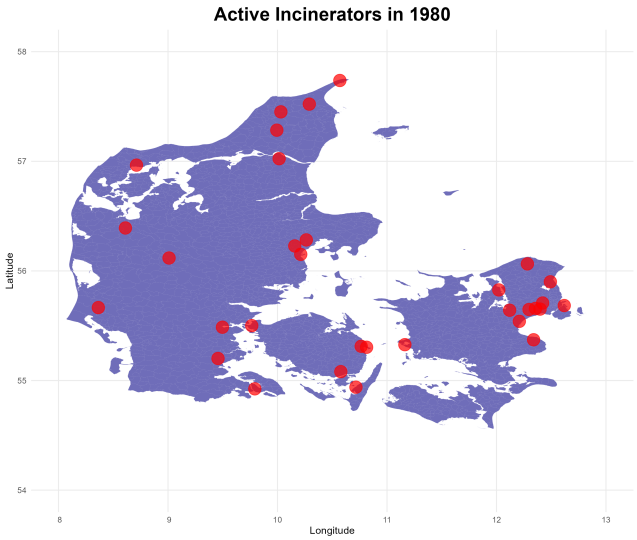
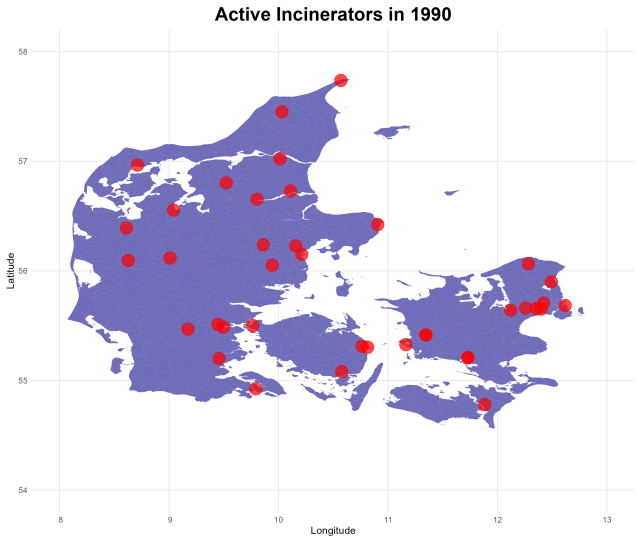


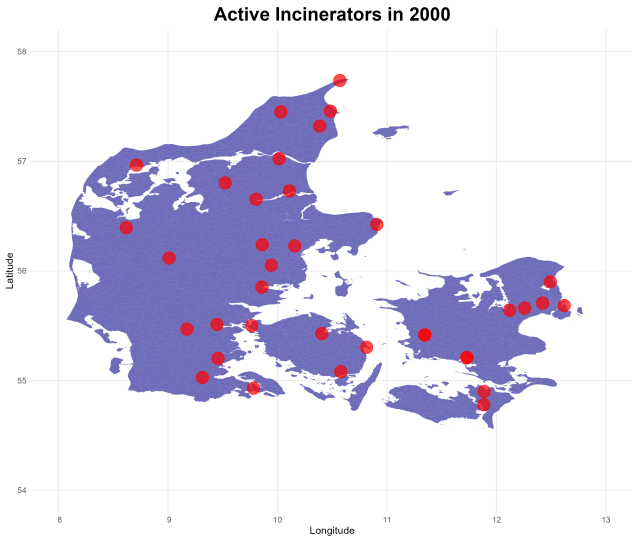
Figure 3: Distribution of the Active Incinerators in Denmark



