Noise Pollution and Infant Health

Click here to download the latest version of the full paper.

Simon Greenhill*

February 6, 2026

Abstract

Noise, or unwanted sound, is ubiquitous in urban environments. I introduce a new approach to measuring noise at scale using seismometers and compile the largest publicly available database of ambient noise. Using a research design leveraging idiosyncratic variation in electric passenger rail noise exposure, I estimate the impact of noise pollution on infant health. In utero noise pollution exposure harms health at birth. A 2 decibel increase in average noise levels during pregnancy—a small but perceptible increase—lowers an overall index measure of infant health by 4 percent of a standard deviation, equivalent to one-third of the Black-White gap in the index. This effect is driven by nighttime noise, suggesting disruptions to maternal sleep as a main mechanism. Overnight rail services, which account for under 10 percent of overall ridership, generate an average externality of $14 per trip. In per passenger-mile terms, overnight rail noise externalities are comparable to rush hour traffic congestion externalities from private vehicles. Using seismic data and machine learning, I produce a novel map of noise for the contiguous United States and use this map to assess noise exposure and costs nationally. Eighty percent of urban residents are exposed to potentially harmful levels of nighttime noise. I estimate that the cost of noise pollution due to harms to health at birth is $8.4 billion per year. Urban, Black, and Hispanic Americans disproportionately bear these costs.

*Agricultural and Resource Economics, UC Berkeley, and Global Policy Laboratory, Stanford University Doerr School of Sustainability. Contact: sgreenhill@berkeley.edu. I thank Joseph Shapiro, Solomon Hsiang, and Maximilian Auffhammer, and Reed Walker for their advice and support throughout this project. Thanks to the many people who provided helpful feedback, especially Michael Anderson, Thomas Bearpark, Lucy Hackett, Dylan Hogan, Kendra Marcoux, Enrico Moretti, Aaron Watt, and members of the Global Policy Laboratory. The California Policy Lab and the Giannini Foundation provided generous financial support.

> "When [most economists] are prevented from sleeping at night by the roar of jet planes overhead [...], are unable to think (or rest) because of the noise and vibration from passing trains [...], their nerves and mental balance disturbed, they proceed to declaim about the disadvantages of private enterprise and the need for Government regulation."
>
> —Ronald Coase, "The Problem of Social Cost," 1960.

1 Introduction

Noise, or unwanted sound, is a canonical externality. Ronald Coase's seminal article introducing the Coase Theorem uses noise as an example 19 times, focusing on noise as a cause of annoyance, lost productivity, and disturbed sleep (Coase, 1960). The medical literature finds that short-term noise exposure elevates blood pressure, stress hormone levels, and oxidative stress, suggesting potential long-term health risks from noise exposure (Munzel et al., 2014). However, the health impacts, economic costs, and distributional inequities in ambient urban noise exposure remain poorly understood.

This paper studies the health impacts and economic costs of noise pollution. To do so, it builds the largest publicly available database of ambient noise in the United States (US), estimates the effect of noise exposure on health outcomes, assesses the economic cost and distribution of noise exposure nationwide, and evaluates implications for transportation policy. I compile seismic data from nearly a thousand seismometers to construct a new database measuring ambient noise across the country. I combine these data with a research design exploiting quasi-experimental variation in noise to estimate the effect of in utero noise exposure on infant health and the resulting economic costs due to changes in infant mortality risk, healthcare expenditures, and lifetime income. Using seismic data and a frontier geospatial machine learning foundation model, I produce a map predicting ambient noise exposure nationally. I use this map to characterize the distribution of noise exposure across the population and estimate the national cost of noise pollution due to harms to infant health. Finally, I compare the magnitude of the noise externality to other transportation externalities and discuss implications for transportation policy.

In utero noise exposure has a large and precisely estimated negative effect on infant health. A 2 decibel increase in average prenatal noise exposure—equivalent to the difference in average noise levels between six a.m. and noon, or between a Sunday and a Tuesday—decreases an index of infant health by 4 percent of a standard deviation, about one-third of the Black-White gap in the index. The effect is concentrated in overnight exposure, consistent with disruptions to maternal sleep as a main mechanism. Nationally,

urban, Black, and Hispanic Americans experience higher nighttime noise levels than rural and White populations. I estimate that the national cost of noise pollution due to harms to infant health is $8.4 billion per year. Expressed per passenger-mile, peak overnight rail noise externalities are comparable in magnitude to rush hour traffic congestion externalities from private vehicles.

A key challenge in studying noise pollution is a lack of readily accessible ambient noise data. I overcome this challenge by using data from seismometers, devices used to measure vibrations in the Earth's surface that function similarly to microphones. Seismometers have captured anthropogenic noise for over a century; seismologists now use these data to measure human-caused noise from transportation to concerts (e.g., Gutenberg, 1911; Riahi and Gerstoft, 2015; Caplan-Auerbach et al., 2024). In the absence of a national noise monitor network, the nearly 1,000 seismometers that continuously record across the US present a potentially useful and untapped trove of data for measuring ambient noise. I compile, process, and analyze 5 terabytes of publicly available seismic waveform data, creating what I believe is the first large-scale database of seismic noise for use outside seismology.

To estimate the causal effect of noise on health, I exploit quasi-random variation in exposure to noise from electric rail systems in four major California urban areas: Los Angeles, Sacramento, the San Francisco Bay Area, and San Diego. Electric trains are a major urban noise source, but unlike freight trains, road traffic, or aircraft, they do not emit local air pollution, allowing me to isolate the health effects of noise from the effects of air quality. These subway and light-rail passenger lines operate predictable, high-frequency service across densely populated neighborhoods—over 2.5 million infants were born between 1999 and 2021 to mothers living within 5 km of an electric rail line, nearly a quarter of all infants born in California over the same period. I combine hourly wind-direction data with detailed geospatial information on rail lines to measure excess in utero noise exposure as the share of pregnancy spent downwind of a line, building on theoretical and empirical results from the acoustics literature showing that wind increases sound propagation downwind. The outcome data are drawn from confidential administrative records on the universe of California births between 1999 and 2021, which include rich measures of health at birth such as birth weight, gestation length, and Apgar scores; parental demographics; and maternal residence addresses.

I provide several tests supporting the validity of this research design. First, seismometers register higher noise levels on hours when they are downwind of an electric rail line, confirming that wind direction is a major source of short-run variation in noise exposure. Second, air pollution measures from US Environmental Protection Agency (EPA) air quality monitors are uncorrelated with the downwind measure, providing evidence that the research

design is able to separate changes in noise from changes in air pollution. Third, fixed maternal characteristics are uncorrelated with the share of the pregnancy spent downwind of a line, conditional on census block and year by month fixed effects, suggesting that the noise exposure measure is not confounded by omitted household characteristics.

I translate health impacts into monetary terms using existing estimates linking poor health at birth to higher medical costs and lower lifetime income. I value increased mortality risk via the value of a statistical life. The social cost of the harms to infant health caused by electric rail noise in California is $136 million per year. Because the effect is concentrated in overnight noise——and nighttime passengers account for only about 8 percent of total ridership——this implies a large per-trip externality of $14. Using the estimated relationship between downwind status and noise levels and methods from the heterogeneous two-sample instrumental variables literature (Angrist and Krueger, 1992; Zhao et al., 2019), I estimate that the marginal cost of a 1 decibel (2.5 percent) increase in average in utero noise exposure is just under $2,800 per infant.

To assess transportation policy implications, I calculate hourly externalities per passenger-mile for trains, buses, and cars due to noise and traffic congestion. Congestion provides a useful benchmark as it represents the largest externality in the transportation literature (Parry and Small, 2009). During overnight hours, the external cost of noise per passenger-mile is comparable to rush hour congestion externalities from private vehicles. Two mechanisms drive this result: high noise emissions from rail and low background noise during overnight hours combine to cause large marginal noise increases, and ridership during these hours is sparse, yielding high external costs per passenger-mile. These findings add nuance to classical results in the transportation policy literature that consistently identify rail subsidies as welfare-improving (Parry and Small, 2009; Barwick et al., 2024; Almagro et al., 2024).

Finally, I combine the noise data with a geospatial foundation model developed by Google DeepMind (Brown et al., 2025) to produce a new map of average noise exposure across the contiguous US.[1] This model achieves predictive performance an order of magnitude higher than that of the U.S. Department of Transportation's existing map, as measured by the out-of-sample R^2 between predictions and observed noise levels. The resulting estimates reveal substantial inequality in noise exposure: block groups with above-median Black or Hispanic populations are exposed to 10-20 percent more noise than those with below-median shares. A vast majority (80 percent) of urban residents are exposed to nighttime noise levels that could disrupt their sleep.

[1]Foundation models are frontier machine learning models pre-trained on large amounts of data to perform a range of tasks, in contrast so specialized, task-specific machine learning models Google (2025). Large language models such as ChatGPT and Google Gemini are examples of foundation models.

Related literature

This paper provides a credible quantification of the health effects of noise. Foundational work in empirical environmental economics measured greenhouse gas, air, and water pollution externalities (e.g., Mendelsohn et al., 1994; Chay and Greenstone, 2003; Keiser and Shapiro, 2019a), spawning large literatures across subfields of economics. Noise pollution has received little attention despite being a similarly ubiquitous externality.

While existing observational studies estimate cross-sectional correlations between health outcomes and noise exposure, omitted variable bias concerns limit the credibility of these estimates (e.g., Gehring et al., 2014; Smith et al., 2020; Roswall et al., 2021; Thacher et al., 2022; Jephcote et al., 2023). Combined emission of noise and air pollution is a first-order concern given the well-established negative health consequences of air pollution exposure and the fact that most major noise sources, including road traffic and airplanes, also emit air pollution (e.g., Chay and Greenstone, 2003; Currie and Walker, 2011; Deschênes et al., 2017; Deryugina et al., 2019). Even in the absence of correlated changes in air pollution, households may sort in response to noise and correlated neighborhood characteristics, introducing bias to cross-sectional comparisons that fail to account for systematic differences between louder and quieter places.

Prior research in economics has studied short-term noise exposure in a lab-in-the-field setting, showing that noise reduces factory worker productivity and cognitive function (Dean, 2024). Other studies have measured the impacts of airliner flight paths and road noise on infant health, crime, and home prices (Argys et al., 2020; Hener, 2022; Allroggen et al., 2025; Moretti and Wheeler, 2025), though these likely capture the joint impacts of noise and air pollution (Hudda et al., 2014; Schlenker and Walker, 2016; Du and Taylor, 2024). The present analysis isolates variation in noise pollution from variation in air pollution, allowing estimation of the effect of noise alone.

Beyond environmental economics, I believe this study is the first to document a plausibly causal relationship between noise pollution and health at birth. A large literature documents the causes and consequences of infant health. Many studies have measured impacts of in utero economic, social, and environmental conditions on health at birth (e.g., Chay and Greenstone, 2003; Currie, 2009; Almond and Mazumder, 2011; Currie and Walker, 2011; Hoynes et al., 2015; Currie et al., 2013, 2022), and another demonstrates the long-term consequences of poor infant health for human capital accumulation and earnings (Almond et al., 2005; Almond, 2006; Black et al., 2007; Maccini and Yang, 2009; Figlio et al., 2014; Isen et al., 2017; Almond et al., 2018; Bharadwaj et al., 2018). This study introduces a new environmental channel through which in-utero conditions affect fetal health—and, consequently, long-run economic outcomes.