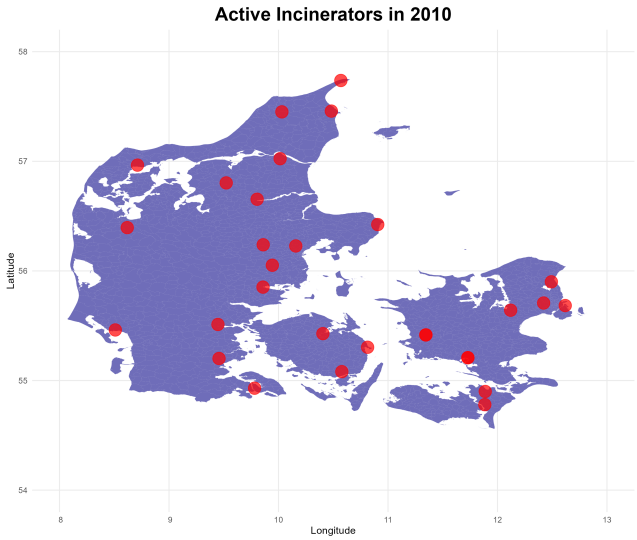
(a) Active Incinerators in 1980



(b) Active Incinerators in 1990



(c) Active Incinerators in 2000



(d) Active Incinerators in 2010

Figure 4: Number of openings, closings and total active incinerators by year

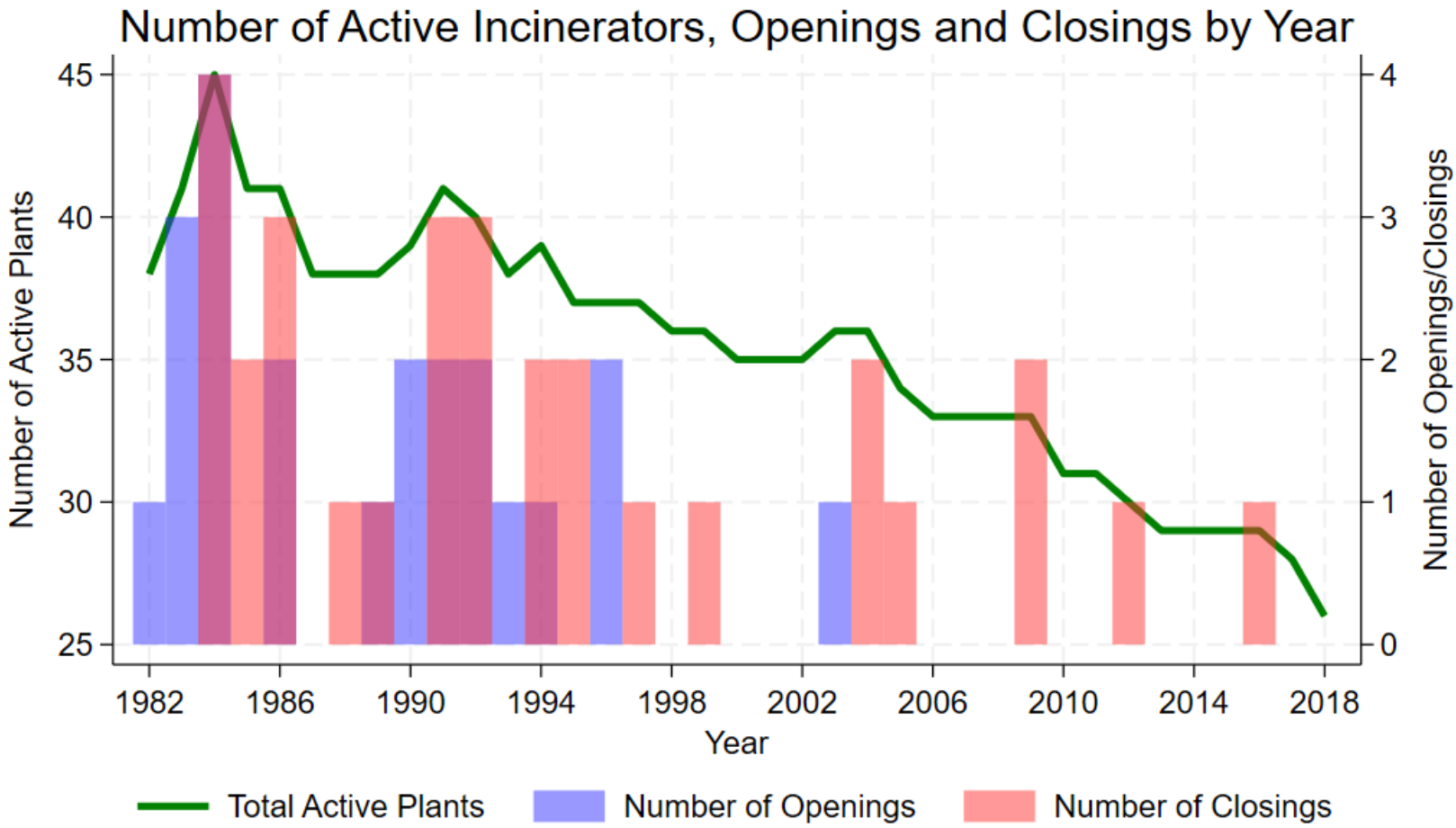


Figure 5: Contemporaneous Cox Regressions by Regulation Period

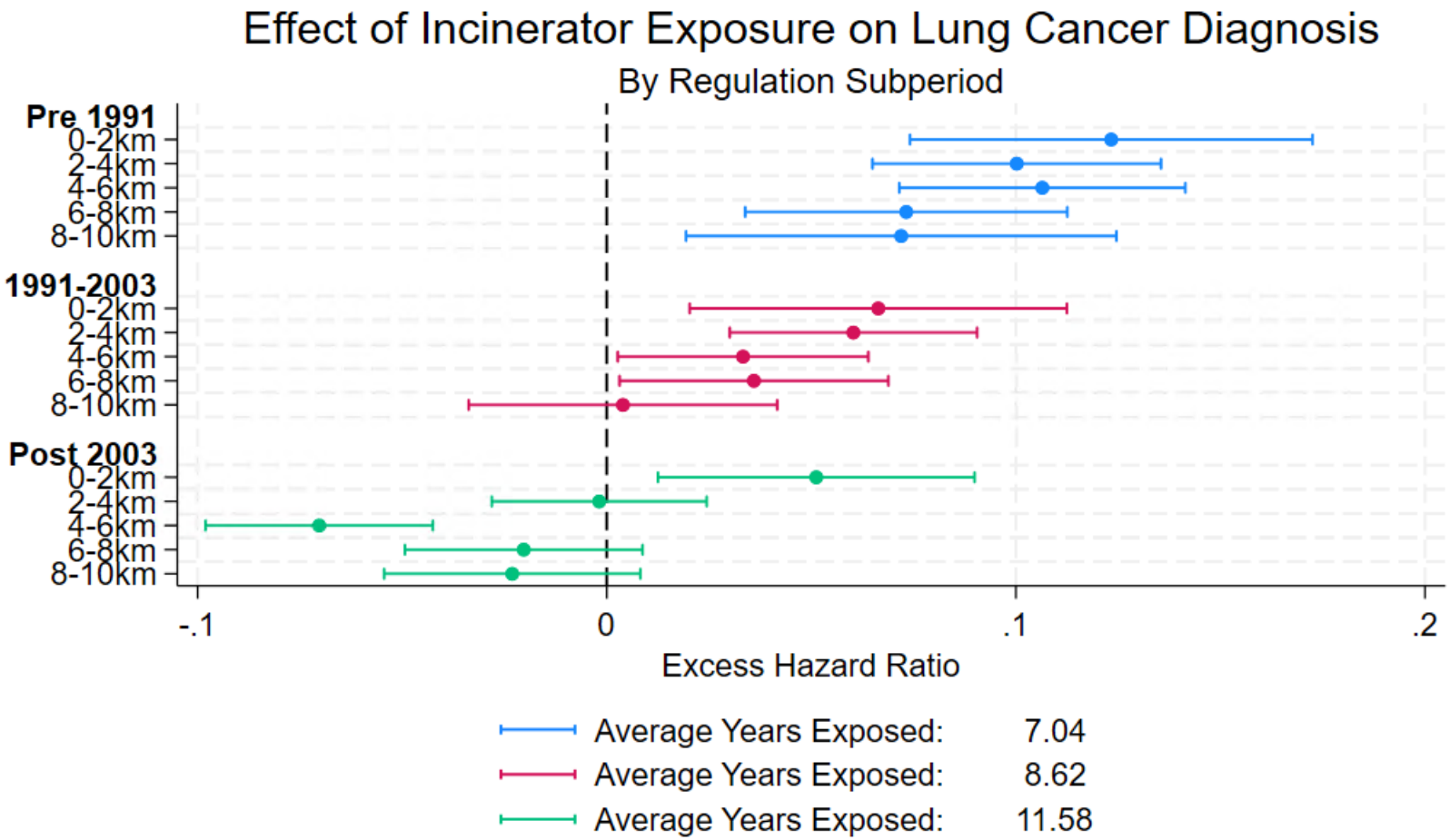


Figure 6: Contemporaneous Cox Regressions by Opening Year

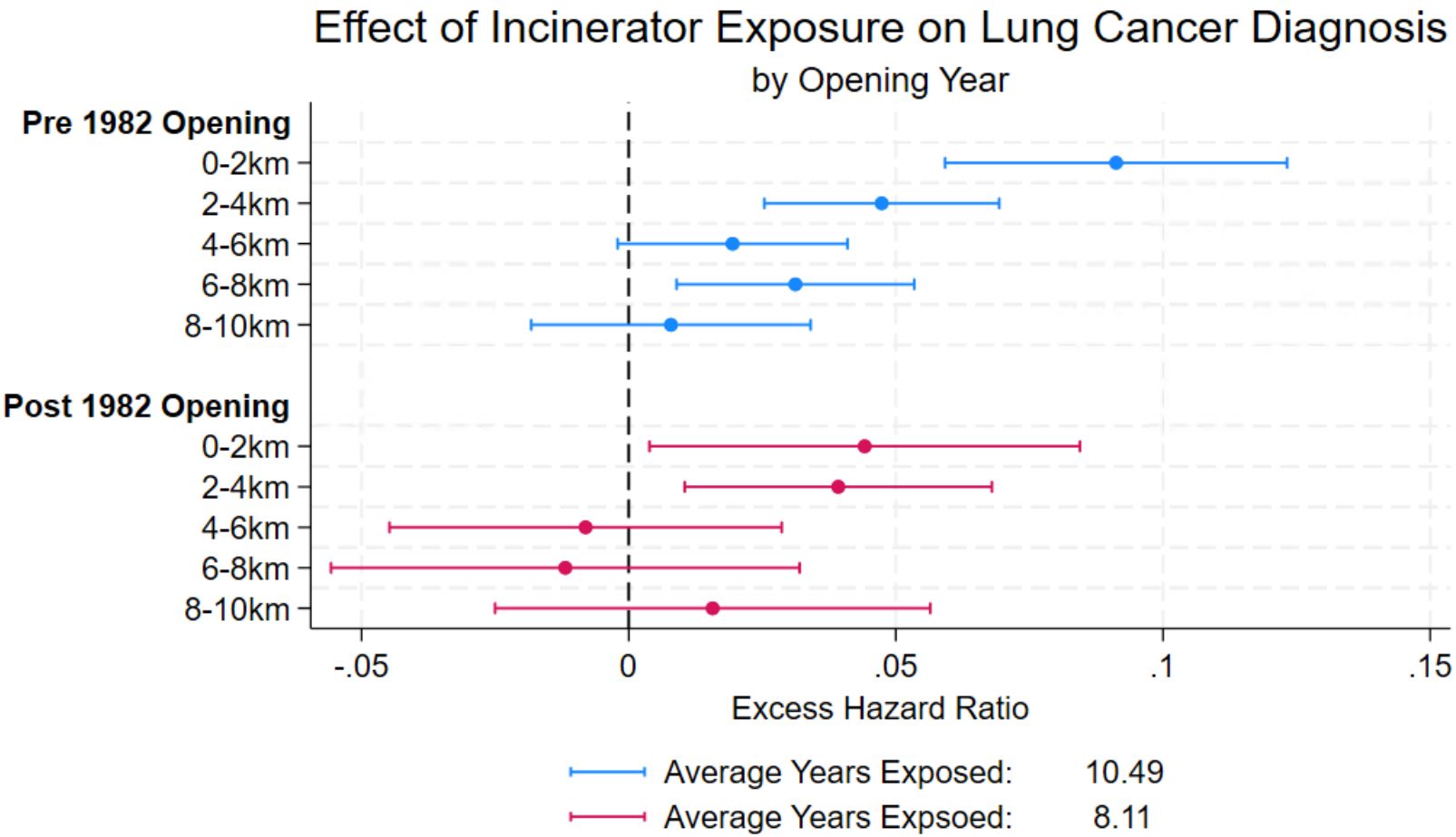


Figure 7: Contemporaneous Cox Regressions by Age Range

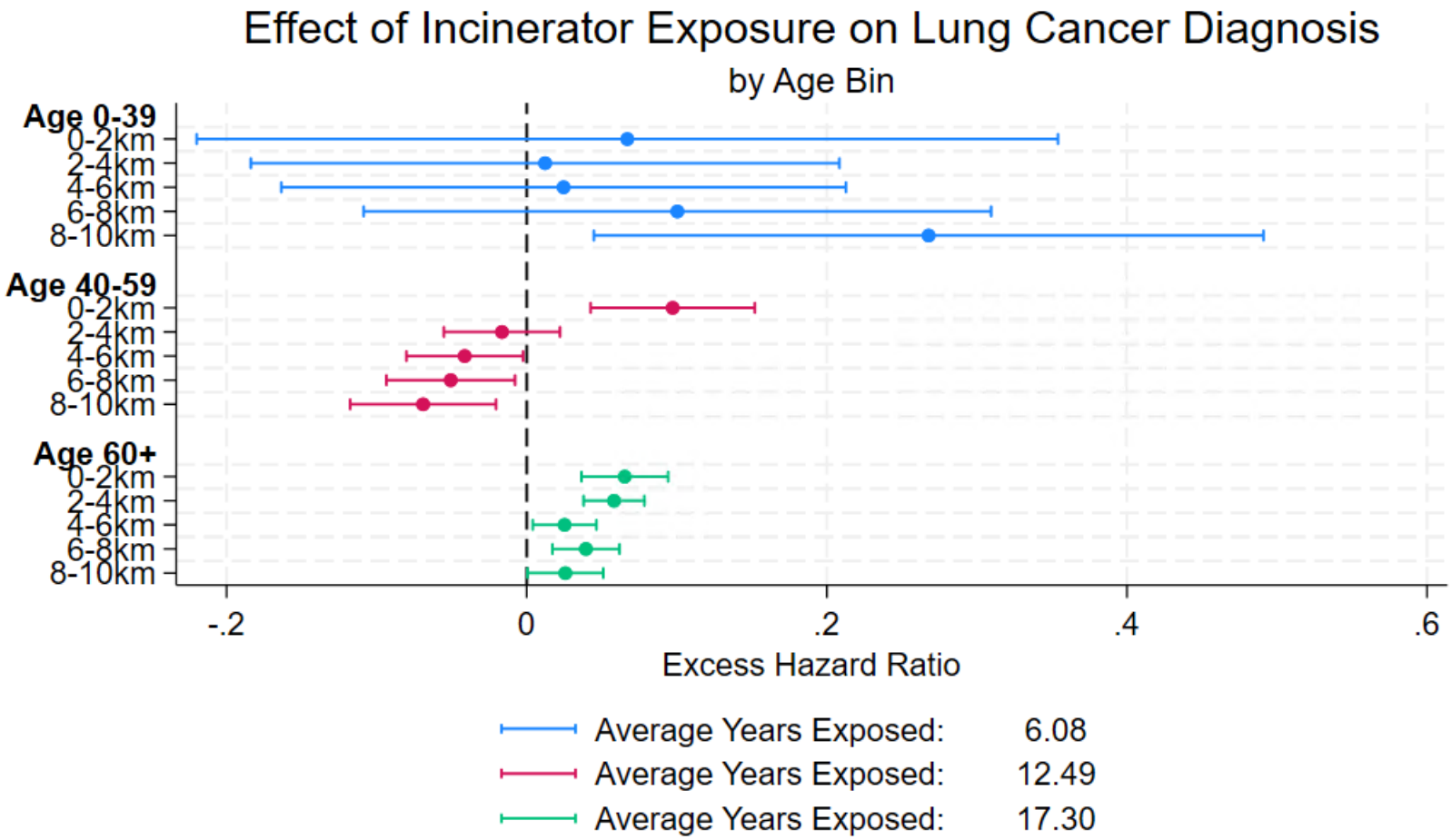


Figure 8: Map of the Openings

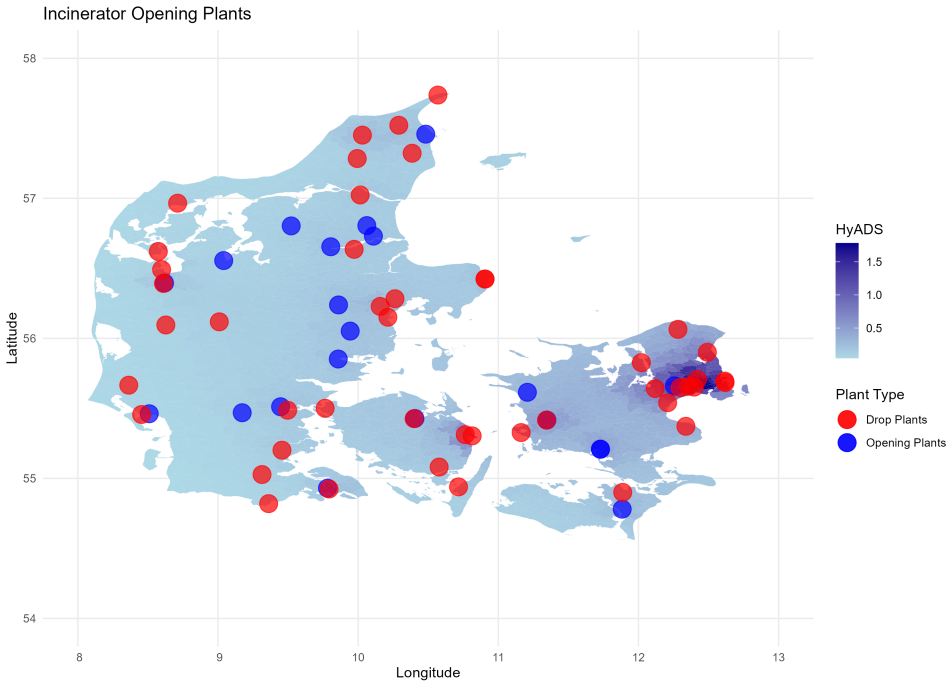


Figure 9: Map of the Closings

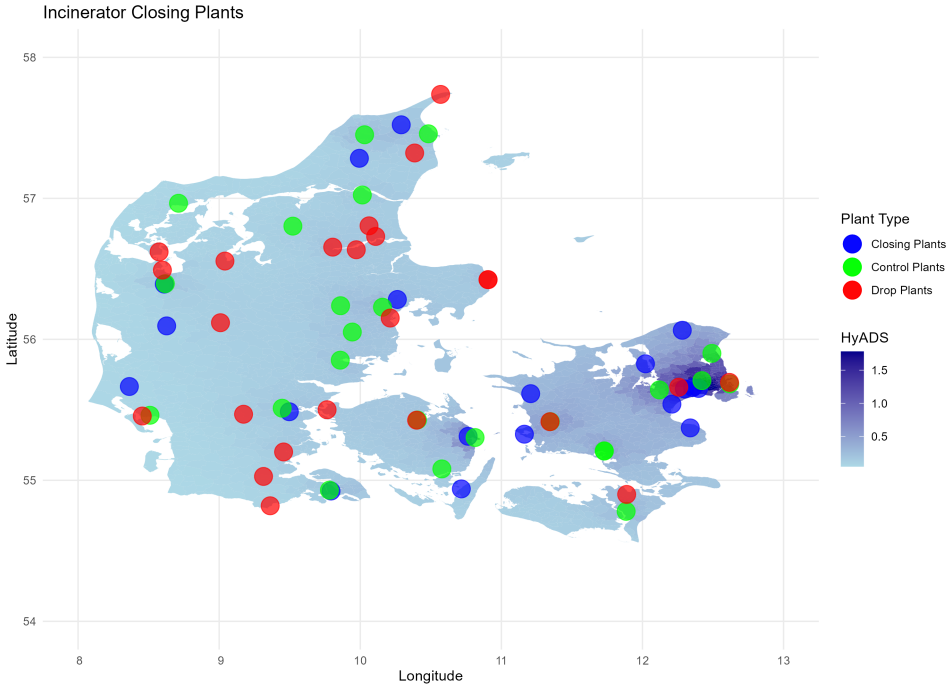


Figure 10: Main Openings Event Study

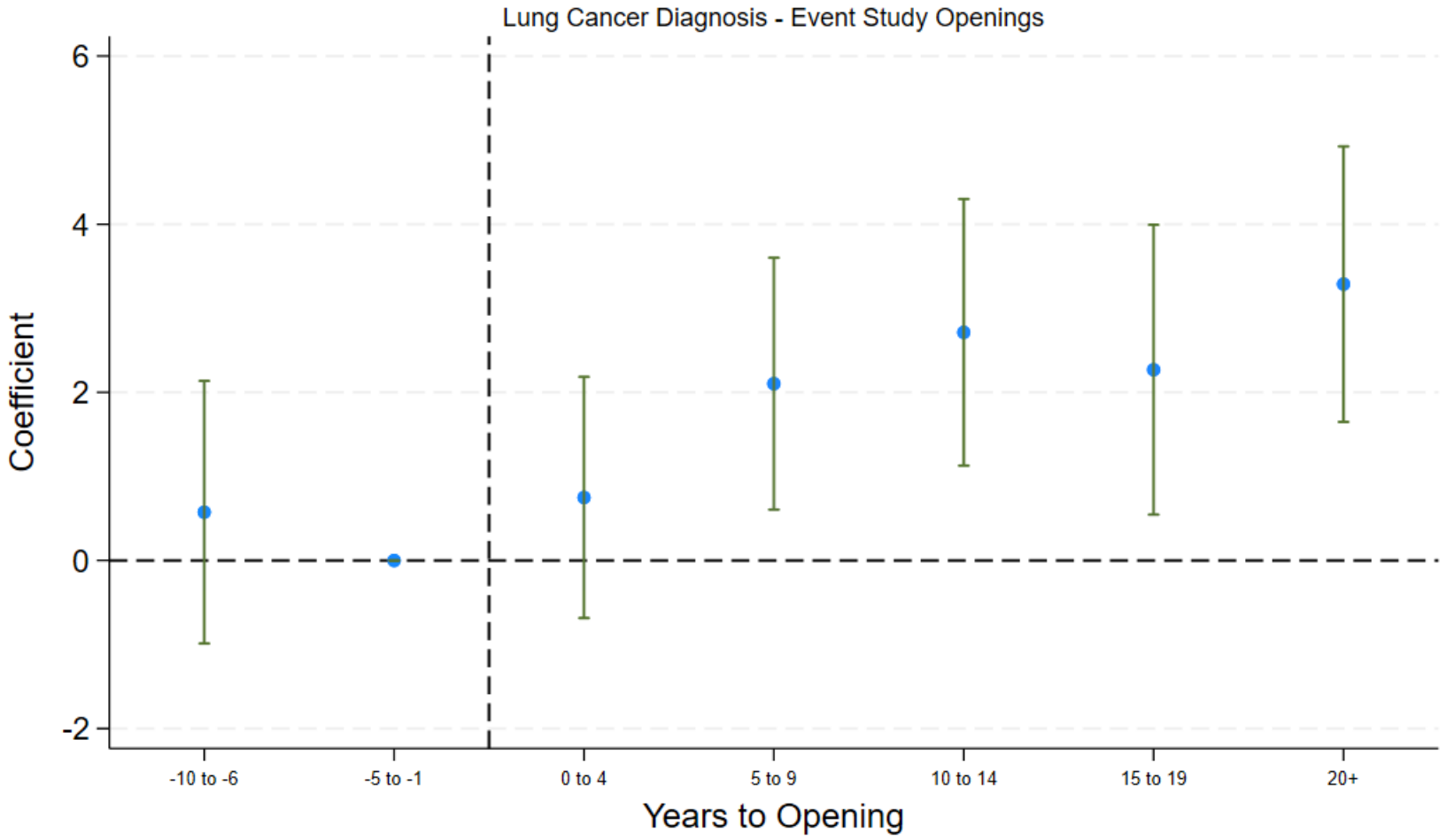


Figure 11: Main Closings Event Study

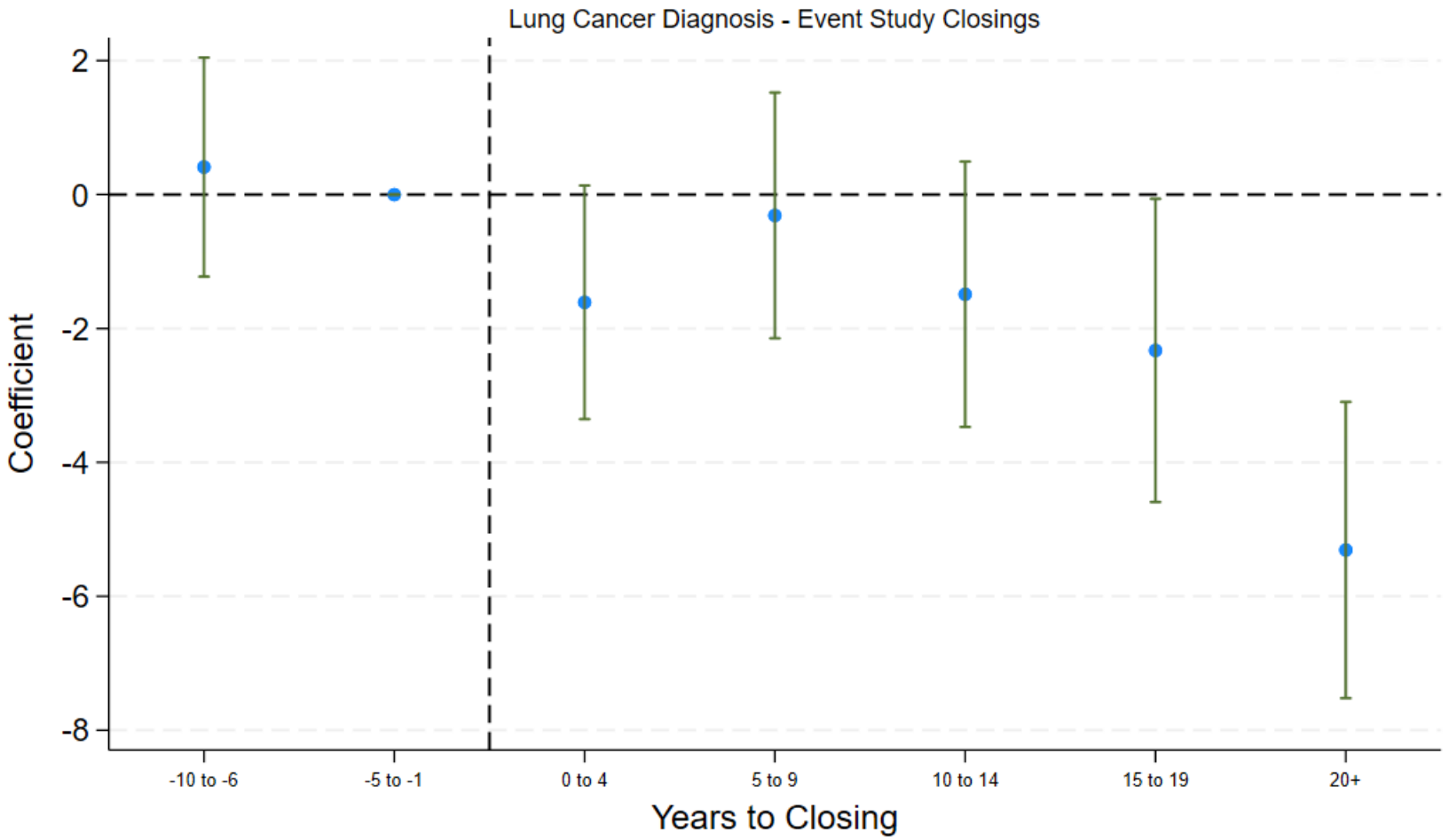


Figure 12: Openings Event Study: Probability of Smoking

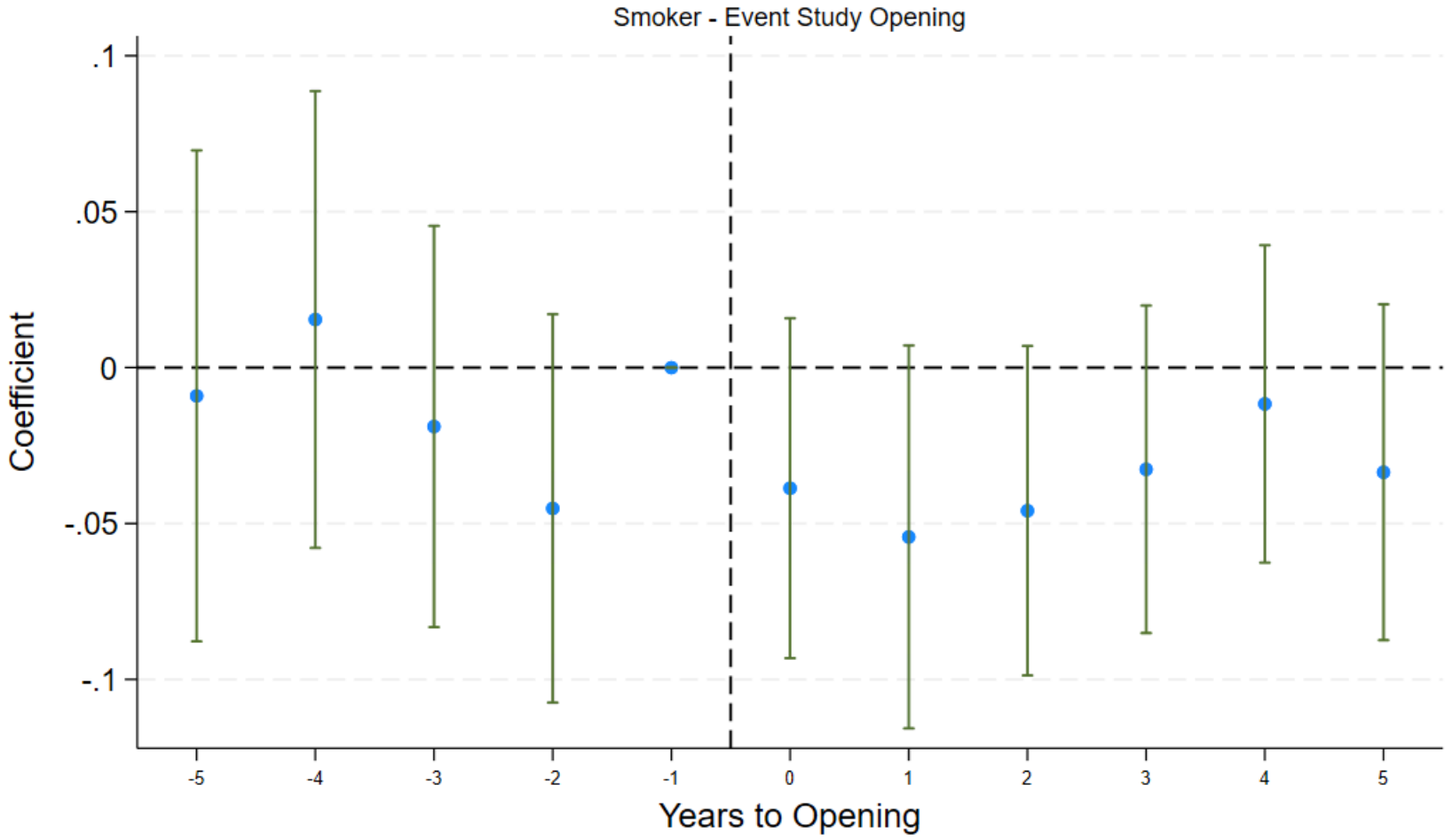


Figure 13: Closings Event Study: Probability of Smoking

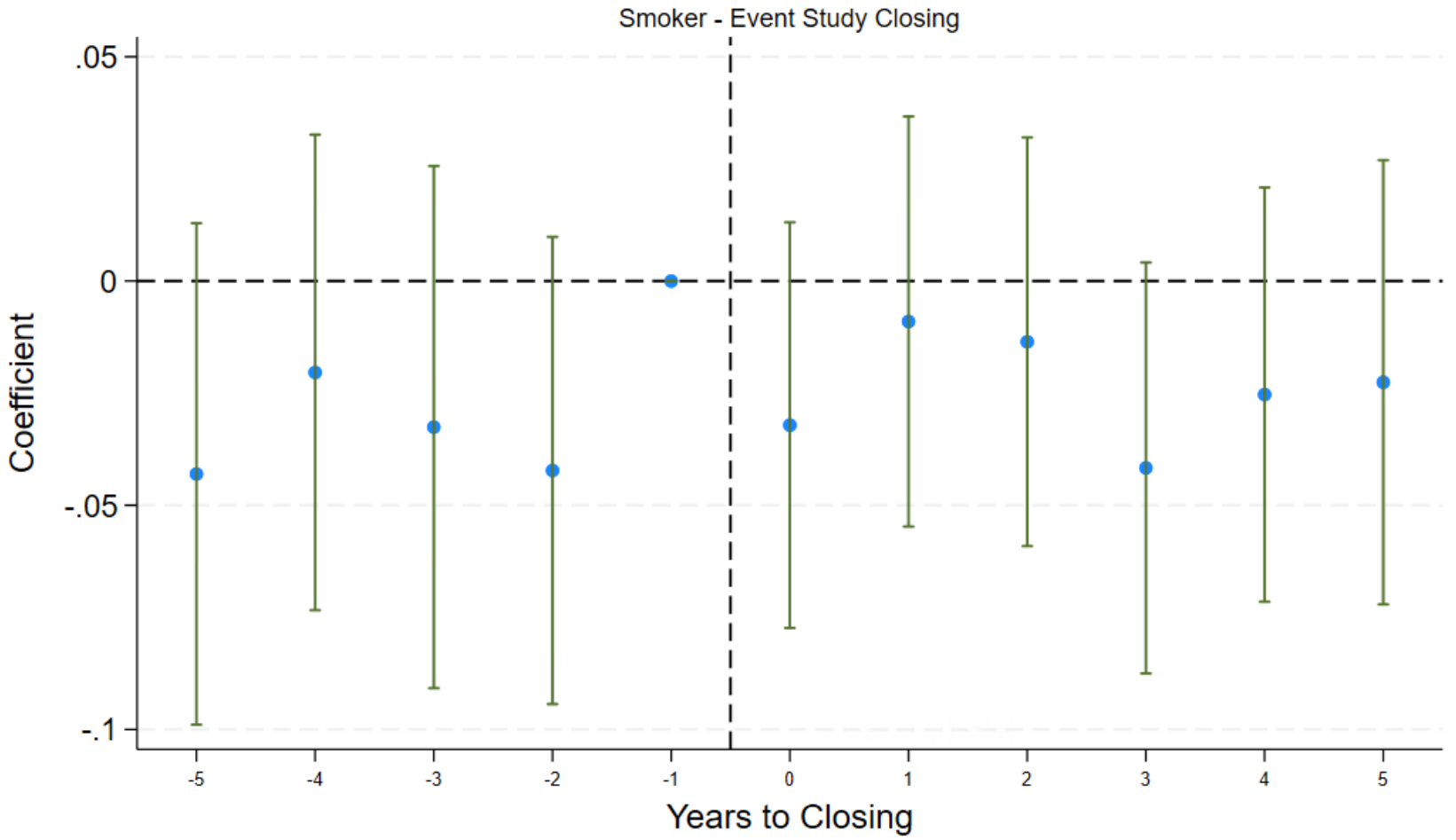


Figure 14: Openings Event Study: Log Selling Price

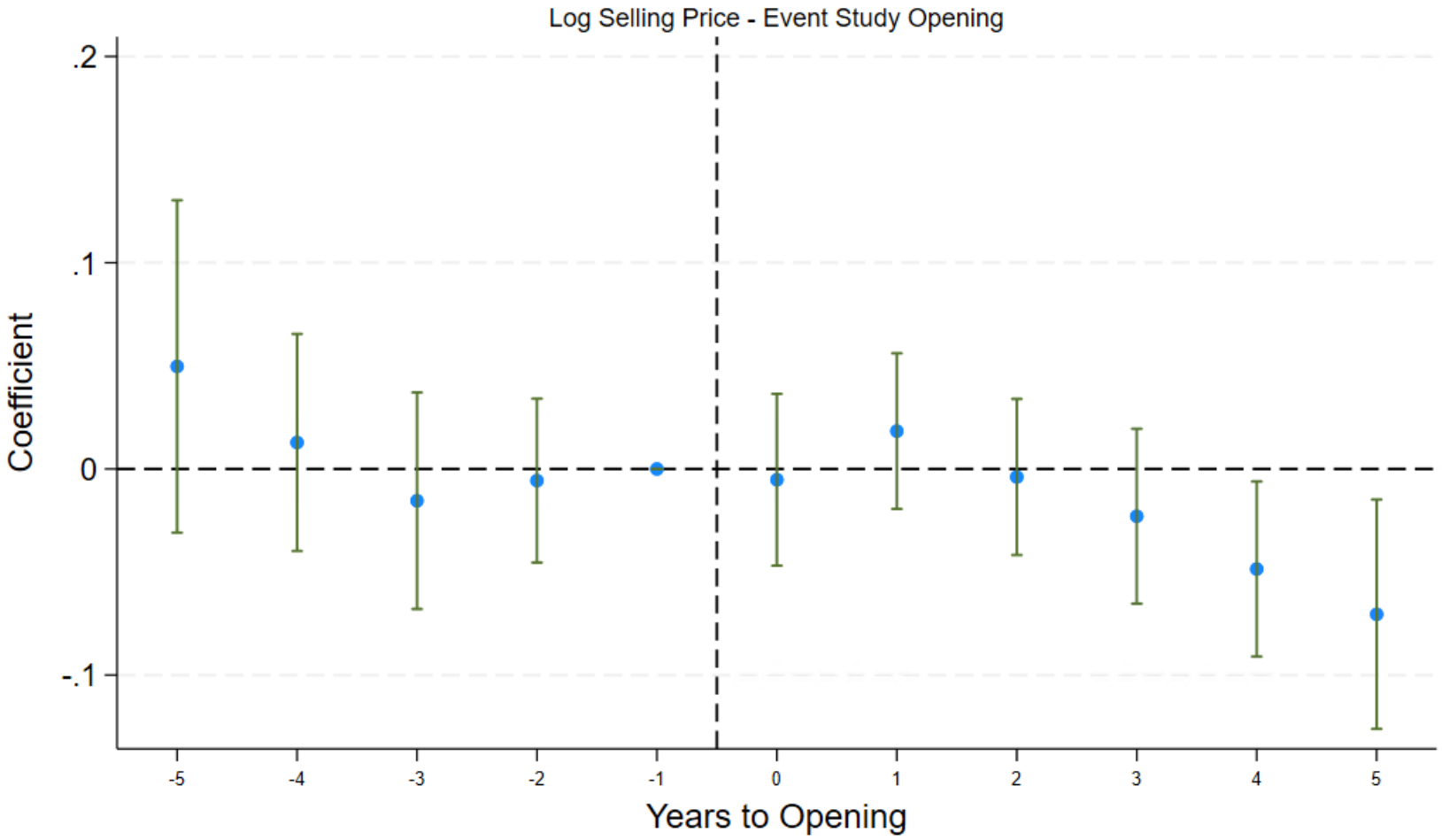


Figure 15: Closings Event Study: Log Selling Price

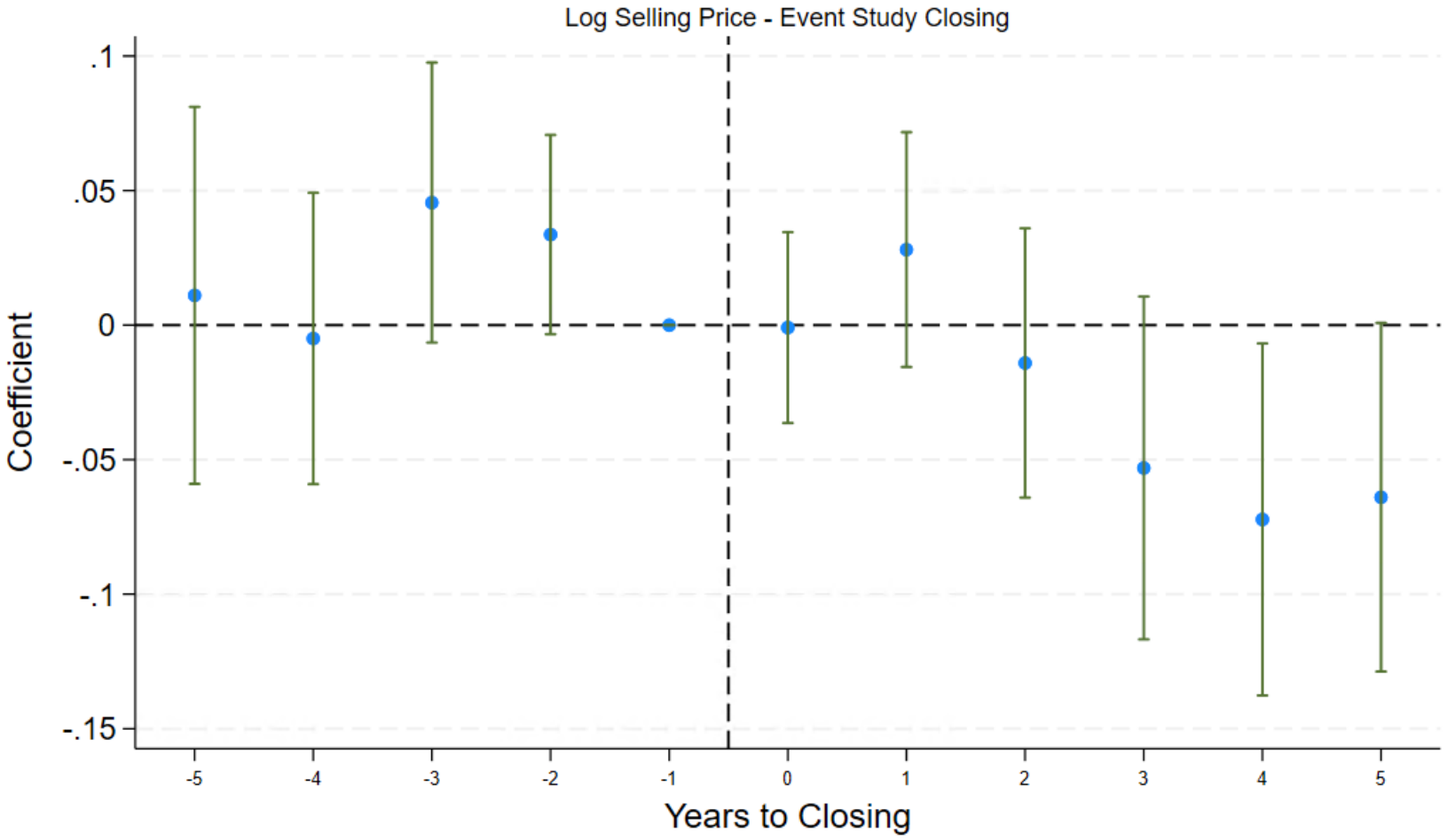


Figure 16: Openings Event Study: Share of Movers Out

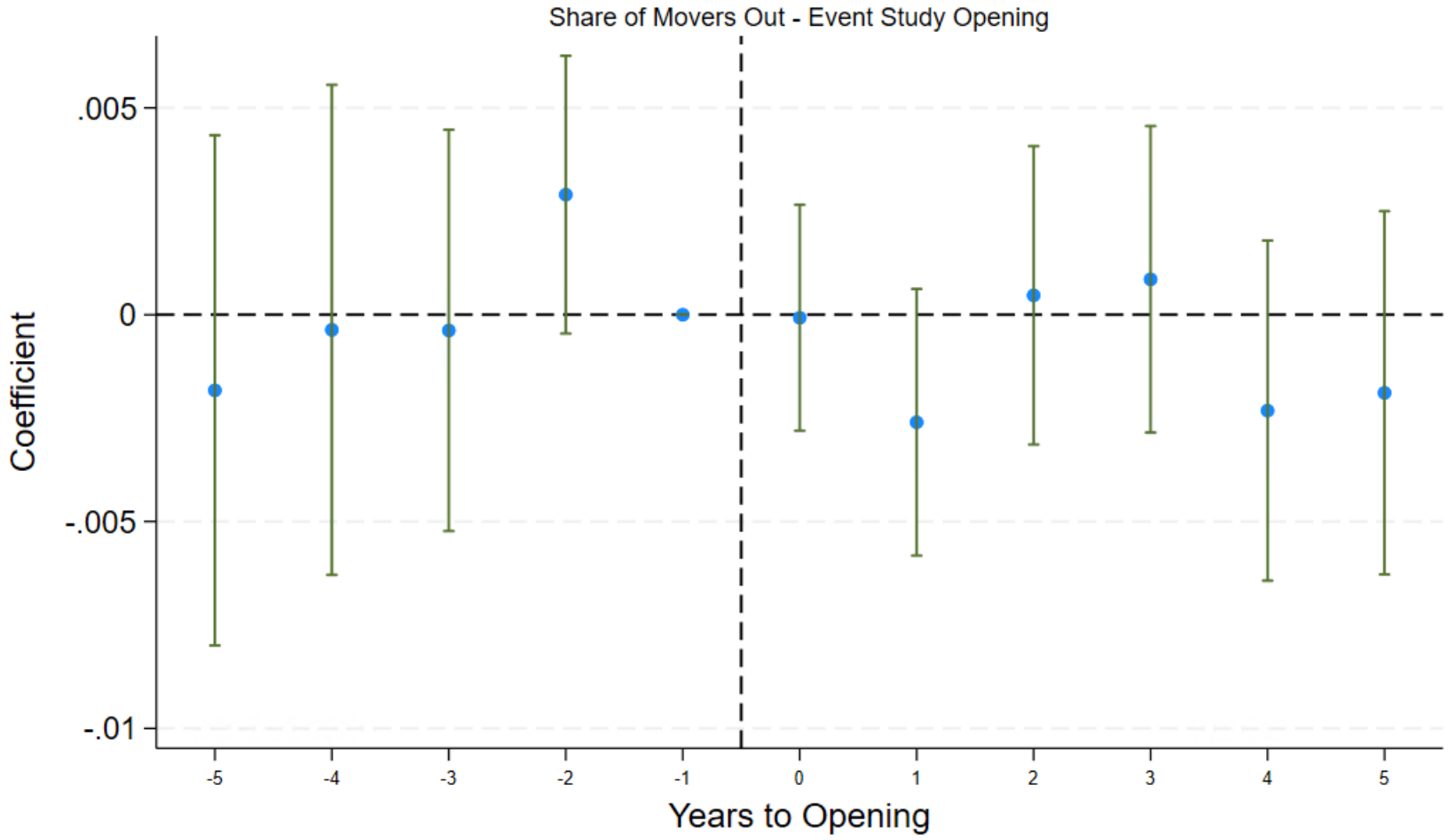


Figure 17: Openings Event Study: Share of Movers Out

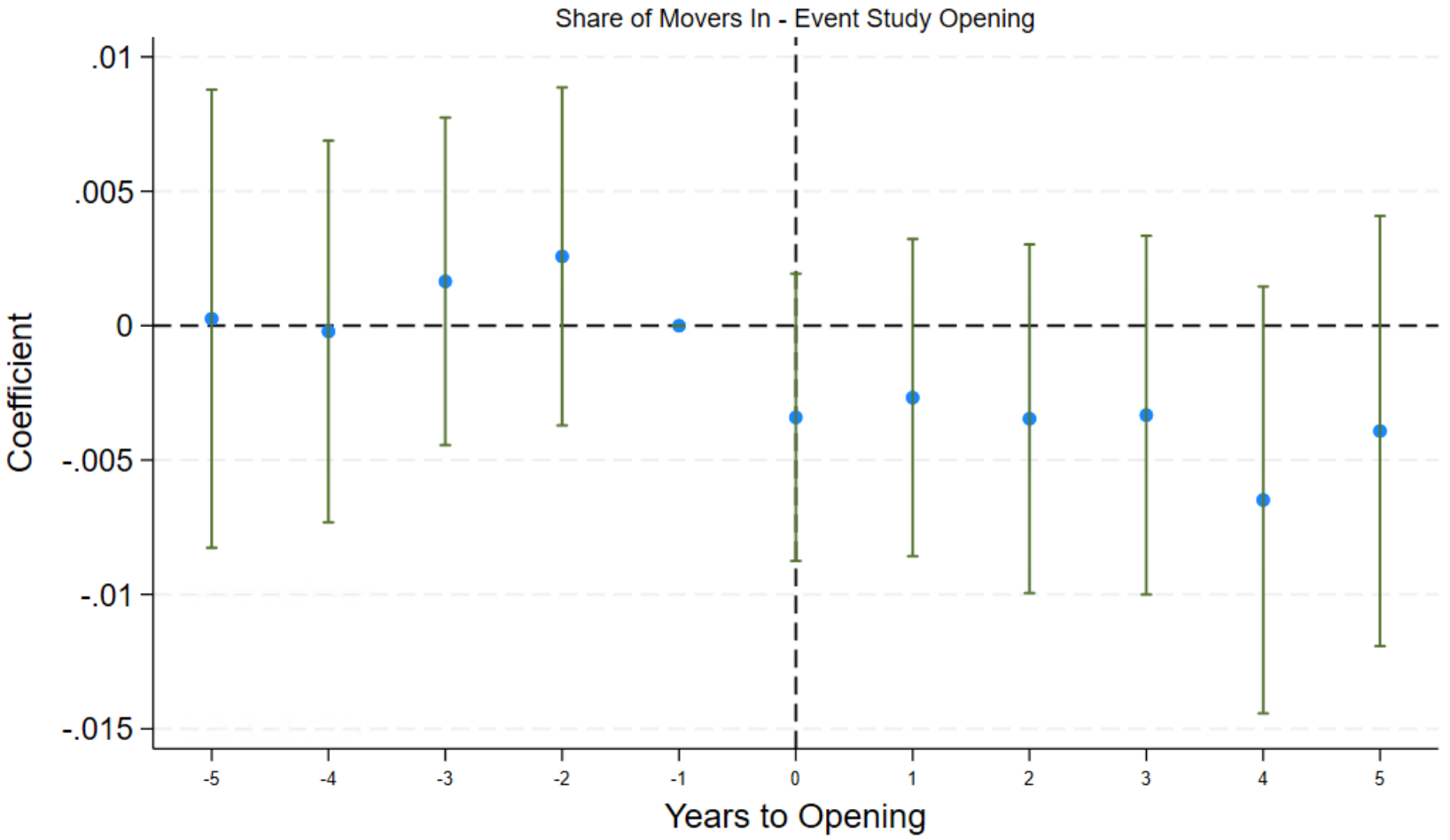


Figure 18: SO2 Emissions from Incinerators This picture shows the sum of emissions of SO2 (in tonnes) from incinerators between 1980 and 2018. For the incinerator-years associated with available