The remainder of the paper proceeds as follows: Section 2 gives brief background on noise pollution; Section 3 describes the data; Section 4 introduces the empirical framework for estimating the effect of noise on infant health; Section 5 presents the results, including the impact of noise on health, the costs of noise pollution due to harms to infant health, and implications for transportation policy; Section 6 develops the machine learning model of noise exposure and describes noise exposure and costs nationally; Section 7 concludes.

2 Background: What is Noise Pollution?

Noise pollution is defined as unwanted sound (Münzel et al., 2024). This definition has two equally important components. The first is purely physical: *sound* is an energetic vibration propagated via acoustic waves through air and other mediums. The second part, that to qualify as noise pollution a sound must be *unwanted*, is more psychological: one person's late-night dance party might be their neighbor's noise pollution. Concerts, sports games, and even schoolyards during recess are all examples of sound sources that can be desirable to those producing and directly experiencing the sound, but might be considered noise pollution by others nearby who experience it merely as an interruption or distraction. Other sources of sound, such as transportation noise (e.g., the sound of trucks on a highway), may be undesirable to all those who experience it, but costly to abate. In this paper, I use the terms "noise" and "noise pollution" interchangeably.

Randomized lab studies have documented physiological responses to short-run noise exposure, finding that the body has a stress response to noise. Noise exposure, especially during the nighttime, causes increases in blood pressure, stress hormone levels, and oxidative stress, reducing sleep quality and quantity (see Munzel et al., 2014, for a review). Several lab studies have gone a step further and investigated the effect of nighttime noise exposure on both sleep and measures of wellbeing and productivity in subsequent days. For example, Elmenhorst et al. (2010) find nighttime noise exposure reduces reaction times and short-term memory the following day. In a related economics study, Dean (2024) randomly exposed Kenyan textile workers to increased noise during work, finding that increased noise reduced worker productivity and cognitive function.

The negative physiological effects of noise begin below perceptive thresholds. For example, (Basner et al., 2006) found that noise negatively affected sleep even in study subjects who did not recall their sleep being disturbed by the noise.

There is some evidence that the body habituates over time to nighttime noise exposure, reducing the probability of awakening. However, adaptation is not complete: the effects of noise on heart rate and blood pressure do not appear to attenuate with repeated exposure

Basner et al. (2011). This suggests that long-run noise exposure may have negative health effects, even if those exposed are not acutely aware of the noise.

Policymakers have sought to control noise pollution since at least 44 BCE, when Caesar banned wagons from entering Ancient Rome at specific times of day due to the noise they made; during the US constitutional convention, in 1787, the framers had dirt spread over Philadelphia's cobblestone streets to muffle the sound of passersby (Johnson et al., 2003; National Archives, 2015). Three major noise sources in modern contexts are trains, planes, and cars. Other noise sources include construction, outdoor air conditioning units, eletrical substations, and industrial plants (National Academy of Engineering, 2010).

Noise regulation is virtually nonexistent in the US.[2] Congress passed the Noise Control Act in 1972, the same year it passed the Clean Water Act and nine years after passing the Clean Air Act. While the Clean Air and Clean Water Acts are cornerstones of US environmental regulation and have been the subject of detailed study (Currie and Walker, 2019; Keiser and Shapiro, 2019b), the Noise Control Act has languished in obscurity. In 1982, the Reagan administration defunded the Environmental Protection Agency (EPA)'s Office of Noise Abatement and Control, effectively ending federal noise regulation in the US.

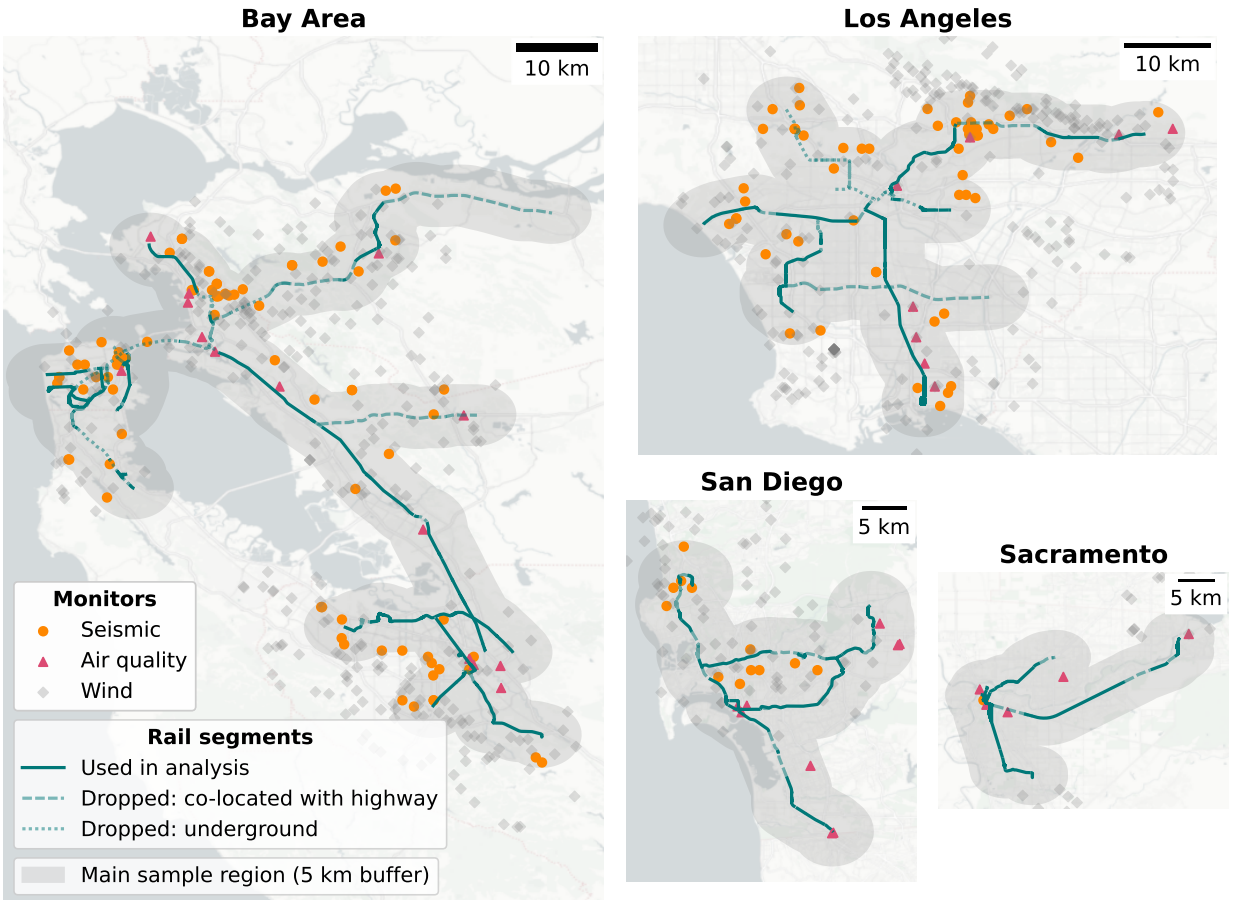
3 Data

3.1 Noise Data

There are no large-scale efforts to measure noise exposure in the US. The EPA maintains national sensor networks measuring air and water pollution, ensures consistent measurement standards, and makes the resulting data readily available to the public. In contrast, noise monitoring is ad hoc and haphazard: different jurisdictions operate their own small noise monitor networks with little if any coordination. For example, in Santa Clara County, California, the San Jose Mineta International Airport operates a network of 11 noise monitors directly under the airport's landing and takeoff flight paths, while the neighboring cities of Santa Clara and Sunnyvale operate networks of four monitors each. Santa Clara's monitors are clustered near Levi's Stadium, where the San Francisco 49ers NFL team plays, while Sunnyvale's are scattered throughout the city and primarily report noise from the San Jose airport and Moffett Federal Airfield, a joint military-civil airport abutting Sunnyvale to the North. Even though each jurisdiction operates noise monitors for similar reasons—and even to track noise from the exact same sources—there appears to be little data sharing between

[2] In contrast, other wealthy countries have extensive noise regulation. For example, Europe's Environmental Noise Directive requires all member countries to assess noise exposure and develop noise mitigation plans for all major roads, railways, and airports, as well as for all agglomerations of 100,000 people or more.

Figure 1: California Electric Rail Lines, Seismometers, Air Quality Monitors, and Weather Stations

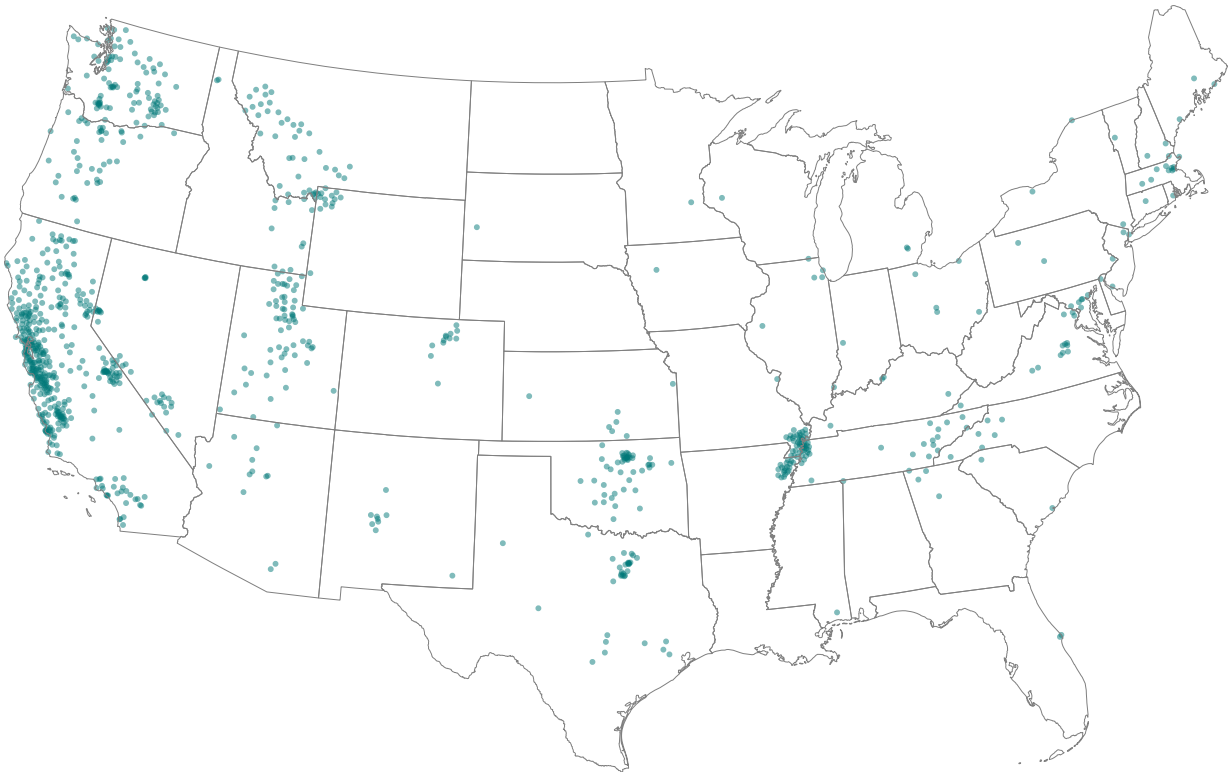


Notes: The map shows the locations of electric rail lines, seismometers, air quality monitors, and weather stations in my analysis sample. I keep rail lines shown in solid teal for the main analysis. I drop underground rail lines and rail lines that are co-located with a highway. These are shown as dotted and dashed lines, respectively. A 5 km buffer around all rail lines is shown in light gray. Seismometers and EPA air quality monitors within 5 km of any electric rail line are marked by orange circles and pink triangles, respectively. Weather stations recording wind data within 10 km of either a rail line, a seismometer, or an air quality monitor are marked by light gray diamonds. Panels are not shown to scale; see scale bars in the top right of each panel.

the jurisdictions, and historical data is not made available to researchers or policymakers.

Imperfect models of noise exposure take the place of direct noise measurement in many applications. Modeled noise estimates suffer from significant measurement error due to simplifying assumptions. In the US, the main noise model for the US is the National Transportation Noise Map, developed by the US Department of Transportation. The model makes drastic simplifying assumptions, including ignoring topography and buildings and assuming all ground cover absorbs noise. The model is not validated using direct noise measurement.

Figure 2: Seismometers in the Continental US



Notes: The map shows the locations of the 975 seismometers that were recording continuously across the US in 2018.

For these reasons, the Web page hosting the National Transportation Noise Map reads, "[the National Transportation Noise Map] is intended for the tracking of trends and should **not** be used to evaluate noise levels in individual locations and/or at specific times" (Bureau of Transportation Statistics, 2022, formatting in original). In many studies, researchers simultaneously rely on modelled air pollution and noise exposure (e.g., Gehring et al., 2014; Smith et al., 2020; Thacher et al., 2022). Measurement problems may be particularly pernicious in these cases, as models of noise and air pollution transport often take the same variables as inputs, such as the amount, composition, and proximity of road traffic, and omit other variables, which could lead to correlated errors (Munzel et al., 2014). Appendix B.1 further discusses measurement issues in the National Transportation Noise Map and other noise models.

3.1.1 Using Seismometers to Measure Noise

To overcome the absence of data on ambient noise exposure, I turn to measurements from seismometers.[3] Seismometers, which are designed to measure the ground's vibration (including earthquakes), consist of a suspended mass and a housing which is fixed to the ground. When the ground shakes, the housing moves relative to the mass. The monitor measures quantities such as the displacement, velocity, and acceleration of this movement to quantify the strength of the ground's shaking.

The way seismometers measure waves is analogous to the way microphones measure them. Microphones contain a diaphragm inside a fixed housing, and measure waves by measuring the movement of the diaphgram caused by the changes in pressure due to the wave. Indeed, the component of a seismometer that directly measures vibrations is called a geophone, reflecting similarity to microphones. Figure 1 is a schematic illustrating this similarity.

When sound waves moving through the atmosphere come into contact with the ground, they cause vibrations which may be detectable by seismometers. Empirically, seismologists have observed human noise sources such as traffic in seismic records since the early 20th century (Gutenberg, 1911). Recent studies have demonstrated that seismometers capture noise from traffic, public transportation, and sports games, documented overall noise reductions during COVID-19 lockdowns, and even identified specific songs played at a nearby Taylor Swift concert (Denton et al., 2018; Lecocq et al., 2020; Riahi and Gerstoft, 2015; Caplan-Auerbach et al., 2024).

Seismologists largely view "seismic noise" caused by humans as a nuisance, impeding the study of natural phenomena such as earthquakes and gravitational waves (Lecocq et al., 2020). In the absence of data measuring environmental noise directly, however, seismometers can provide a valuable record of noise exposure: there are thousands of seismometers recording continuously around the world.

3.1.2 Seismic Data Sample Selection and Processing

I gather two samples from the global seismic record. First, I gather all publicly available data from monitors within 5 km of an electric rail line in California between 1999 and 2021. Second, in ongoing work, I gather data from all monitors in the US in 2018.[4] I use these data to assess average noise pollution exposure in the US. For both samples, I require each

[3]Seismometers are also referred to as seismometers and seismographs. While these can refer to different components of instruments used to measure and record ground motion, the terms are often used interchangeably.

[4]I focus on 2018 data for US-wide measurement to match the timing of the satellite imagery I use to train my machine learning model predicting average noise levels across the US.

monitor to record continuously for at least one month of the desired period and that the data be publicly available.

Publicly available seismic data come from two main sources: networks of seismometers operated by governments and academic institutions for research and earthquake monitoring, and single monitors belonging to schools, museums, or private individuals. Data from both sources are automatically uploaded to a public online repository. The repository is maintained by the International Federation of Digital Seismograph Networks (FDSN), a non-governmental organization with member institutions in over 50 countries. In addition to providing access to seismic data, FDSN maintains standards for the quality and format of seismic data. FDSN data are the backbone of rapid earthquake response and tsunami warning efforts and are widely used in seismologic research (Suarez et al., 2008). I programmatically accessed FDSN metadata to identify seismometers and download one file per day per seismometer, resulting in an archive of one million files measuring 5.4 terabytes.

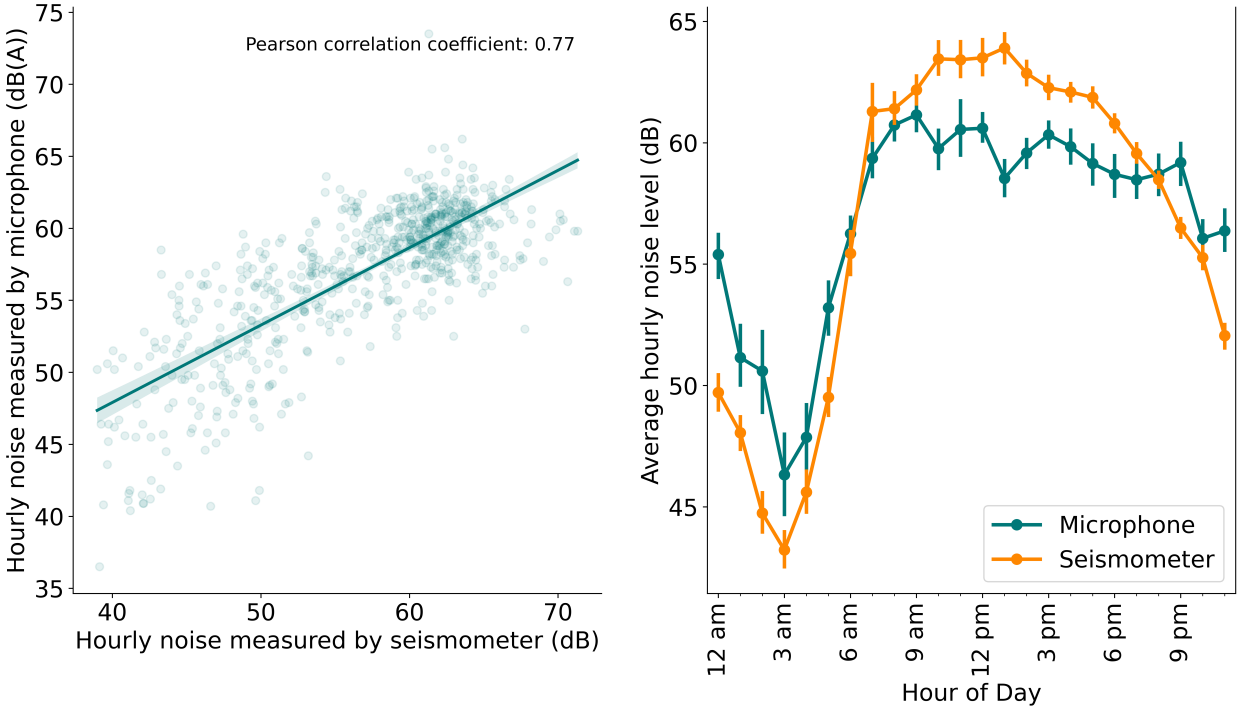
Figure 1 maps the locations of the California seismometers and Figure 2 maps the locations of the 1,837 seismometers across the US I use data from. Each of the four metropolitan areas I focus on in my main empirical analysis has at least one seismometer; three out of the four have 10 or more. Seismometer locations across the country are primarily determined by seismological factors. For example, seismometers are clustered in seismically active areas such as California and Oklahoma. In some cases, seismometers are densely clustered, such as along the border between Massachusetts and Connecticut, reflecting research efforts to study the Earth's outer shell in a specific geographic area. In other cases, seismometers are placed near population centers to measure seismic risk to people and their homes. Individually owned monitors are typically placed in private homes, or in schools, museums, community centers, and other locations where they are used for science education.

I use signal processing methods developed by seismologists to process raw waveform data into a dataset of hourly average noise levels at each seismometer's location. Appendix C.1 offers details and equations; I summarize the methods here. Raw seismic data are a time series of waveforms measuring ground motion at a high frequency, typically 100 observations per second. To calculate hourly average noise for each monitor, I estimate a power spectral density of the seismometer's vertical displacement for each one-hour period in the data. This provides an empirical distribution of the amount of energy measured at each frequency. I focus on noise at frequencies between 4.0 and 20.0 Hertz, corresponding to the frequencies of anthropogenic noise sources documented in the seismology literature (Bonnefoy-Claudet et al., 2006). I then integrate the empirical distribution to obtain the mean squared displacement of the seismometer during the hour. The square root of this quantity is the root mean squared displacement d measured by the seismometer.