Green Accounts (red area), I consider emissions as reported in these documents. For the other incinerator-years (blue area), I consider CORINAIR SO2, computed as the product between waste burnt and CORINAIR emissions factors. For the period prior to 1994, I discount 1994 waste by real GDP growth. For the period prior to 1990, I assume the 1990 emissions factors to be constant. In total, there are 1,435 incinerator-years between 1980 and 2018; 516 are associated with available Green Accounts.

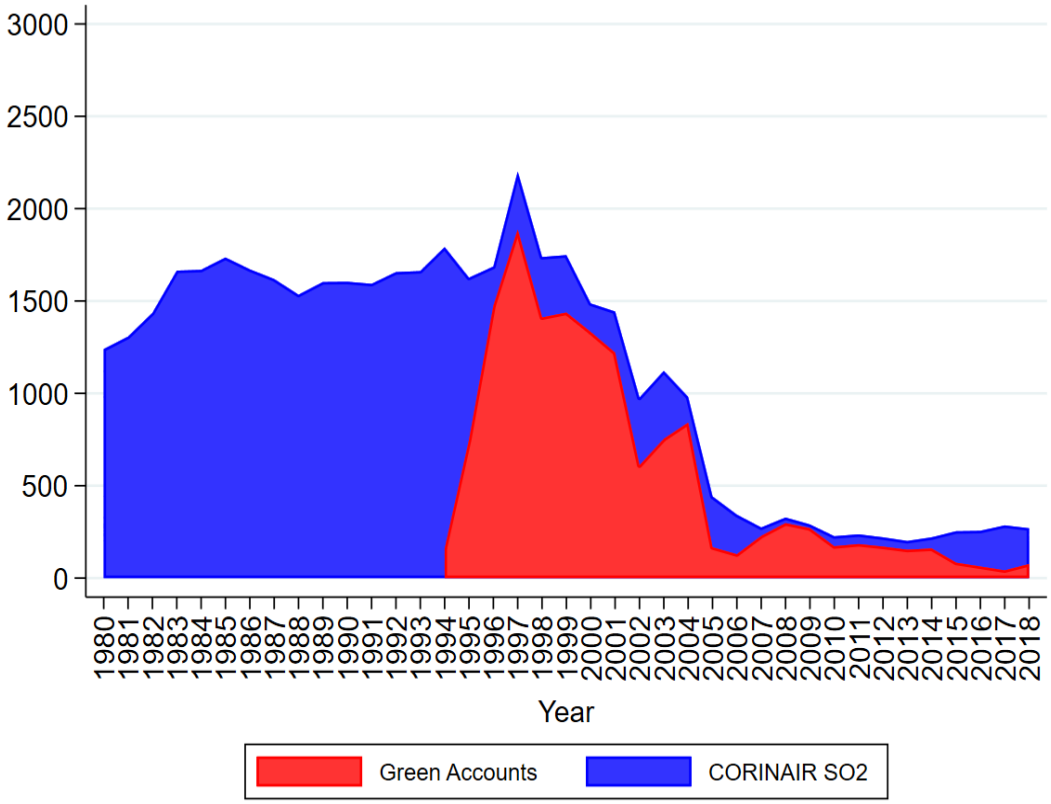
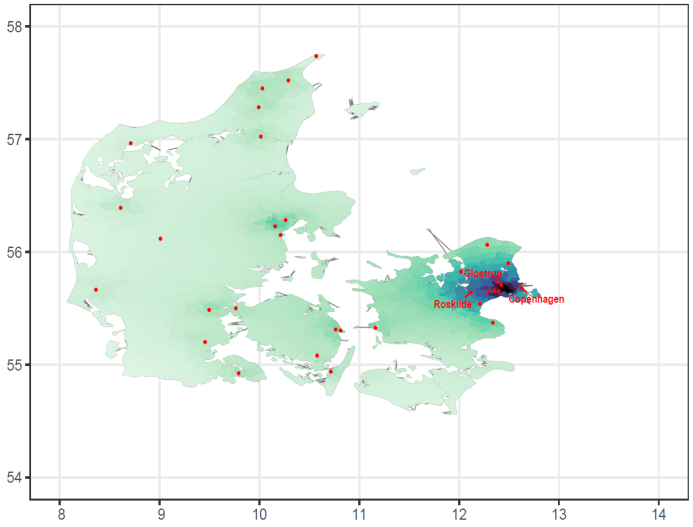
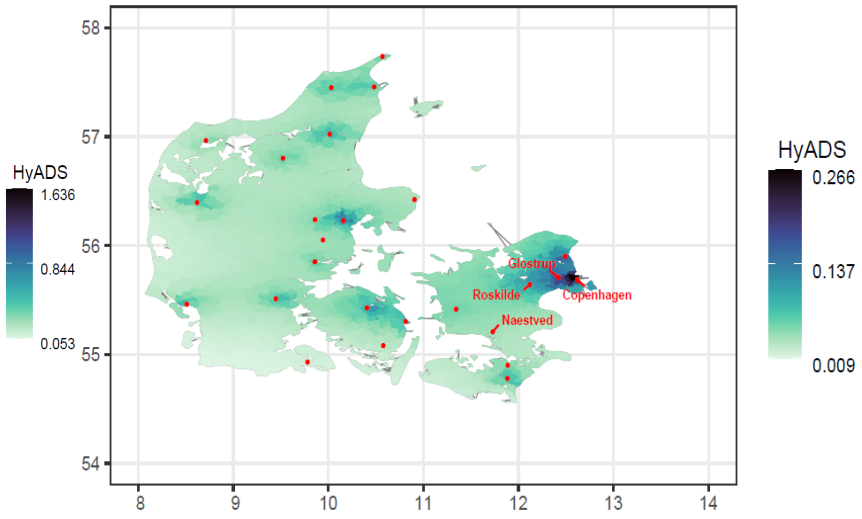


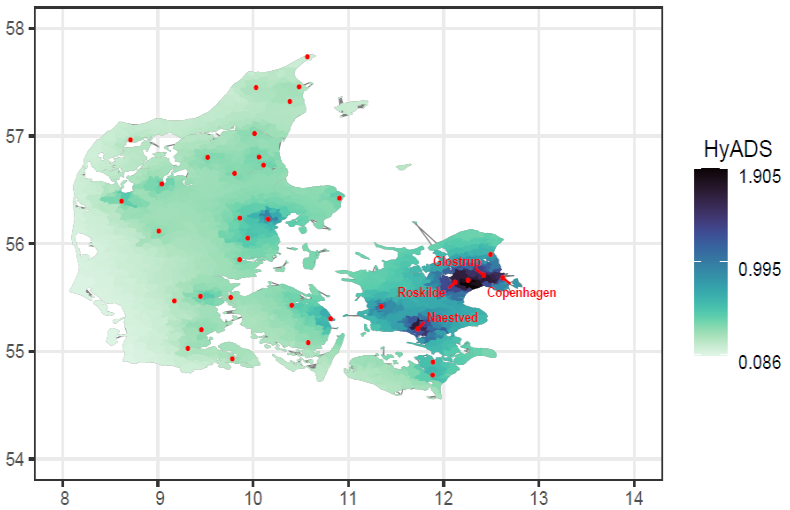
Figure 19: HyADS in 1980, 1997, and 2018. This figure shows the map of Denmark in 1980, 1997, and 2018, with each parish being colored depending on the value taken by HyADS and each red dot representing the location of an active incinerator. HyADS is computed by running HYSPLIT starting from emissions of SO2 from incinerators; every year, the release of 100 particles is simulated from each active plant, 4 times per day; the concentration of the particles from each plant is computed at the parish level and weighted by emissions of SO2. The latter are either drawn from the Green Accounts (available from 1996) or computed as the product between waste burnt and CORINAIR emissions factors; for the period before 1990, we assume that the 1990 emissions factors stayed constant; for the period before 1994, we discount the 1994 data on waste burnt by real GDP growth. we highlight the names of the 4 biggest plants (Glostrup, Copenhagen, Naestved, and Roskilde).



(a) HyADS in 1980.



(b) HyADS in 2018.



(c) HyADS in 1997.