3.1.3 Validating the Use of Seismometers for Measuring Noise

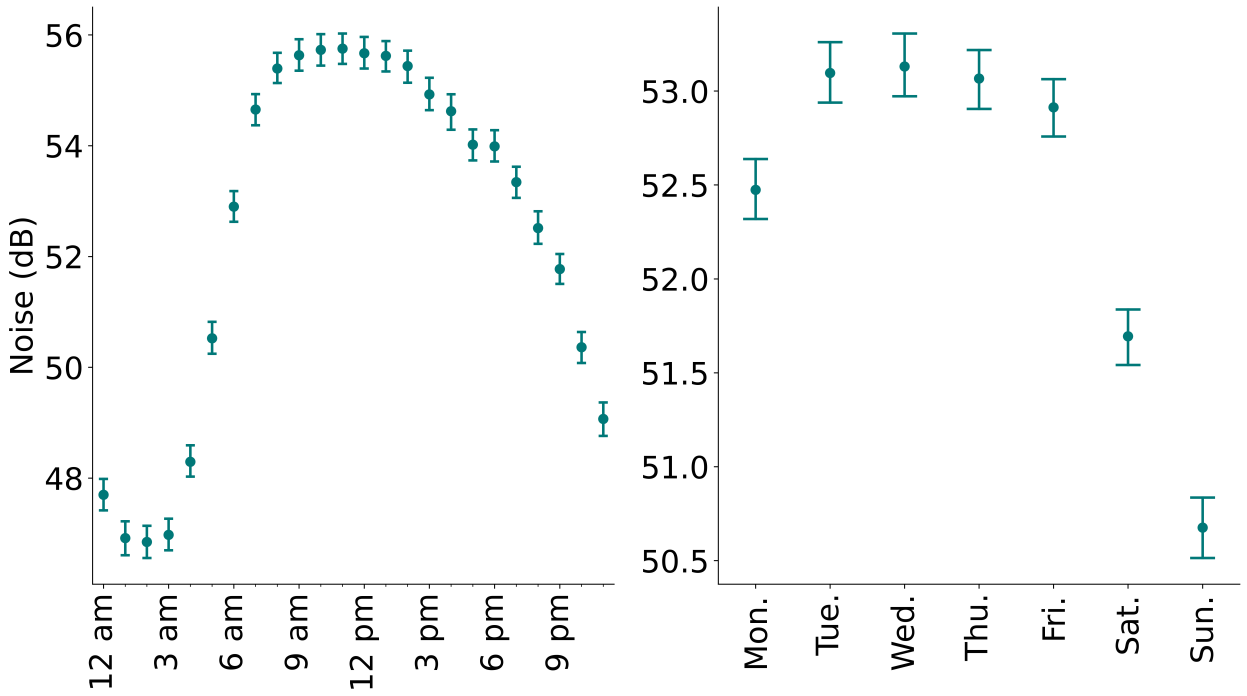
I conduct two empirical exercises to validate the use of seismometers for measuring noise pollution, building on findings from the seismology literature that seismometers capture anthropogenic noise sources. First, I conduct a case study of a location where a seismic noise monitor and a microphone noise monitor are nearly co-located. I use temporally matched hourly data from a microphone monitor and a seismometer located approximately 500 m apart in Foster City, California. The seismic data are publicly available. The microphone data are collected by the San Francisco International Airport; I obtained the data for January 2023 via a request to the Airport. While a full investigation of the relationship between seismic noise measures and microphone noise measures would require long-term co-located measurements of seismic and microphone noise at a variety of locations, this analysis offers suggestive evidence.

Figure 3: Comparison of seismic and microphone noise measures



Notes: The figure shows comparisons of seismic and microphone noise measures using hourly data from January 2023. Microphone data come from a noise monitoring station operated by the San Francisco International Airport. Seismic noise data come from a RaspberryShake seismometer located about 500 m away. Both monitors are in Foster City, California. The left panel shows a scatter plot and regression line of the base-10 logarithm of hourly seismometer displacement, in nanometers (nm) on the x-axis, and audible sound level, in decibels (dB), on the y-axis. The right panel shows means and 95% confidence intervals for the noise levels in each hour of the day for the month. The teal series corresponds to microphone noise, with units on the left axis. The orange series corresponds to seismic noise, with units on the right axis.

Figure 4: Within-Day and Within-Week Noise Patterns



Notes: The figure plots the mean and 95% confidence interval of noise levels at each hour of the day across all monitors in my sample. I take logs then normalize the noise level before calculating means.

Figure 3 summarizes the case study. The left panel shows that hourly seismic and microphone noise are positively linearly correlated, with a Pearson correlation coefficient of 0.77. The right panel shows that the magnitude and pattern of the within-day variation in both measurements is similar. Noise levels are low overnight, reaching their lowest point in the 3 a.m. hour, and are elevated during the day, staying relatively constant between 7 a.m. and 7 p.m. before falling again.

I complement this case study by investigating patterns in the full sample of seismic noise data. While I do not have co-located microphone measures for each location, observing similar patterns to those in Figure 3 provide evidence that the seismic noise measurements in my full sample similarly reflect the measurements that would be captured by a microphone. The left panel of Figure 4 plots the average hourly noise across all seismometers in my sample. As in Figure 3, seismic noise follows a strong diurnal pattern. Noise levels are at their lowest in the early morning hours, then rise gradually between 4 and 8 am, and remain elevated until 5 pm. They then decline gradually over the course of the evening and into the night. Seismic noise also follows an intuitive weekly pattern, which is plotted in the right panel of Figure 4. Average daily noise is higher on weekdays than on weekends, with the highest values on Tuesday, Wednesday, Thursday, and Friday. Saturday and especially are

significantly quieter.

3.2 Outcome Data: Infant Health

I focus on health at birth for three reasons. First, health at birth is an important determinant of early- and later-life health, medical spending, educational attainment, and economic well-being (e.g., Black et al., 2007; Currie, 2009; Figlio et al., 2014). Social policies including Medicaid and the Women, Infants, and Children program directly target and evaluate their effectiveness on infant health outcomes, reflecting the importance policymakers place on these outcomes. Second, a pregnancy is a well-defined exposure period over which to assess the impact of noise on health. I measure in utero noise exposure, and health outcomes are measured at precise end of the look-back period. Third, data on infant health outcomes are consistently tracked for large populations over long periods. In this study, I use data on the universe of births in California between 1999 and 2021. I obtain these data from California vital statistics files accessed from the California Department of Public Health via a research data request.[5] The vital statistics database contains measures of infant health such as gestation length, birth weight, and Apgar scores, rich parental characteristics, and mothers' home addresses.

Noise, like other in utero shocks, could potentially affect many birth outcomes. Canonical studies in the fetal origins literature have focused on birth weight (Almond et al., 2005), although more recent work has examined a wider range of outcomes, such as Apgar scores, congenital anomalies, delivery complications, fetal death, gestation length, neonatal intensive care unit admissions, and sex ratios, among others (see Almond et al., 2018, for a review). The wide range of outcomes of potential interest introduces both statistical and expository challenges: a researcher studying many outcomes should adjust p-values to account for multiple hypothesis testing; displaying and synthesizing results across many outcomes succinctly is difficult or impossible. To avoid these issues, I focus primarily on an infant health index, which I construct following methods from the multiple inference literature (Anderson, 2008; Kling et al., 2007), as is common in the recent infant health literature (Aizer and Currie, 2014; Currie et al., 2017, 2022). In addition to solving statistical and expository issues, using the index maximizes my statistical power. Intuitively, if noise marginally affects a wide range of outcomes, parameters from a model where an index of those outcomes is on the left-hand side will be more precisely estimated than parameters from models where each outcome is considered separately (Anderson, 2008).

I construct the infant health index by taking a weighted average of birth weight, an

[5]Vital statistics records prior to 1999 did not include mother's addresses, which I need to accurately map noise exposure to pregnancies. I therefore do not use data on births occurring before 1999.

indicator for low birth weight (birth weight less than 2500 grams), an indicator for very low birth weight (birth weight last than 1500 grams), an indicator for preterm birth (gestation less than 37 weeks), an indicator for very preterm birth (gestation less than 34 weeks), an indicator for whether the infant died, and 5- and 10-minute Apgar scores. The weights are given by the inverse of the covariance matrix of the outcomes after transforming the outcomes by standardizing them and changing their signs so that the positive direction always indicates a "better" outcome. Further details on the construction of the index are in appendix section C.2.1. Appendix table 12 reports the weights. I also report impacts on each of the components of the infant health index. Because Apgar scores are only available from 2007 on, the infant health index is defined for the universe of births in California between 2007 and 2021.

The California vital statistics data allow me to observe other information pertaining to each birth in addition to measures of the infant's health at birth. Most importantly, I observe the mother's residential address, which I use to assign noise exposure. I also observe maternal demographics, such as age, race, and educational attainment. Appendix C.2.2 details how I clean the data and construct covariates.

To investigate how the main analysis sample compares to the overall California population, Appendix Table 2 reports the means and standard deviations of key health outcomes and maternal demographic characteristics. Column (1) reports statistics for births in the main analysis sample, within five km of an electric rail line in California. Column (2) reports statistics for California births outside the main analysis sample. Column (3) reports the differences between the in sample and out of sample means divided by the variable's standard deviation in the full sample. Panel A describes birth outcomes. These are quite similar across both the in sample and out of sample populations, differing by no more than 3.4% of a standard deviation. Panel B describes maternal characteristics, which differ somewhat more between the in-sample and out-of-sample groups. The share of Black mothers has the largest standardized difference of just under 20%, reflecting the fact that most of California's Black population is urban. Conversely, the share of Hispanic mothers is lower in the analysis sample, reflecting the large rural Hispanic population in California. Other variables differ only modestly between the analysis sample and the remainder of the population.

3.3 Other data sources

Rail lines and highways. I obtain spatial data on rail lines from OpenStreetMap, an open-source geographical database maintained by volunteer contributors. I observe the locations of all rail lines, whether they are electrified, and what segments are underground. I also

gather opening dates and operating hours for each electric rail line segment in California. Spatial data on the locations of highways comes from the National Highway System. I use these data to identify rail segments that are co-located with highways. Appendix C.3.1 details how I compile and validate these data.

I focus on interstate and non-interstate highways as well as highways in the Department of Defense's Strategic Highway Network. To identify highways that are co-located with rail lines, I create 50 m buffers around each highway and rail line segment, and label any rail line segment overlapping more than 10% of its area with any buffered highway segment as coincident with a highway.

Figure 1 maps California's electric rail lines and notes which rail segments are underground or co-located with a highway. There are six major electric rail systems in California: Bay Area Rapid Transit, Los Angeles Metro Rail, Sacramento Regional Transit Light Rail, San Diego Metropolitan Trolley, San Francisco Muni Metro, and Valley Transit Authority Light Rail (serving the southern Bay Area, including San Jose).[6] All are passenger light rail or mass transit systems. There are 1,270 km of electric rail in California, of which 320 km (25%) are co-located with highways and 190 km (15%) are underground. The remaining 760 km (60%) define the geographic focus of my main analysis.

Weather. Hourly temperature and precipitation come from ERA5-Land, a standard gridded reanalysis weather dataset with 0.1 degree ($\approx$ 11 km) resolution (Muñoz-Sabater et al., 2021). ERA5-Land has been found to perform similarly to higher resolution data such as PRISM in econometric analyses.[7] However, analyses of ERA5-Land and ERA5's accuracy for near-surface winds have found large and systematic discrepancies between the reanalysis model's predictions and wind observations, particularly in rugged terrains (Ramon et al., 2019; Gualtieri, 2021; Gadal et al., 2022). I therefore turn to weather station data to measure wind. I access hourly wind direction and speed data from weather station data via the Synoptic Weather API, which harmonizes data from over 150,000 weather stations globally; I select weather stations that are within 10 km of a rail line. Figure 1 maps the 653 weather stations in my sample.

I match temperature and precipitation data to seismometers, air quality monitors, and mothers' residential addresses using the nearest ERA5-Land grid cell centroid. I match wind data to seismometers and air quality monitors by taking an inverse distance weighted average

[6]Caltrain, a light rail system serving the San Francisco Peninsula, transitioned from diesel to electric locomotives in fall 2024. The infant health data I use end in 2021; Caltrain is therefore not considered an electric rail system in this analysis.

[7]For example, econometric analysis using ERA5-Land correctly recovers the nonlinear temperature-crop yield relationship (Hogan and Schlenker, 2024).

of the wind observations for all stations within 10 km of each monitor. It is computationally infeasible to calculate inverse distance weighted wind for each of the 2.6 million births in my sample. Instead, I create a 1 km grid over California and calculate hourly wind at each grid cell centroid, using an inverse distance weighted average of all weather stations within 10 km of the grid cell centroid. To investigate the extent to which this aggregation could introduce measurement error, I calculate wind data for 10 randomly sampled points within 1000 randomly sampled 1 km grid cells and compare the wind direction at each point to the wind direction at the centroid of the grid cell in each hour.

Air pollution. I obtain daily data on concentrations of particulate matter less than 2.5 microns in diameter ($PM_{2.5}$), Carbon Monoxide (CO), and the Air Quality Index (AQI) at EPA air quality monitors.[8] My sample selection for air quality monitors is identical to my sample selection for seismometers: I gather data from all monitors within 5 km of an electric rail line in California. This yields 36 monitors I additionally obtain satellite-derived gridded data on monthly $PM_{2.5}$ concentrations on a 0.01 degree ($\approx$ 1 km) resolution grid between 1998 and 2021 from Shen et al. (2024).[9]

4 Empirical framework

I introduce a research design to estimate the effect of noise on health in the context of rail noise in California. An ordinary least squares regression of noise exposure on an economic outcome will almost certainly suffer from omitted variable bias. There are two main sources of potential omitted variables bias in this setting: combined emission of noise pollution with other pollutants and sorting due to heterogeneous willingness to pay for quiet or for correlated amenities. I discuss each in turn.

Noise is frequently co-emitted with air pollution. Many noise sources, including most airplanes, cars, trucks, and trains, are powered by internal combustion engines which emit airborne pollutants including carbon monoxide, nitrogen oxides, sulfur dioxide, and fine particulate matter. A vast literature has found these pollutants negatively affect many economic outcomes at every stage of the lifecycle (e.g. Chay and Greenstone, 2003; Currie and Walker, 2011; Schlenker and Walker, 2016; Deschênes et al., 2017; Deryugina et al., 2019;

[8]The Air Quality Index is a summary measure constructed from measures of ozone, $PM_{2.5}$, particulate matter less than 10 microns in diameter (PM_{10}), carbon monoxide, sulfur dioxide, and nitrogen dioxide. An index value is calculated for each pollutant separately, and the Air Quality Index is the maximum of the pollutant-specific index values. See EPA (2018) for details.

[9]These data are often colloquially referred to as the "Van Donkelaar" data, in reference to van Donkelaar et al. (2010), the study which introduced the first version of these data. I use the 2024 version.

Anderson, 2020). Thus, a research design seeking to recover the effect of noise alone must isolate variation in noise which is unrelated to variation in air pollution.

Separately estimating the effects of noise pollution and air pollution is an important research objective because the policies and technologies to regulate and mitigate noise and air pollution are distinct. A first-best set of policies would thus address each externality separately (Goulder and Parry, 2008). For example, above 20 miles per hour, over 90% of road noise is due to cars' aerodynamics, tires, and the road surface, and is unrelated to the car's engine type (van Blokland and Peeters, 2009). This suggests that policies aimed at internalizing the global climate and local air pollution externalities of driving internal combustion engine cars will have limited co-benefits in terms of mitigating noise externalities for most vehicles.

Sorting-related omitted variable bias concerns remain even if a noise source does not co-emit air pollution. Consider a thought experiment in which a set of neighborhoods in an urban area are randomly allocated a new electric light rail commuter line. While noise pollution would increase in these neighborhoods, the neighborhoods would also experience changes in other characteristics such as amenities and market access. These changes might affect the neighborhoods' economic outcomes, for example through changes in the neighborhood's demographic composition in response to the change in nearby transport infrastructure. A comparison of neighborhoods with and without the new rail line, even under random assignment and without a change in air pollution, would thus reflect not only the change in noise pollution, but also the change in the neighborhood's market access, amenities, and demographic composition.

4.1 Variation in Noise Exposure Due to Changes in the Wind Direction

The wind direction affects sound waves' movement through the atmosphere, causing changes in sound levels upwind and downwind of noise sources such as rail lines, roads, and airports. A noise source emits sound waves in all directions; the wind direction affects how the sound waves bend (refract) as they move away from their source, with opposite effects in the upwind and downwind directions. Downwind of a noise source, sound waves bend toward the ground, leading to more total noise at ground level. Conversely, sound waves upwind of a noise source bend away from the ground, leading to less total noise at ground level. Field measurements have found that the wind direction can lead to a 30 dB difference in loudness between equidistant locations upwind and downwind of a noise source, roughly equivalent to the difference between light rain and a vacuum cleaner (Wiener and Keast,

1959; Kriebel, 1972; Willshire, 1985; Hornikx et al., 2018).[10] Appendix D.1 discusses the physical mechanism causing changes in noise exposure depending on the wind direction in further detail.

Changes in the wind direction are a leading source of hourly, daily, and seasonal variation in noise exposure. Appendix Figure 15 reproduces a diagram from a fact sheet released by the Illinois Department of Transportation titled "Why is it so Loud Today?" which explains the wind direction's role in variation in noise exposure and illustrates the physical mechanism (Sperry, 2023). Similar fact sheets have been released by the Texas and Wisconsin Departments of Transportation, as well as by the National Academies of Sciences, Engineering, and Medicine (National Academies of Sciences, Engineering, and Medicine, 2018; Wisconsin Department of Transportation, 2019; Sperry, 2023).

A potential concern with using changes in the wind direction to generate variation in noise exposure is that the wind also carries air pollution (e.g., Anderson, 2020; Deryugina et al., 2019; Herrnstadt et al., 2021). To address this concern, I focus on noise from electric rail systems by measuring increased noise exposure downwind of electric rail lines. Electric trains do not produce air pollution, getting their locomotive power from an electrified third rail or overhead line instead of from onboard combustion. Moreover, electric trains are a major noise source.

Even though electric trains themselves do not produce air pollution, it is possible that electric rail lines are co-located with other sources of air pollution. For example, the San Diego Trolley's Blue and Green lines are partially coincident with Interstate 5, and Los Angeles Metro Rail's C-Line is almost entirely coincident with Interstate 105. I exclude rail line segments that are co-located with highways, as described in Section 3.3, and test directly for the possibility that my noise exposure measure is correlated with air pollution. I also exclude underground sections of rail lines, where the wind direction has no influence on the propagation of noise from passing trains.[11]

[10] In some cases, upwind sound attenuation generates an "acoustic shadow," where a noise source is inaudible upwind. Historians and acoustical researchers have argued that acoustic shadows played an important role in determining the outcome of military battles prior to the development of modern communications technologies, when armies relied on the sound of nearby troops and battles to make strategic decisions. For example, acoustic shadows have been documented in at least six battles in the US Civil War, including in the battle of Gettysburg, which was reportedly inaudible in locations 10 miles from the battlefield, yet simultaneously audible in Pittsburgh, 150 miles away (Ross, 1998).

[11] A possible additional concern is that highways sometimes are parallel to electric rail lines even if they are no co-located. To address this concern, a robustness check restricts the estimation sample to include only areas between the rail line and highway if there is a highway within 5 km.

4.2 Measuring Wind-Driven Changes in Noise Exposure

To measure variation in electric rail noise exposure driven by the wind direction, I must assign each hour of an expectant mother's pregnancy as either downwind or not downwind of an electric rail line. I implement an analogous procedure to assign seismometers and air quality monitors as downwind or not downwind. Throughout this section, I refer only to maternal addresses for simplicity. Because I do not observe parents' real-time locations during pregnancy, I assume their location is fixed at their home address, which I do observe.

I begin with the simplest definition of downwind: an address is downwind of an electric rail line if the wind direction is within 45 degrees of a perpendicular ray from the address to the nearest rail segment. This definition follows Anderson (2020), and I refer to it as such. It is also closely related to the definition in Herrnstadt et al. (2021).

However, this definition may introduce measurement error because it only considers the closest rail segment to an address. In cases where an address is near a curve in the rail line, or where there are multiple rail lines nearby, defining downwind based only on the nearest rail segment can lead to measurement error in the range of wind directions for which the address is downwind of a rail line.

I calculate the perpendicular ray from each address to each rail segment within five km of the address, then average the directions of the rays, weighting by the inverse of the squared distance between the address and the segment and by the segment's length.[12] An address is downwind if the wind direction is within 45 degrees of the resulting perpendicular ray. I additionally report robustness to the Anderson (2020) definition as well as to a more expansive definition of downwind that takes the union of all 90 degree quadrants centered on perpendicular rays to all rail segments within 5 km of each address. My preferred definition balances between these polar cases by considering all nearby rail segments while accounting for rail segment length and the distance to each segment.

To choose the five km cutoff, I regress downwind status on noise separately for monitors within one km bins from the nearest electric rail line. Appendix Figure 2 reports the results of this exercise. I detect elevated downwind noise that is statistically different from zero for all distance bins under six km. The pattern of the point estimates is consistent with what a physical model of downwind noise propagation would predict: noise increases are largest at intermediate distances, where the wind direction will have the greatest influence on the total amount of noise from the electric train in that location (Salomons, 2001).[13] Within one km of

[12] I use the inverse of the squared distance, rather than inverse distance, to reflect the fact that sound dissipates quadratically as it moves away from its source.

[13] Within about one km of the noise source, sound propagation is primarily driven by non-atmospheric factors. Further away (e.g., two to three km), sound level amplification due to sound waves bending downwind

an electric rail line, being downwind elevates noise levels, but the effect is less than monitors slightly further away, between one and three km from the nearest line. For monitors further away from the nearest line, effects taper off monotonically. The point estimate falls below one dB, the difference in sound pressure level generally accepted as the threshold of human perception, at distances larger than five km. I therefore consider areas within five kilometers of an electric rail line as potentially exposed to elevated noise due to changes in the wind direction.

4.3 Temporal Aggregation

The downwind measure described in the previous section allows me to measure changes in noise exposure due to hourly changes in the wind direction. However, I observe health outcomes at birth, reflecting the accumulation of fetal health over the entire term of the pregnancy. I thus define my in utero electric train noise exposure measure T_{it} for each birth i with pregnancy start date t as the share of pregnancy hours in which the mother's home address was exposed to elevated electric rail noise due to the wind direction:

$$T_{it} \coloneqq \frac{1}{H} \sum_{h=t}^{t+H} D_{ih}. \tag{1}$$

The variable D_{ih} is equal to 1 if the wind direction is such that infant i's mother's residence is downwind of an electric rail line, as defined in the previous section. I additionally require that at least one rail line within 5 km of the address be operating during hour h, and that the wind be blowing at least 0.5 meters per second ($\approx$ 1 mile per hour) during hour h.[14] If any of these conditions are not met, D_{ih} is equal to 0.