Tables

Table 1: Cox Regressions Sample: Characteristics by Group

Group	Full Sample (1)	0-2km from Incinerator (2)	2-4km from Incinerator (3)	4-6km from Incinerator (4)	6-8km from Incinerator (5)	8-10km from Incinerator (6)	10+km from Incinerator (7)
Number of Person/Years	168,113,550	7,858,595	20,290,519	21,377,049	17,371,974	13,274,319	87,941,094
Panel A:				Outcomes			
Lung Cancer Diagnosis	7.834	8.385	8.438	8.084	7.961	7.180	7.659
Lung Cancer Mortality	3.847	4.310	4.292	4.162	4.046	3.414	3.652
Panel B:				Individual Characteristics			
Age	40.233	39.931	40.105	39.534	40.265	39.728	40.530
Income (USD)	36,783.233	35,736.193	36,904.850	37,264.711	38,907.252	40,309.311	35,779.871
Gross Wealth (USD)	97,781.455	77,635.123	80,167.493	80,111.574	106,254.909	124,455.017	102,240.965
Male	0.494	0.483	0.483	0.484	0.487	0.493	0.501
Married	0.411	0.389	0.365	0.352	0.385	0.420	0.441
Divorced	0.072	0.081	0.084	0.086	0.078	0.069	0.063
Widow	0.070	0.075	0.072	0.068	0.071	0.063	0.070
Higher Education	0.189	0.173	0.212	0.227	0.235	0.228	0.160
Panel C:				Municipality Characteristics			
Average Municipality Income (USD)	35,784.916	35,250.848	35,406.202	35,992.652	37,406.637	38,599.209	35,124.362
Average Municipality Wealth (USD)	95,295.211	90,496.634	80,821.720	77,648.195	100,404.740	119,091.817	98,751.836
Municipality Share of Immigrants	0.076	0.070	0.090	0.105	0.098	0.087	0.061
Municipality Share of Higher Education	0.199	0.184	0.210	0.232	0.246	0.242	0.175
N of Hospitals in the Municipality	0.990	0.912	1.415	1.399	1.162	0.795	0.795
Panel D:				Parish Characteristics			
Average Parish Income	36,056.690	34,928.228	35,961.300	36,147.811	37,596.050	39,411.542	35,346.903
Average Parish Wealth	97,920.029	77,250.428	78,131.114	77,784.984	103,556.200	128,304.773	103,527.658
Parish Share of Immigrants	0.075	0.082	0.094	0.105	0.093	0.082	0.058
Parish Share of Higher Education	0.199	0.185	0.219	0.238	0.245	0.243	0.171
Panel E:				Distance from a non-incinerator LPS			
0-4km from an LPS	0.299	0.401	0.565	0.503	0.415	0.305	0.154
4-10km from an LPS	0.238	0.133	0.157	0.321	0.389	0.449	0.185
10-20km from an LPS	0.247	0.216	0.090	0.093	0.115	0.160	0.363

In this table, we show the characteristics of the main sample. In Column 1, we show the full sample. In Columns 2-6, we restrict to the person/years who live 0-2km, 2-4km, 4-6km, 6-8km and 8-10km from the nearest active incinerator, respectively. Finally, in Column 7, we restrict to those who live at least 10km from the nearest active incinerator.

Table 2: Recap of the Incinerator Openings

Incinerator Name (District/Municipality) (1)	Waste Burnt in the Opening Year (2)	Opening Year (3)	Closing Year (4)	% of Treated (5)
Aars	152.956	1986		5.64
Hadsund	27.373	1984	2005	3.73
Hammel	106.549	1986		4.38
Jerslev Sjælland	11.314	1991	1995	1.25
Næstved	205.255	1983		5.72
Hobro	75.972	1984	2017	6.74
Nykøbing F	197.801	1983		11.32
Skanderborg	182.203	1984		8.96
Slagelse	281.158	1990		0.15
Vejen	288.073	1990	2010	6.10
Horsens	599.035	1992		4.93
Holstebro	1301.524	1993		2.16
Kolding	316.817	1982		2.19
Skive	18.849	1984	2000	4.83
Taastrup	336.285	1989	2005	5.53
Rønne	171.615	1991		2.26
Terndrup	20.25	1992	1998	0.62
Frederikshavn	242.388	1994		2.09
Sønderborg	561.183	1996		3.58
Odense	685.51	1996		10.59
Esbjerg	1312.095	2003		7.24

In this table, we make a recap of all the incinerators that are used to construct the sample of openings. In column 1, we report the incinerator's municipality. In column 2, we report the amount of waste burnt in the opening year. Notice that, for all the pre-1994 openings, this amount is imputed and it may be underestimated. In columns 3 and 4, we report the opening year and the closing year, if any. Finally, in column 5, we report the percentage of treated who are exposed to that plant.

Table 3: Recap of the Incinerator Closings: Pre 1998 (**In the Final Sample**)

Incinerator Name (District/Municipality) (1)	Closing Year (2)	Opening Year (3)	% of Treated (4)
Gerlev (Frederikssund)	1985	1968	4.52
Magleby (Stevns)	1985	1972	0.58
Rosenholm (Syddjurs)	1985	1968	1.52
Sindal (Hjorring)	1985	1970	0.30
Kærup (Varde)	1986	1970	0.21
Klemensker (Bornholm)	1987	1969	0.34
Poulsker (Bornholm)	1987	1969	0.56
Rudkobing (Langeland)	1987	1966	1.46
Taastrup	1989	1968	0.28
Solrød	1990	1978	5.25
Albertslund	1992	1967	9.40
Brøndby	1992	1971	14.22
Kolding	1993	1969	2.89
Korsør (Slagelse)	1993	1971	4.40
Videbæk (Ringkøbing-Skjern)	1993	1981	1.38
Graested (Gribskov)	1995	1950	1.51
Jerslev Sjælland (Kalundborg)	1995	1991	1.17
Nyborg	1996	1972	0.34

In this table, we make a recap of all the incinerators that are used to construct the sample of openings. In column 1, we report the incinerator's municipality. In columns 2 and 3, we report the closing year and the opening year. Finally, in column 4, we report the percentage of treated who are exposed to that plant.

Table 4: Main Distance-Based Cox Regressions

Outcome	Lung Cancer Diagnosis (1)	Lung Cancer Mortality (2)
0-2km Distance	0.0720***	0.0760***
	(0.013)	(0.014)
2-4km Distance	0.0427***	0.0390***
	(0.009)	(0.010)
4-6km Distance	0.0115	0.0085
	(0.010)	(0.010)
6-8km Distance	0.0219**	0.0198*
	(0.010)	(0.011)
8-10km Distance	0.0083	0.0085
	(0.012)	(0.012)
Observations	174,981,686	175,188,207
Avg Outcome	7.698	7.105
Avg Distance (Nearest Plant)	8160	8160
Birth Cohort Strata	Yes	Yes
Gender Strata	Yes	Yes
Individual Characteristics	Yes	Yes
Municipality Characteristics	Yes	Yes
Parish Characteristics	Yes	Yes

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Individual controls include: education level dummies, lagged income quintiles, lagged wealth quintiles, age range dummies, birth year dummies, gender, widow, marriage and divorced status. Municipality and parish level controls include: number of hospitals in the municipality, share of immigrants, share of people with higher education, log population, average income and average wealth. **Excess hazard ratios reported.**

Table 5: IV Sample: Baseline Specification

Outcome	Lung Cancer	Lung Cancer	N Yrs Exposed	Lung Cancer
Specification	Cox (Hazard Ratios)	Linear (OLS)	Linear (1st Stage)	Linear (2nd Stage)
	(1)	(2)	(3)	(4)
N Yrs Exposed to Active Plants	0.0022***	0.0661***		0.0702***
	(0.000)	(0.004)		(0.015)
Post Opening in the Initial Address			7.3082***	
			(0.013)	
Observations	174,981,572	174,969,903	175,186,611	174,969,903
R-squared		0.002	0.357	0.002
Lung Cancer Rate	7.698	7.698	7.698	7.698
Avg. Cumul. Exposure	4.023	4.023	4.023	4.023
Avg. Cumul. Exposure (Ever Exposed)	9.463	9.463	9.463	9.463
Share of Instrumented			0.0741	
Cragg-Donald F Statistic				8.863e+06
Kleibergen-Paap F Statistic				5.815e+06
Individual Characteristics	Yes	Yes	Yes	Yes
Municipality & Parish Characteristics	Yes	Yes	Yes	Yes
Municipality FEs	Yes	Yes	Yes	Yes
Initial Municipality FEs	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses. Individual controls include: distance from the nearest hospital, education level dummies, lagged income quintiles, lagged wealth quintiles, age fixed effects, gender, widow, marriage and divorced status. Municipality and parish level controls include: number of hospitals in the municipality, share of immigrants, share of people with higher education, log population, average income and average wealth.

Table 6: Openings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of Person/Years	14,016,394	7,065,766	6,950,628
Panel A:	Outcomes		
Lung Cancer Diagnosis	10.206	10.184	10.228
Lung Cancer Mortality	9.185	9.095	9.277
Panel B:	Individual Characteristics		
Age	45.758	45.831	45.684
Income	34,678.470	34,238.071	35,126.165
Gross Wealth	87,799.711	83,534.405	92,135.674
Male	0.484	0.485	0.484
Married	0.495	0.491	0.499
Divorced	0.079	0.082	0.076
Widow	0.090	0.090	0.089
Higher Education	0.160	0.155	0.164
Panel C:	Municipality Characteristics		
Average Municipality Income	33,902.349	33,165.759	34,651.140
Average Municipality Wealth	86,061.746	84,662.931	87,483.733
Municipality Share of Immigrants	0.067	0.059	0.075
Municipality Share of Higher Education	0.175	0.163	0.188
N of Hospitals in the Municipality	0.976	0.885	1.068
Panel D:	Parish Characteristics		
Average Parish Income	34,106.706	33,418.829	34,805.979
Average Parish Wealth	84,269.705	79,008.653	89,617.907
Parish Share of Immigrants	0.067	0.062	0.072
Parish Share of Higher Education	0.173	0.160	0.186
Panel E:	Distance from a non-incinerator LPS		
0-4km from an LPS	0.514	0.575	0.452
4-10km from an LPS	0.133	0.097	0.169
10-20km froman LPS	0.118	0.085	0.152

In this table, we split the first matching diff-in-diff sample in different categories and we show the corresponding characteristics. This is the sample in which we impose that the treated are exposed for the first time to an active plant in its opening year or after. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 7: Openings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean (1)	Control Group Mean (2)	Standardized Difference (3)	T-Test (4)
Panel A:		Outcomes		
Lung Cancer Diagnosis	7.58	8.34	-0.00	-0.76**
Lung Cancer Mortality	6.77	7.52	-0.00	-0.75**
Panel B:		Individual Characteristics		
Age	39.13	39.04	0.00	0.09***
Income	28,370.98	29,007.16	-0.03	-636.18***
Gross Wealth	56,515.63	58,615.34	-0.01	-2,099.71***
Male	0.49	0.49	0.00	0.00
Married	0.45	0.47	-0.03	-0.01***
Divorced	0.06	0.05	0.04	0.01***
Widow	0.08	0.08	0.01	0.00***
Higher Education	0.10	0.11	-0.06	-0.02***
Panel C:		Municipality Characteristics		
Average Municipality Income	28,005.25	29,330.61	-0.43	-1,325.36***
Average Municipality Wealth	61,472.03	62,653.32	-0.07	-1,181.29***
Municipality Share of Immigrants	0.04	0.06	-0.52	-0.01***
Municipality Share of Higher Education	0.12	0.14	-0.36	-0.01***
N of Hospitals in the Municipality	0.93	1.09	-0.25	-0.16***
Panel D:		Parish Characteristics		
Average Parish Income	28,222.60	29,457.56	-0.31	-1,234.96***
Average Parish Wealth	56,346.42	63,389.96	-0.26	-7,043.53***
Parish Share of Immigrants	0.05	0.05	-0.24	-0.01***
Parish Share of Higher Education	0.11	0.13	-0.43	-0.02***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.53	0.43	0.21	0.10***
4-10km from an LPS	0.07	0.12	-0.15	-0.04***
10-20km from an LPS	0.06	0.12	-0.22	-0.06***
Observations	1,820,209	1,804,772		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the openings Diff-in-Diff sample. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 8: Openings Diff-in-Diff Results

Specification	Full Sample	Age in the Opening Year: 20-39	Age in the Opening Year: 40+	Opening Year (Age: 40+): Pre 1991	Opening Year (Age: 40+): Post 1991	Wealth (Age: 40+): 1st Quintile	Wealth (Age: 40+): 5th Quintile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated × 0-4 Yrs After	0.7495	0.4618*	0.8657	2.9121	-0.9727	3.3649	-0.8570
	(0.732)	(0.243)	(1.152)	(1.788)	(1.516)	(2.994)	(2.117)
Treated × 5-9 Yrs After	2.1037***	-0.0830	3.3806***	2.2826	4.3868***	7.7923**	1.4413
	(0.764)	(0.266)	(1.242)	(1.888)	(1.664)	(3.278)	(2.254)
Treated × 10-14 Yrs After	2.7140***	0.8380**	3.7461***	4.4179**	3.1344*	7.9921**	0.0853
	(0.809)	(0.407)	(1.357)	(2.081)	(1.800)	(3.641)	(2.445)
Treated × 15-19 Yrs After	2.2697***	0.5086	3.1697**	2.7697	3.5419*	16.0813***	-4.1498
	(0.880)	(0.542)	(1.542)	(2.288)	(2.104)	(4.185)	(2.892)
Treated × 20+ Yrs After	3.2878***	-0.0194	5.9374***	5.1566**	7.4507***	14.2353***	1.9024
	(0.836)	(0.629)	(1.510)	(2.054)	(2.358)	(4.289)	(2.769)
Observations	10,211,522	4,491,089	5,720,433	2,430,175	3,290,258	1,106,966	1,149,042
R-squared	0.002	0.001	0.001	0.001	0.001	0.002	0.001
Avg Outcome	13.31	4.041	20.59	21.57	19.86	27.99	14.66
Avg Age	53.35	40.39	63.54	65.70	61.93	62.19	64.17
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality and Parish Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the person level, in parentheses. Individual controls include: education level dummies, lagged income quintiles, lagged wealth quintiles, age range dummies, birth year dummies, gender, widow, marriage and divorced status. Municipality and parish level controls include: number of hospitals in the municipality, share of immigrants, share of people with higher education, log population, average income and average wealth. Group fixed effects are the interaction between the opening year FEs (either actual or fictitious) and the *at risk* group status.