The set of hours H denotes the period over which I consider exposure. In my baseline specifications, I consider the full pregnancy. Because the duration of the pregnancy may be correlated with T, I instrument the actual pregnancy duration with the 39-week (6,552-hour) period from the date of the beginning of the pregnancy (Currie et al., 2013). The choice of 39 weeks for this period reflects the mean pregnancy length in my sample.

I consider other definitions of H in extensions to my main results, allowing me to calculate different exposure measures T. I calculate T separately for exposure at different times of day, allowing elevated noise exposure to have differential effects during the daytime, evening,

becomes the dominant factor. Note the contrast to downwind propagation of air pollution, which decreases monotonically with distance from the source (Zhu et al., 2006; Hu et al., 2009). Section 5.2 shows that the effect of noise exposure on infant health mirrors the pattern shown in Appendix Figure 2.

[14] Requiring that the rail line be operating accounts for the fact that California's electric rail systems do not operate 24 hours a day.

and nighttime. Measuring exposure separately for different times of day helps address the possible concern that T is measured with error because expectant mothers likely do not spend all their time at their home address. Individuals are more likely to be at their home addresses in the evenings and at night, so exposure measures during those hours will more accurately reflect actual variation in noise exposure. Moreover, the effects of evening and nighttime exposure are of independent interest, because the medical literature suggests that we are most vulnerable to noise exposure during these periods due to the increased likelihood that noise will cause sleep disturbances (Munzel et al., 2014). Air pollution is also lower during nighttime (LaDochy and Witiw, 2023), reducing the likelihood that the nighttime exposure measure is contaminated by air pollution. I also consider exposure before and after the pregnancy as placebos, because noise exposure after birth can by definition not affect health at birth, and noise exposure prior to the pregnancy is unlikely to affect health at birth.

4.4 Estimating Equations

An ideal dataset would allow me to observe hourly noise at every location in California. I could then implement an instrumental variables research design where I instrument for in utero noise using downwind status as defined in Equation (1). However, my noise measurement is spatially and temporally sparse relative to the universe of California births between 1999 and 2021. Figure 1 illustrates the spatial sparsity of seismometers and associated challenges: assigning measured noise to nearby locations would require me to either apply a model of noise transport or make arbitrary assumptions about the radius for which a monitor's noise measurements apply. Appendix Table 1 illustrates the temporal sparsity: many seismometers do not enter the record until the later part of the vital statistics sample, and do not record continuously for more than a few months.

4.4.1 Main Equation of Interest

Due to this data sparsity, I focus on the reduced form effect of the electric rail noise exposure proxy T on measures of infant health Y by running regressions of the form:

$$Y_{it} = \beta^Y T_{it} + X_{it}\gamma^X + W_{it}\gamma^W + AQ_{it}\gamma^{AQ} + \lambda_{y\times m} + \mu_b + \varepsilon_{it}. \tag{2}$$

In addition to the main regressor of interest T_{it} defined in Equation (1), the share of the pregnancy spent downwind of an electric rail line, Equation (2) includes several control variables and fixed effects: X is a matrix of demographic controls consisting of indicator variables for birth order, infant sex, and maternal race, age, and education. I include these

controls primarily to improve the precision of $\hat{\beta}^Y$. The matrix W consists of controls for in utero weather exposure, including binned counts of daily maximum temperature and means of fourth-order polynomials in daily precipitation and hourly wind speed. These controls account for the possibility that downwind status is correlated with weather conditions, which may also have a direct effect of infant health outcomes. I also control directly for remotely-sensed monthly average $PM_{2.5}$ concentration AQ. I include fixed effects at the year by month ($\lambda_{y \times m}$) level, which account for seasonality, trends, and shared time-varying shocks across the whole sample, and at the census block level (μ_b), which account for time-invariant factors correlated with health outcomes and downwind status. I show robustness to the exclusion of all covariates as well as to different levels of spatial and temporal fixed effects.

The research design leverages changes in electric rail noise exposure due to within-census block variation in downwind status across time, conditional on year-by-month fixed effects. Appendix Figure 17 shows example variation for a single location, Appendix Figure 18 summarizes residualized variation for the whole sample, and Appendix D.3 discusses meteorological drivers of variation in wind direction.

4.4.2 Tests of Internal Validity

Equation (2) provides an unbiased estimate of β^Y if downwind status is orthogonal to the error term ε, conditional on the fixed effects and controls. To test this assumption, I run a version of equation (2) with maternal characteristics on the left-hand-side:

$$x_{it} = \beta^X T_{it} + \lambda_{y \times m} + \mu_h + \varepsilon_{it}. \tag{3}$$

I focus on a GLS-weighted average of indicators for maternal race, age, and education, as well as infant sex and birth order. The appendix reports results for each variable separately. A statistically significant or economically meaningful estimate of $\hat{\beta}^X$ would indicate that my measure of noise exposure is correlated with a maternal characteristic, suggesting that estimates of the main specification in Equation (2) suffer from omitted variable bias if they do not adequately control for these characteristics.

As a second test of orthogonality between downwind status and the error term, I report the results of regressions that include exposure after birth. If T_{it} measures changes in in utero noise exposure, versions of T_{it} that capture noise exposure after the pregnancy should be uncorrelated with outcomes at birth.[15]

Interpreting β^Y from Equation (2) as the effect of in utero noise exposure on infant health

[15]Exposure prior to the start of the pregnancy could have an impact on fetal health through its impact on the mother's health. I test for this possibility by including exposure before the beginning of the pregnancy.

requires an exclusion restriction: downwind status affects health at birth only through the increase in noise exposure it causes. I test this restriction in two ways. I additionally implement placebo tests and model heterogeneity in the effect of downwind status on health to elucidate patterns that further probe this assumption.

First, I estimate the relationship between downwind status and measured noise using hourly data on noise levels and the wind direction:

$$\text{Noise}_{jh} = \beta^N D_{jh} + W_{jh}\gamma^N + \lambda_{y\times m} + \mu_j + \varepsilon_{ih}, \tag{4}$$

where j indexes seismometers, h indexes hours, and all other variables are defined as in Equation (2). Equation (4) provides a test of whether downwind status affects noise exposure.

Second, I estimate the relationship between status downwind of an electric rail line and contemporaneous air quality using daily air quality data from EPA monitors:

$$\text{AQ}_{kd} = \beta^A D_{kd} + W_{kh}\gamma^A + \lambda_{y\times m} + \mu_k + \varepsilon_{kh}, \tag{5}$$

where k indexes EPA air quality monitors, d indexes days, and all other variables are defined as in Equation (2). I run versions of Equation (5) that use $PM_{2.5}$, Carbon Monoxide, and the EPA's Air Quality Index (AQI) as measures of air quality. Equation (5) provides a test of whether status downwind of an electric rail line—which itself does not produce air pollution but may be co-located or correlated with air pollution sources—is correlated with changes in air quality.

Several heterogeneity analyses and an additional placebo test further probe the assumption that status downwind of an electric rail line affects health at birth only through noise. First, I estimate versions of Equations (2) and (4) that allow the coefficients on downwind status, β^Y and β^N respectively, to vary by distance to the electric rail line. Similar patterns across the coefficients would suggest that noise indeed drives the effect of noise on health, whereas diverging patterns would suggest an omitted variable. Second, I allow the effect of downwind status on health to vary by time of day. Larger impacts of nighttime noise would be consistent with the medical literature on noise causing sleep disruption, and would help rule out impacts due to worse air quality, which is generally better overnight (LaDochy and Witiw, 2023). Finally, I include placebo exposure, which includes hours when maternal addresses are downwind of an electric rail line, but no trains are operating on that line. During these times, downwind status should have no impact on noise exposure, since the noise source—the train's movement over the rail line—is not present. Thus, under the assumption that downwind status affects health only through noise, there should be no correlation

between this placebo exposure measure and health outcomes.

4.4.3 Interpreting Magnitudes Using Two-Sample Instrumental Variables

While estimating β^Y from Equation (2) provides evidence on whether noise has a non-zero effect on infant health, a more policy relevant parameter is the effect of an additional decibel of average in utero noise exposure on health at birth. To estimate this parameter, I turn to two-sample instrumental variables (TSIV) methods (Angrist and Krueger, 1992). In particular, I use the estimated coefficient $\hat{\beta}^N$ from the regression of hourly noise on hourly downwind status (Equation (4)) to rescale the $\hat{\beta}^Y$ from Equation (2). The ratio $\frac{\beta^Y}{\beta^N}$ is the impact of an additional decibel in average in utero noise exposure on health at birth.

This interpretation requires several additional assumptions. First, standard instrumental variables assumptions about instrument relevance, exogeneity, and monotonicity are required. Second, I assume that the contribution of downwind status to total noise over a period T is additively separable in time, so that the share of hours spent downwind of an electric rail line is a sufficient statistic for the total increase in average noise over the period:

$$\sum_{h \in T} \text{Noise}_{jh} = \beta^N \sum_{h \in T} D_{jh}.$$

Finally, I assume structural invariance (Zhao et al., 2019): the parameter $\hat{\beta}^N$ I estimate in the sample where I observe noise is the same as the parameter I would estimate if I measure noise exposure for every birth in the sample I use to estimate the effect of downwind status on health. I use a block bootstrap to estimate the standard errors on the ratio of parameters $\frac{\hat{\beta}^Y}{\hat{\beta}^N}$.

5 Results

5.1 Validating the Research Design

Table 1 summarizes the results of the three key tests of internal validity discussed in Section 4.4.2. Column (1) corresponds to the balance test described in Equation (3), which regresses a GLS-weighted index of fixed maternal and infant demographic characteristics on the in utero noise exposure measure. The index consists of indicators for whether the mother is Black, Hispanic, under 20 years old, has less than a high school education, an indicator for whether the infant is female, and birth order. Appendix Table 3 reports results of regressing each characteristic on the exposure measure separately.

I do not find evidence that households differ systematically as a function of wind-driven

Table 1: Tests of Internal Validity

	Demographic index	Noise		$PM_{2.5}$	CO	AQI
	(1)	(2)	(3)	(4)	(5)	(6)
Noise exposure	0.011	2.426***	1.371***	-0.254	-0.020	-0.948
	(0.011)	(0.457)	(0.524)	(1.700)	(0.076)	(4.950)
Dep. var. mean	0.00	50.86	50.77	12.01	0.70	44.21
Temporal aggregation	Full pregnancy	Hourly	Daily	Daily	Daily	Daily
Observations	1,960,153	1,041,029	48,022	126,105	142,700	126,105
R^2	0.32	0.93	0.95	0.34	0.62	0.40

Notes: The table shows the results of three tests of internal validity. Column (1) regresses a demographic index on the in utero noise exposure measure. The demographic index is a GLS-weighted average of standardized indicators for whether the mother is Black, Hispanic, under 20 years old, or has less than a high school education, an indicator for whether the infant is female, and birth order. Column (2) regresses hourly noise on an hourly downwind indicator. Columns (3) through (4) regress daily noise, $PM_{2.5}$, carbon monoxide (CO), and the air quality index (AQI) on the daily share of hours downwind. All specifications include year by month fixed effects. Column (1) includes census block fixed effects, while Columns (2) through (6) include monitor fixed effects. Columns (2) through (6) also include fourth-order polynomials in daily temperature, precipitation, and wind speed. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

changes in noise exposure. The coefficient in column (1) of Table 1 shows that going from no excess downwind noise exposure to spending the full pregnancy downwind is correlated with a change in the demographic index of 1.1% of a standard deviation. The coefficient is not statistically different from zero at conventional levels. Nonetheless, I control nonparametrically for maternal characteristics in my preferred specification.

I now turn to tests that probe my interpretation of the coefficient of interest in Equation (2) as the effect of noise on health at birth. Columns (2) and (3) of Table 1 report the results of estimating versions of Equation (4). These columns show that seismometers register more noise on hours and days when they are downwind of an electric rail line. During hours when a seismometer is downwind of an electric rail line (column 2), it is 2.4 dB louder than during an hour when it is not downwind. This effect corresponds to a small but perceptible increase in loudness, and is statistically significant at the 1% level. One way to interpret the magnitude of this coefficient is in terms of variation in average noise levels across days of the week as reported in Figure 4: the increase in noise in a downwind hour is comparable to the difference in typical loudness between 6 a.m. and noon, or between an average Sunday and an average Tuesday. Aggregating to the daily level yields slightly attenuated but otherwise similar results: on a day when a seismometer is downwind of an electric rail line for every hour in the day, it is 1.4 dB louder. This estimate is also statistically

significant at the 1% level.

Finally, columns (4) through (6) of Table 1 report the results of estimating versions of Equation (5), which regresses measures of air quality on an indicator for status downwind of an electric rail line. I find no relationship between downwind status and daily measures of air quality. Point estimates on the AQI, the ambient concentration of $PM_{2.5}$ and carbon monoxide are negative and not statistically distinguishable from zero.

5.2 The Effect of Noise on Infant Health

5.2.1 Average Treatment Effects

Table 2: The Effect of Noise on Infant Health

	Infant health index			
	(1)	(2)	(3)	(4)
Noise exposure	-0.046***	-0.048***	-0.042***	-0.041***
	(0.014)	(0.013)	(0.013)	(0.013)
Maternal characteristics		✓	✓	✓
Weather			✓	✓
PM 2.5				✓
Observations	1,960,153	1,960,153	1,960,153	1,960,153

Notes: the table shows the results of estimating equation (2). The dependent variable for all regressions is the infant health index, which is composed of birth weight, an indicator for low birth weight (< 2500 g), an indicator for very low birth weight (< 1500 g), an indicator for preterm birth (< 37 weeks gestation), an indicator for very preterm birth (< 34 weeks gestation), an indicator for whether the infant died, and 5- and 10-minute Apgar scores. Column (1) includes only year by month and census block fixed effects. Column (2) adds controls for maternal age decade bins, race indicators, and educational attainment indicators. Column (3) add controls for fourth-order temperature, precipitation, wind speed polynomials. Column (4) adds a control for monthly $PM_{2.5}$ concentrations. The unit of observation for all regressions is a birth in California to a mother living within 5 km of an electric rail line, 2007-2021. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Table 2 reports the results of estimating equation (2) for a variety of specifications. Column (1) includes only year by month and census block fixed effects. Columns (2) through (4) add increasingly stringent controls: mother age decade bins, race indicators, and educational attainment indicators; controls for temperature, precipitation, and wind speed; and finally a control for $PM_{2.5}$. Across these specifications, the point estimate for the coefficient on noise exposure is stable between -0.04 and -0.05. The coefficient is statistically significant at the 1% level across all specifications. Going forward, I use the most stringent specification,

reported in column (4), as my preferred specification.

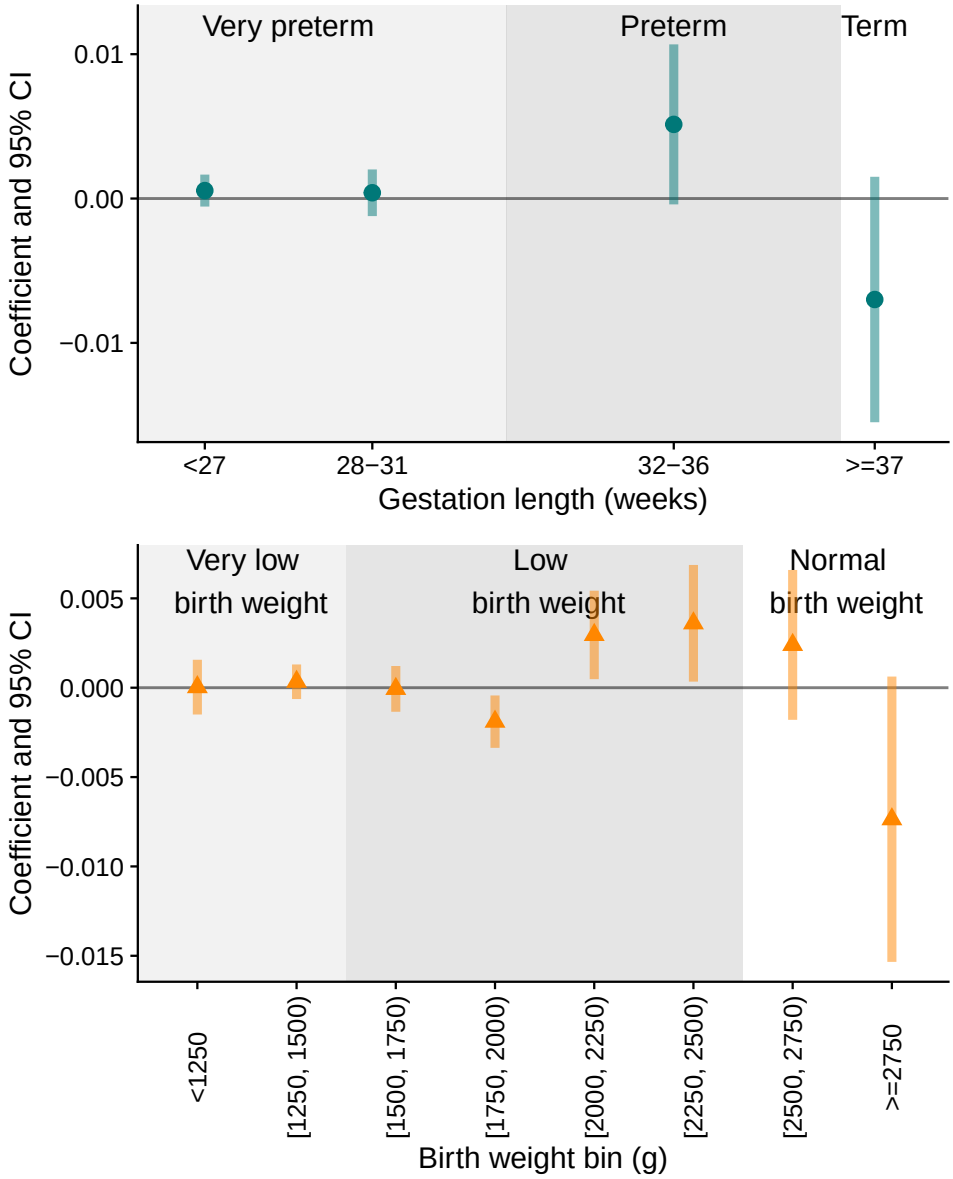
The infant health index is a weighted average of standardized variables, and can therefore be interpreted in terms of standard deviations. For example, Table 2 column (4) shows that going from no additional in utero noise exposure to spending the full pregnancy exposed to elevated noise causes a 4.1% of a standard deviation reduction in the infant health index. To help give further context for this number, a difference of 0.041 in the infant health index is about one-third of the gap in the average infant health index value for infants born to White mothers compared to Black mothers.

To help interpret these results, Appendix Table 4 reports the results of estimating equation 2 separately for each component of the infant health index. The point estimates suggest that noise exposure has a negative effect on all components of the infant health index. While statistical significance should be interpreted cautiously as they are not adjusted for multiple hypothesis testing, the precision of these effects is significantly lower than the overall infant health index, with the absolute value of the t-statistics ranging from 0.58 to 2.5, compared to an absolute t-statistic of 3.3 on the infant health index. These results are consistent with the notion that noise exposure marginally affects a range of measures of health, and validates the focus on a summary index as the main outcome of interest, as in addition to guarding against false discoveries, using the infant health index maximizes statistical power (Anderson, 2008).

While differences in units complicate rigorous comparisons of the infant health effects of noise and other environmental and socioeconomic exposures, rough comparisons help build intuition for the magnitudes reported here. The 7 percent increase in the probability of low birth weight I estimate from a 2 decibel increase in noise exposure in utero is roughly 60 percent of the effect of air pollution improvements due to the adoption of E-ZPass studied by Currie and Walker (2011), and is similar in magnitude to the effect of Safe Drinking Water Act violations studied by Currie et al. (2013). Beyond environmental exposures, the effect I estimate is roughly equivalent to a post-tax transfer of $2,500 (Hoynes et al., 2015). I emphasize that the impacts on individual outcomes, as well as costs, remain fairly uncertain, suggesting additional caution in these comparisons.

To characterize how elevated exposure to electric rail noise affects the distribution of birth outcomes, I estimate regressions with bins of birth weight and pregnancy length as the dependent variables. Figure 5 reports the results of these regressions. Noise exposure shifts birth weights and gestation lengths downward from the middle and upper ends of their respective distributions, increasing the probability that infants are born preterm and with low birth weights, but having little or no effect on the likelihood that infants are born at the lowest end of the distributions, where consequences on later health are most severe.

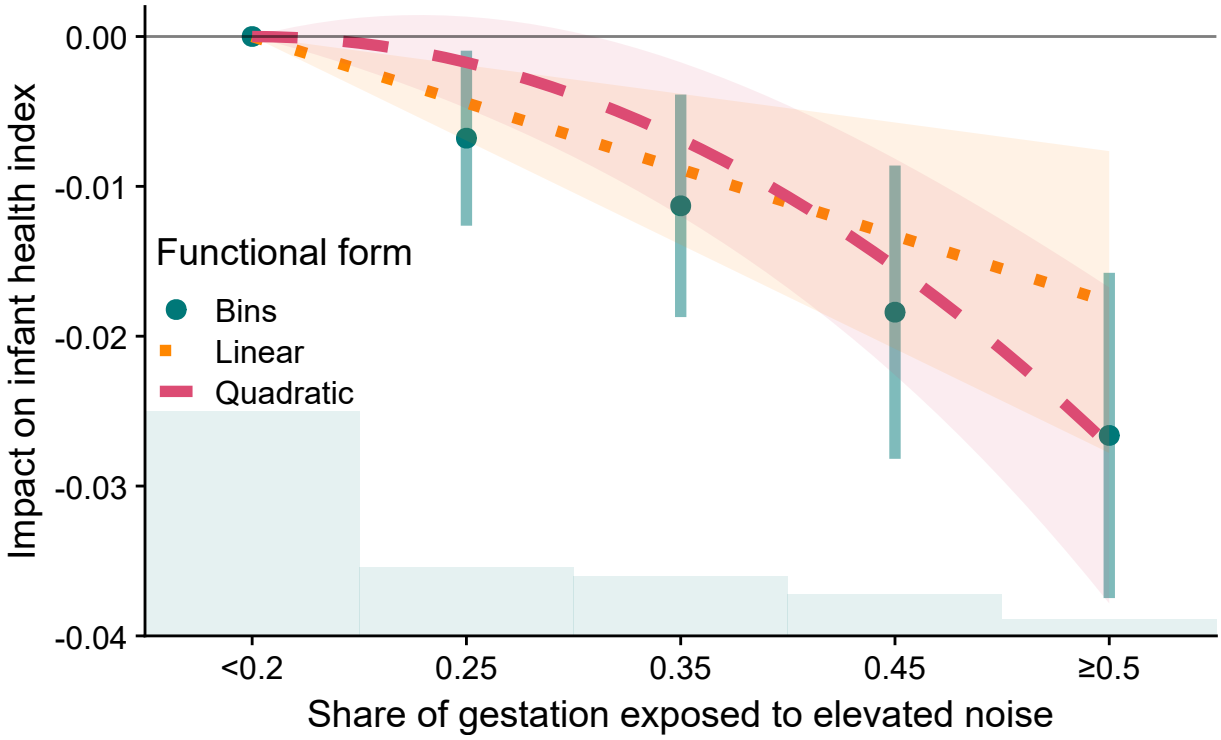
Figure 5: Noise Impacts on Birth Weight and Gestation Length Distributions