Table 9: Closings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of Person/Years	8,090,673	3,969,589	4,121,084
Panel A:		Outcomes	
Lung Cancer Diagnosis	10.753	11.595	9.942
Lung Cancer Mortality	9.486	10.378	8.627
Panel B:		Individual Characteristics	
Age	47.759	47.446	48.061
Income	36,713.669	37,393.407	36,059.537
Gross Wealth	95,547.588	98,552.788	92,655.596
Male	0.485	0.484	0.486
Married	0.539	0.548	0.531
Divorced	0.090	0.088	0.092
Widow	0.086	0.083	0.089
Higher Education	0.173	0.162	0.183
Panel C:		Municipality Characteristics	
Average Municipality Income	35,296.166	36,299.150	34,330.96
Average Municipality Wealth	87,115.503	85,663.775	88,512.543
Municipality Share of Immigrants	0.075	0.085	0.066
Municipality Share of Higher Education	0.176	0.173	0.180
N of Hospitals in the Municipality	0.782	0.585	0.971
Panel D:		Parish Characteristics	
Average Parish Income	35,664.252	36,713.533	34,654.499
Average Parish Wealth	84,771.121	87,588.361	82,060.008
Parish Share of Immigrants	0.075	0.080	0.070
Parish Share of Higher Education	0.178	0.172	0.183
Panel E:		Distance from a non-incinerator LPS	
0-4km from an LPS	0.285	0.138	0.426
4-10km from an LPS	0.304	0.454	0.159
10-20km from an LPS	0.188	0.290	0.090

In this table, we split the Diff-in-Diff distance-based closings sample in different categories and we show the corresponding characteristics. This is the sample in which we impose that the treated are exposed for the first time to an active plant in its opening year. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 10: Closings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean	Control Group Mean	Standardized Difference	T-Test
	(1)	(2)	(3)	(4)
Panel A:		Outcomes		
Lung Cancer Diagnosis	8.48	5.68	0.01	2.80***
Lung Cancer Mortality	7.64	4.76	0.01	2.89***
Panel B:		Individual Characteristics		
Age	4048	40.83	-0.02	-0.35***
Income	33,100.31	31,503.05	0.08	1,597.26***
Gross Wealth	64,690.99	65,354.65	-0.00	-663.66***
Male	0.49	0.49	-0.00	-0.00
Married	0.54	0.53	0.02	0.01***
Divorced	0.07	0.07	0.01	0.00***
Widow	0.06	0.07	-0.02	-0.00***
Higher Education	0.13	0.14	-0.03	-0.01***
Panel C:		Municipality Characteristics		
Average Municipality Income	32,016.25	29,226.49	0.70	2,789.76***
Average Municipality Wealth	60,183.90	65,887.81	-0.28	-5,703.91***
Municipality Share of Immigrants	0.06	0.05	0.63	0.01***
Municipality Share of Higher Education	0.13	0.13	-0.16	-0.01***
N of Hospitals in the Municipality	0.58	0.94	-0.54	-0.36***
Panel D:		Parish Characteristics		
Average Parish Income	32,413.99	29,744.04	0.57	2,669.95***
Average Parish Wealth	60,485.29	63,016.82	-0.09	-2,531.53***
Parish Share of Immigrants	0.06	0.05	0.43	0.01***
Parish Share of Higher Education	0.13	0.13	-0.18	-0.01***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.08	0.35	-0.70	-0.27***
4-10km from an LPS	0.44	0.19	0.55	0.25***
10-20km from an LPS	0.34	0.10	0.61	0.24***
Observations	1,203,708	1,226,009		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the closings Diff-in-Diff sample. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 11: Closings Diff-in-Diff Results

Specification	Full Sample	Age in the Closing Year: 20-39	40+	Yrs of Exposure (Age: 40+): 1st Quintile	5th Quintile	Wealth (Age: 40+): 1st Quintile	5th Quintile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated × 0-4 Yrs After	-1.6085*	0.2989	-2.2647*	2.1225	-0.9317	4.3786	-1.3028
	(0.890)	(0.262)	(1.253)	(2.644)	(2.569)	(3.211)	(2.462)
Treated × 5-9 Yrs After	-0.3107	0.1524	-0.2320	2.7481	0.5914	1.3679	-3.2516
	(0.936)	(0.416)	(1.337)	(2.923)	(2.756)	(3.545)	(2.561)
Treated × 10-14 Yrs After	-1.4884	0.2534	-1.9240	5.2539	2.5174	-0.7887	-0.1785
	(1.011)	(0.603)	(1.470)	(3.268)	(3.104)	(3.932)	(2.781)
Treated × 15-19 Yrs After	-2.3272**	-0.0780	-2.9012*	6.0375	-6.7073*	-0.8875	-2.4456
	(1.156)	(0.857)	(1.721)	(3.766)	(3.670)	(4.725)	(3.258)
Treated × 20+ Yrs After	-5.3084***	-0.9177	-7.2706***	-8.6384**	-14.2823***	-11.8072**	-3.3583
	(1.129)	(0.933)	(1.760)	(3.996)	(3.897)	(4.957)	(3.295)
Observations	6,344,924	2,113,901	4,231,023	714,883	954,313	847,763	846,519
R-squared	0.001	0.001	0.001	0.001	0.001	0.002	0.001
Avg Outcome	13.45	3.392	18.47	16.39	17.80	25.10	13.60
Avg Age	54.09	39.47	61.39	60.61	59.82	60.42	61.82
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality and Parish Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Standard errors, clustered at the person level, in parentheses. Individual controls include: education level dummies, lagged income quintiles, lagged wealth quintiles, age range dummies, birth year dummies, gender, widow, marriage and divorced status. Municipality and parish level controls include: number of hospitals in the municipality, share of immigrants, share of people with higher education, log population, average income and average wealth. Group fixed effects are the interaction between the closing year FEs (either actual or fictitious) and the *at risk* group status.

Table 12: Smokers Openings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of Person/Years	60,588	30,294	30,294
Panel A:	Outcomes		
Smoker	0.239	0.249	0.229
Panel B:	Mother's Characteristics		
Mother's Age	28.564	28.564	28.564
Mother's Income	263,496.369	262,390.905	264,601.833
Mother's Gross Wealth	471,401.777	457,224.183	485,579.371
Mother Higher Education	0.310	0.311	0.309
Panel C:	Municipality Characteristics		
Average Municipality Income (USD)	36,013.750	35,952.608	36,074.892
Average Municipality Wealth (USD)	94,398.920	92,618.940	96,178.901
Municipality Share of Immigrants	0.077	0.077	0.077
Municipality Share of Higher Education	0.200	0.200	0.200
N of Hospitals in the Municipality	0.963	0.958	0.968
Panel D:	Parish Characteristics		
Average Parish Income (USD)	36,141.141	35,894.658	36,387.623
Average Parish Wealth (USD)	95,086.619	89,093.413	101,079.824
Parish Share of Immigrants	0.076	0.081	0.072
Parish Share of Higher Education	0.197	0.192	0.202
Panel E:	Distance from a non-incinerator LPS		
0-4km from an LPS	0.412	0.416	0.408
4-10km from an LPS	0.155	0.149	0.161
10-20km from an LPS	0.107	0.107	0.107

In this table, we split the birth-level diff-in-diff openings sample of smokers in different categories and we show the corresponding characteristics. This is the sample in which we impose that the treated are exposed for the first time to an active plant in its opening year or after. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 13: Smoker Openings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean	Control Group Mean	Standardized Difference	T-Test
	(1)	(2)	(3)	(4)
Panel A:		Outcomes		
Smoker	0.42	0.35	0.14	0.07***
Panel B:		Mother's Characteristics		
Mother's Age	26.61	26.61	-0.00	-0.00
Mother's Income	33,442.61	35,621.64	-0.16	-2,179.04***
Mother's Gross Wealth	31,608.11	40,122.44	-0.07	-8,514.33***
Mother Higher Education	0.14	0.19	-0.12	-0.04***
Panel C:		Municipality Characteristics		
Average Municipality Income (USD)	31,582.62	31,423.46	0.08	159.16***
Average Municipality Wealth (USD)	54,909.76	56,967.86	-0.18	-2,058.10***
Municipality Share of Immigrants	0.06	0.06	0.27	0.00***
Municipality Share of Higher Education	0.15	0.15	0.16	0.01***
N of Hospitals in the Municipality	0.98	0.94	0.13	0.04***
Panel D:		Parish Characteristics		
Average Parish Income (USD)	30,854.07	31,696.93	-0.26	-842.86***
Average Parish Wealth (USD)	45,470.32	57,389.24	-0.56	-11,918.91***
Parish Share of Immigrants	0.07	0.06	0.46	0.02***
Parish Share of Higher Education	0.13	0.15	-0.44	-0.02***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.46	0.39	0.13	0.06**
4-10km from an LPS	0.15	0.20	-0.13	-0.05**
10-20km from an LPS	0.14	0.19	-0.12	-0.05**
Observations	3,505	3,505		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the birth-level diff-in-diff openings sample of smokers. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 14: Smokers Closings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of Person/Years	68,030	34,015	34,015
Panel A:	Outcomes		
Smoker	0.240	0.251	0.230
Panel B:	Mother's Characteristics		
Mother's Age	28.481	28.481	28.481
Mother's Income	40,950.100	40,709.382	41,190.818
Mother's Gross Wealth	70,657.528	68,506.530	72,808.525
Mother Higher Education	0.294	0.273	0.315
Panel C:	Municipality Characteristics		
Average Municipality Income (USD)	37,331.779	37,485.880	37,177.677
Average Municipality Wealth (USD)	95,829.658	95,172.405	96,486.91
Municipality Share of Immigrants	0.084	0.090	0.079
Municipality Share of Higher Education	0.195	0.192	0.199
N of Hospitals in the Municipality	0.802	0.754	0.851
Panel D:	Parish Characteristics		
Average Parish Income (USD)	37,596.307	37,721.110	37,471.504
Average Parish Wealth (USD)	93,226.814	93,212.471	93,241.157
Parish Share of Immigrants	0.085	0.089	0.081
Parish Share of Higher Education	0.197	0.193	0.202
Panel E:	Distance from a non-incinerator LPS		
0-4km from an LPS	0.407	0.333	0.481
4-10km from an LPS	0.283	0.306	0.260
10-20km from an LPS	0.230	0.293	0.166

In this table, we split the birth-level diff-in-diff closings sample of smokers in different categories and we show the corresponding characteristics. This is the sample in which we impose that the treated are exposed for the first time to an active plant in its opening year or after. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 15: Smoker Closings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean	Control Group Mean	Standardized Difference	T-Test
	(1)	(2)	(3)	(4)
Panel A:		Outcomes		
Smoker	0.31	0.30	0.02	0.01
Panel B:		Mother's Characteristics		
Mother's Age	28.08	28.08	0.00	0.00
Mother's Income	39,172.92	38,817.20	0.03	355.71*
Mother's Gross Wealth	59,427.28	58,315.69	0.01	1,111.59
Mother Higher Education	0.20	0.23	-0.07	-0.03***
Panel C:		Municipality Characteristics		
Average Municipality Income (USD)	34,928.33	34,762.16	0.05	166.17***
Average Municipality Wealth (USD)	79,189.54	76,328.16	0.10	2,861.38***
Municipality Share of Immigrants	0.06	0.06	-0.01	-0.00
Municipality Share of Higher Education	0.15	0.16	-0.32	-0.01***
N of Hospitals in the Municipality	0.79	0.82	-0.05	-0.03***
Panel D:		Parish Characteristics		
Average Parish Income (USD)	35,437.68	35,161.22	0.07	276.46***
Average Parish Wealth (USD)	76,465.95	73,243.65	0.09	3,222.30***
Parish Share of Immigrants	0.06	0.06	0.04	0.00**
Parish Share of Higher Education	0.16	0.16	-0.16	-0.01***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.47	0.52	-0.10	-0.05***
4-10km from an LPS	0.24	0.25	-0.04	-0.02***
10-20km from an LPS	0.25	0.13	0.29	0.11***
Observations	10,687	10,686		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the birth-level diff-in-diff closings sample of smokers. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 16: Housing Openings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of House/Years	61,762	30,881	30,881
Panel A:		Outcomes	
Inflation-Adjusted Selling Price (thousands of DKK)	948.516	892.366	1,004.666
Selling Price (thousands of DKK)	1,140.000	1,072.708	1,207.293
Log Selling Price	6.891	6.858	6.924
Panel B:		Building Characteristics	
Number of Floors	1.579	1.604	1.555
Number of Rooms	3.827	3.778	3.876
Build Year	1,950.436	1,951.329	1,949.544
Building Age	52.127	51.146	53.107
Panel C:		Municipality Characteristics	
Average Municipality Income (USD)	35,583.312	35,185.487	35,981.137
Average Municipality Wealth (USD)	86,851.935	83,027.379	90,676.490
Municipality Share of Immigrants	0.073	0.074	0.073
Municipality Share of Higher Education	0.189	0.187	0.191
N of Hospitals in the Municipality	0.998	0.999	0.997
Panle D:		Parish Characteristics	
Average Parish Income (USD)	35,030.795	35,031.844	35,029.746
Average Parish Wealth (USD)	76,855.531	72,394.178	81,316.884
Parish Share of Immigrants	0.075	0.076	0.075
Parish Share of Higher Education	0.184	0.183	0.184
Panel E:		Distance from a non-incinerator LPS	
0-4km from an LPS	0.740	0.746	0.735
4-10km from an LPS	0.022	0.018	0.025
10-20km from an LPS	0.000	0.000	0.000

In this table, we split the transaction-level diff-in-diff sample of house prices in different categories and we show the corresponding characteristics. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 17: Housing Openings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean	Control Group Mean	Standardized Difference	T-Test
	(1)	(2)	(3)	(4)
Panel A:		Outcomes		
Inflation-Adjusted Selling Price (thousands of DKK)	552.80	631.90	-0.27	-79.10***
Selling Price (thousands of DKK)	786.69	894.27	-0.28	-107.58***
Log Selling Price	6.58	6.67	-0.19	-0.09***
Panel B:		Building Characteristics		
Number of Floors	1.72	1.69	0.02	0.03
Number of Rooms	3.68	3.81	-0.09	-0.13***
Build Year	1,946.94	1,948.54	-0.05	-1.60**
Building Age	48.78	47.46	0.05	1.32**
Panel C:		Municipality Characteristics		
Average Municipality Income (USD)	32,704.84	33,490.41	-0.36	-785.56***
Average Municipality Wealth (USD)	58,297.89	60,731.63	-0.20	-2,433.74***
Municipality Share of Immigrants	0.06	0.06	-0.01	-0.00
Municipality Share of Higher Education	0.16	0.16	-0.18	-0.01***
N of Hospitals in the Municipality	1.00	1.02	-0.08	-0.03***
Panle D:		Parish Characteristics		
Average Parish Income	32,000.12	32,567.84	-0.19	-567.72***
Average Parish Wealth	47,276.69	53,703.61	-0.35	-6,426.92***
Parish Share of Immigrants	0.07	0.06	0.31	0.01***
Parish Share of Higher Education	0.14	0.15	-0.22	-0.01***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.93	0.92	0.03	0.01
4-10km from an LPS	0.01	0.01	0.06	0.01***
10-20km from an LPS	0.00	0.00	0.00	0.00
Observations	6,150	6,150		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the transaction-level diff-in-diff sample of house prices. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 18: Housing Closings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of House/Years	72,630	36,315	36,315
Panel A:		Outcomes	
Inflation-Adjusted Selling Price (thousands of DKK)	1,191.310	1,082.240	1,300.380
Selling Price (thousands of DKK)	1,417.059	1,285.882	1,548.236
Log Selling Price	7.097	7.016	7.179
Panel B:		Building Characteristics	
Number of Floors	1.530	1.604	1.457
Number of Rooms	4.025	3.966	4.084
Build Year	1,958.098	1,958.732	1,957.464
Building Age	44.676	43.622	45.731
Panel C:		Municipality Characteristics	
Average Municipality Income (USD)	38,118.236	37,304.884	38,931.588
Average Municipality Wealth (USD)	104,526.394	100,010.155	109,042.634
Municipality Share of Immigrants	0.077	0.076	0.079
Municipality Share of Higher Education	0.190	0.167	0.214
N of Hospitals in the Municipality	0.730	0.710	0.751
Panle D:		Parish Characteristics	
Average Parish Income (USD)	38,459.319	38,066.890	38,851.748
Average Parish Wealth (USD)	99,365.968	97,687.345	101,044,591
Parish Share of Immigrants	0.077	0.079	0.076
Parish Share of Higher Education	0.193	0.180	0.206
Panel E:		Distance from a non-incinerator LPS	
0-4km from an LPS	0.420	0.411	0.429
4-10km from an LPS	0.256	0.209	0.304
10-20km from an LPS	0.272	0.348	0.195

In this table, we split the transaction-level diff-in-diff sample of house prices in different categories and we show the corresponding characteristics. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 19: Housing Closings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean	Control Group Mean	Standardized Difference	T-Test
	(1)	(2)	(3)	(4)
Panel A:		Outcomes		
Inflation-Adjusted Selling Price (thousands of DKK)	834.31	1,062.84	-0.38	-228.53***
Selling Price (thousands of DKK)	1,074.30	1,365.87	-0.42	-291.57***
Log Selling Price	6.86	7.06	-0.36	-0.20***
Panel B:		Building Characteristics		
Number of Floors	1.77	1.46	0.28	0.31***
Number of Rooms	3.79	4.03	-0.16	-0.24***
Build Year	1,957.84	1,955.40	0.06	2.44***
Building Age	40.64	44.31	-0.13	-3.67***
Panel C:		Municipality Characteristics		
Average Municipality Income (USD)	36,467.47	38,164.99	-0.36	-1,697.52***
Average Municipality Wealth (USD)	85,492.24	95,821.55	-0.25	-10,329.31***
Municipality Share of Immigrants	0.07	0.08	-0.13	-0.00***
Municipality Share of Higher Education	0.15	0.20	-1.02	-0.06***
N of Hospitals in the Municipality	0.59	0.77	-0.38	-0.18***
Panle D:		Parish Characteristics		
Average Parish Income (USD)	37,041.70	38,213.97	-0.23	-1,172.27***
Average Parish Wealth (USD)	81,210.16	90,455.72	-0.20	-9,245.57***
Parish Share of Immigrants	0.08	0.07	0.14	0.00***
Parish Share of Higher Education	0.16	0.20	-0.61	-0.04***
Panel E:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.56	0.41	0.32	0.16***
4-10km from an LPS	0.03	0.31	-0.78	-0.27***
10-20km from an LPS	0.40	0.20	0.45	0.20***
Observations	6,150	6,150		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the transaction-level diff-in-diff sample of house prices. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 20: Movers Openings Diff-in-Diff Sample: Characteristics by Group

Group	Full Sample (1)	Treatment Group (2)	Control Group (3)
Number of Parish/Years	6,240	3,120	3,120
Panel A:		Outcomes	
Share of Movers In	0.089	0.092	0.087
Share of Movers In, Bottom Income Quintile	0.109	0.114	0.105
Share of Movers In, Top Income Quintile	0.087	0.088	0.087
Share of Movers Out	0.091	0.094	0.088
Share of Movers Out, Bottom Income Quintile	0.105	0.109	0.102
Share of Movers Out, Top Income Quintile	0.084	0.087	0.082
Panel B:		Parish Characteristics	
Parish Population	5,796.349	5,674.146	5,918.552
Average Parish Income (USD)	34,977.393	34,510.409	35,444.377
Average Parish Wealth (USD)	90,144.468	88,003.135	92,285.801
Parish Share of Higher Education	0.176	0.169	0.183
Parish Share of Immigrants	0.068	0.067	0.070
Panel C:		Distance from a non-incinerator LPS	
0-4km from an LPS	0.498	0.533	0.463
4-10km from an LPS	0.161	0.121	0.202
10-20km from an LPS	0.165	0.163	0.167

In this table, we split the parish-level diff-in-diff sample of movers in different categories and we show the corresponding characteristics. In Column 1, we show the full sample. In Columns 2 and 3, we split it in two categories: treatment and control group.

Table 21: Movers Openings Diff-in-Diff Sample: Covariate Balance Test

	Treatment Group Mean (1)	Control Group Mean (2)	Standardized Difference (3)	T-Test (4)
Panel A:		Outcomes		
Share of Movers In	0.09	0.07	0.35	0.02***
Share of Movers In, Bottom Income Quintile	0.10	0.08	0.52	0.02***
Share of Movers In, Top Income Quintile	0.10	0.09	0.26	0.01***
Share of Movers Out	0.10	0.08	048	0.02***
Share of Movers Out, Bottom Income Quintile	0.11	0.09	0.44	0.02***
Share of Movers Out, Top Income Quintile	0.10	0.09	0.39	0.02***
Panel B:		Parish Characteristics		
Parish Population	5,562.87	5,488.19	0.02	74.68
Average Parish Income (USD)	29,327.72	30,141.03	-0.16	-813.31***
Average Parish Wealth (USD)	63,766.00	64,048.29	-0.01	-282.29
Parish Share of Higher Education	0.12	0.12	0.04	0.00
Parish Share of Immigrants	0.05	0.05	0.01	0.00
Panel C:		Distance from a non-incinerator LPS		
0-4km from an LPS	0.46	0.39	0.13	0.06**
4-10km from an LPS	0.15	0.20	-0.13	-0.05**
10-20km from an LPS	0.14	0.19	-0.12	-0.05**
Observations	816	816		

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. In this table, we show the covariate balance test performed on the parish-level diff-in-diff sample of movers. In Column 3, we report the standardized difference, while in Column 4 we report the t-test.

Table 22: Comparing the average cumulative effect of one year of incinerator activity to other causes of death

	Point Estimate [95% Confidence Interval]
Panel A	Comparison to Deryugina (Short Run, Probability of Death).
Incinerators Cumul. Prob. of Cancer (Over 50)	5.22 [1.15, 9.41]
Prob. of Death, $1\mu g/m^3$ PM2.5 Increase (Over 65)	2.51 [2.08, 2.95]
Panel B	Comparison to Deryugina (Long Run, Life Expectancy)
Incinerators Expected Life Loss in the Opening Year (50 Year Olds)	0.237 [0.048, 0.456]
Expected Life Improvement at Birth from a Permanent $1\mu g/m^3$ PM2.5 Decrease	0.47-0.57
Panel C	Any Age: Car Accident Deaths
Incinerators Cumul. Prob. of Cancer	1.96 [0.43, 3.53]
Prob. of Car Accident Death	0.28
Panel D	Working Age: Workplace Deaths
Incinerators Cumul. Prob. of Cancer (18-65 Year Olds)	2.36 [0.49, 4.30]
Prob. of Workplace Death	0.14
Panel E	Economic Comparisons (Yearly Costs in USD)
Total Incinerators VSL	724M [159M, 1,306M]
Total Operating Costs	226M
Total Incineration Taxes	210M
% of VSL covered by Taxes	29.01 [16.08, 134.61]

In this table, we compare probability of cancer from our simulations to the probability of death from other causes. In panel A, we compare them to Deryugina PM2.5. In panel C, we consider the entire population and we compare the probability of getting cancer from an incinerator to the probability of car accident death. In panel D, we restrict to people in working age (18-65 years) and we compare the probability of getting cancer from an incinerator to the probability of workplace death. In panel E, we compare the total Value of Statistical Life (we use an estimate of the VSL in Denmark of 31 million DKK, or 4.28 million USD [a]) lost due to the incinerators with the total operating costs of burning waste and the total environmental taxes that are currently paid for burning waste. **Probabilities are expressed in basis points.**

[a] https://dors.dk/files/media/konferencer/konference_2016/DANNERSALONEN/01_merete_kjeldsen_the_value_of_statistical_life_in_denmark.pdf

Table 23: SO2 Concentration from Monitors as a Function of HyADS Based on Actual and Fitted Emissions of SO2. This table shows the results for OLS specifications in which the dependent variable is the annual average concentration of SO2 (in μg/m^3) measured by the monitors of the University of Aarhus. The main explanatory variable is HyADS based on emissions of SO2 from incinerators. Each monitor is matched with a parish depending on its coordinates. HyADS is computed by running HYSPLIT; every year, the release of 100 particles is simulated from each active plant, 4 times per day; the concentration of the particles from each plant is computed at the parish level and weighted by emissions of SO2. The latter are either drawn from the Green Accounts (available in 1996-2001 and in 2007-2018) or computed as the fitted values from a regression in which actual emissions of SO2 are predicted as a function of imputed emissions of SO2 and concentrations of SO2. Imputed emissions are computed as the product between emissions factors (from the CORINAIR Program) and amount of waste burnt (from the Danish Energy Agency); for the period prior to 1990, we assume that the 1990 emissions factors stayed constant; for the period prior to 1994, we discount the 1994 data on waste burnt by real GDP growth. Concentrations are downloaded from the MERRA-2 dataset provided by the NASA; for the period prior to 1980, we assume that the 1980 concentrations stayed constant; the 0.5^{o}x0.625^o resolution of the NASA is matched with the shapefile of the Danish parishes. We control for year and monitor fixed effects. Standard errors are in parentheses; *, **, and *** mean statistically significant at the 10, 5, and 1% levels.

SO2 Monitors	(I)	(II)	(III)
SO2 HyADS	6.395***	3.481***	4.467***
	(0.592)	(0.253)	(0.284)
Constant	2.494***	4.966***	19.019***
	(0.748)	(0.291)	(0.898)
Year FE	No	Yes	Yes
Monitor FE	No	No	Yes
R^2	0.526	0.961	0.986
N	107	107	107