Notes: Each point estimate and Conley confidence interval is the coefficient on the noise exposure measure from a separate regression, regressing the noise proxy on the corresponding binned value of birth weight (upper panel) and gestation length (lower panel). All regressions control for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

5.2.2 Treatment Effect Heterogeneity

Figure 6: The Effect of Noise on Infant Health by Exposure Level

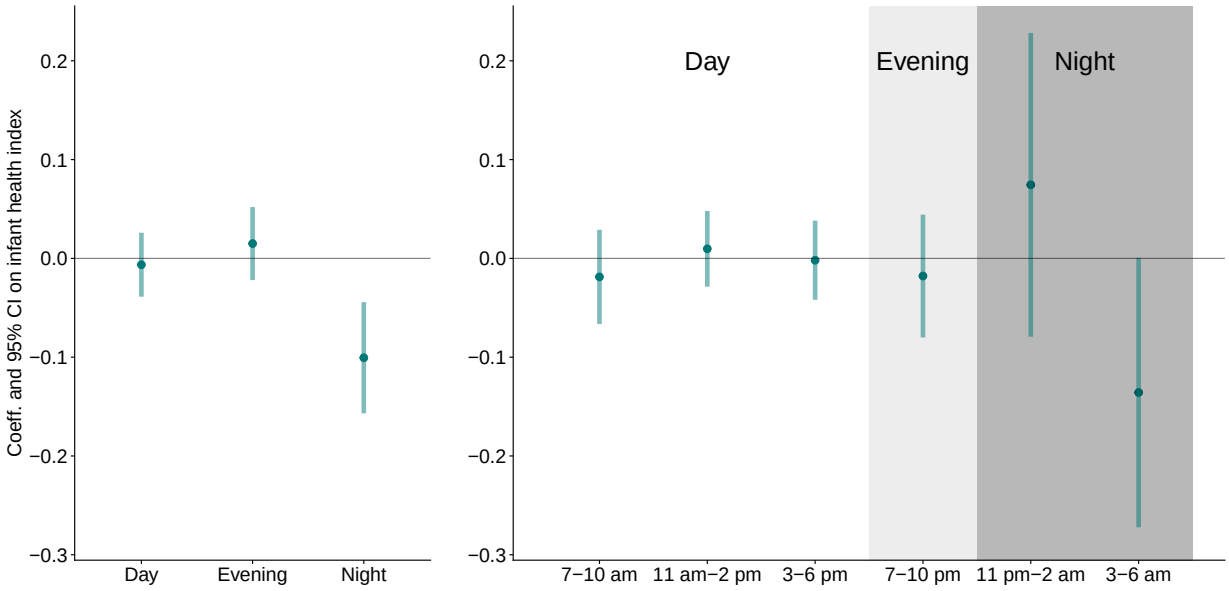


Notes: The figure shows coefficients and 95% confidence intervals from three regressions investigating non-linearity in the impact of noise on the infant health index. The teal dots correspond to 10 percentage point bins. The distribution of observations in each bin is shown at the bottom of the plot. The dotted orange line models exposure linearly. The dashed pink line shows a quadratic in exposure. All coefficients are plotted relative to $< 20\%$ of the gestation exposed to elevated noise. All regressions controls for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

Health damages from noise pollution may be nonlinear in exposure. Figure 6 investigates this possibility by plotting response functions for binned, linear, and quadratic specifications in the share of the gestation exposed to elevated noise. The dotted orange line reports results for the linear specification imposed by equation (2). The other specifications allow for nonlinearity: the dashed pink line is quadratic in exposure, and the teal dots are 10 percentage point bins in exposure. The binned distribution of exposure is shown at the bottom of the plot. Each bin corresponds to roughly one additional month of the pregnancy during which the mother is exposed to elevated noise. The figure reveals that the negative effect of noise pollution on health is increasing in exposure.

Noise exposure may have heterogeneous impacts at different times of day. In particular, noise may be more harmful at nighttime, when it is more likely to cause sleep disturbances.

Figure 7: Effect of Noise On Infant Health by Time of Day



Notes: The figure shows coefficients and 95% confidence intervals from regressions of the infant health index on the share of hours exposed to elevated noise at different times during pregnancy. The left panel shows effects by period of the day. The definitions of these periods follow the noise monitoring literature: daytime (7:00 a.m. to 6:59 p.m), evening (7:00 p.m. to 10:59 p.m.), and night (11:00 p.m. to 6:59 a.m.) (top left panel). The left panel shows effects by four-hour periods. All regressions control for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

To investigate this, I calculate the share of pregnancy hours exposed to elevated noise during three periods of the day: daytime, between 7:00 a.m. and 6:59 p.m.; evening, between 7:00 p.m. and 10:59 p.m.; and nighttime, between 11:00 p.m. and 6:59 a.m. The definitions of these periods follow the standard definitions in the noise monitoring literature (Murphy and King, 2014). I then estimate a version of Equation (2) that allows for different treatment effects for each period of the day. The left panel of Figure 7 describes the results of this regression. Nighttime exposure has a large, negative, and precisely estimated effect on the infant health index. In contrast, the effect of daytime and evening exposure is an order of magnitude smaller and not statistically distinguishable from zero. The right panel of Figure 7 elaborates on this result by estimating the impacts of exposure in each of six four-hour blocks. Noise exposure in the early morning, between 3 am and 7 am, has a large and statistically significant negative effect on infant health, whereas exposure during all other periods of the day have a small and statistically imprecise effect on health. These results are consistent with the medical literature finding that noise causes sleep disruptions which can compound with chronic exposure (Munzel et al., 2014). They also provide further evidence

that my results are not driven by air pollution: air pollution tends to be lowest during the hours I find my noise exposure measure has the largest negative impact on health (LaDochy and Witiw, 2023).

An alternative explanation for the finding that nighttime noise exposure is more detrimental to health is that there is more measurement error in my estimates of daytime and evening exposure. I observe home addresses, but not other locations that mothers may frequent, such as their work addresses. All exposure is therefore measured with respect to the mother's home address, implicitly assuming that mothers are always at home. This assumption is more likely to be satisfied at nighttime, when mothers are likely sleeping in their homes. In contrast, it is less likely to be satisfied during the day, when mothers may leave home. It could be that daytime exposure is harmful but mismeasured, and thus the estimates in Figure 7 are biased toward zero.

To investigate this possibility, Appendix Figure 3 reports the results of a regression that interacts exposure by period of day with and indicator for whether the mother worked during pregnancy. The intuition behind this exercise is that a person who is working is more likely to leave home during the day, leading to more measurement error in the exposure measure. Conversely, someone who is not employed may be more likely to stay at home during the day, and thus would have less measurement error in their exposure measure.[16]

I do not find evidence that daytime noise exposure is harmful for mothers who did not work during their pregnancy. While the point estimates for working mothers are smaller in magnitude than the point estimates for non-working mothers for all periods, the qualitative patterns are the same across both groups, and the coefficients are not statistically different from one another.

Appendix Figures 5 to 8 examine additional dimensions of treatment effect heterogeneity. Appendix Figure 5 reports coefficients from regressions of the infant health index on nighttime noise exposure interacted with one-kilometer bins of distance to the nearest electric rail line. The pattern of effects mirrors the downwind noise gradient in Appendix Figure 2: estimates are U-shaped in distance, with the largest and most precisely estimated effect in the two- to three-kilometer bin. The four- to five-kilometer bin yields the largest point estimate but a wide confidence interval that includes zero. Excluding this bin from the estimation sample (Appendix Table 5, column (3)) leaves the main results unchanged.

Appendix Figures 6 and 7 examine heterogeneity in the estimated effects of nighttime noise exposure across trimesters of pregnancy and by maternal demographics. Appendix

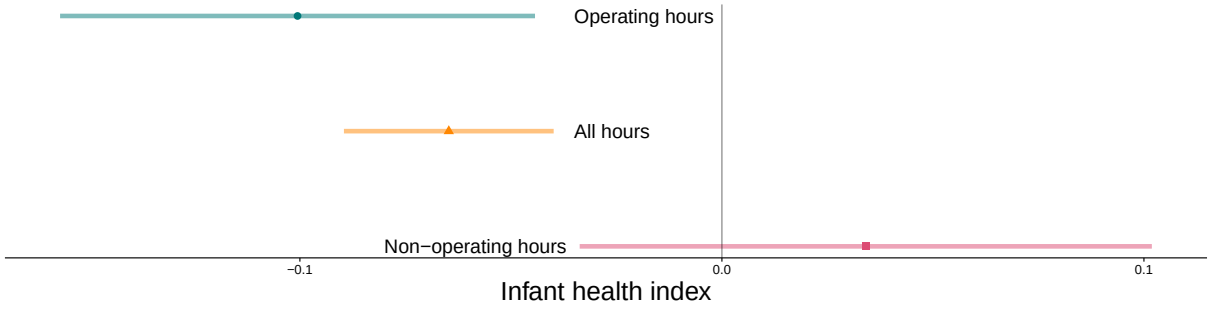
[16] Appendix Figure 4 uses data from the American Time Use Survey to show that urban California women aged 15-50 who are not employed are more likely to be at home than employed women at all hours of the day. This difference is up to 30 percentage points at noon. Over 98% of respondents are at their home address during overnight hours, regardless of employment status.

Figure 6 shows that effects are concentrated in the first and second trimesters, consistent with greater sensitivity of early fetal development to environmental stressors. I find no evidence of heterogeneity according to maternal race, ethnicity, or education (Appendix Figure 7).

Finally, Appendix Figure 8 examines heterogeneity by the age of the mother's home. An important policy question is what factors mediate the relationship between noise exposure and health outcomes. Building quality—particularly the presence of double-pane windows—may attenuate indoor noise levels and thus reduce exposure (Medved, 2022). Although window type is difficult to observe directly, it is strongly correlated with building age (Appendix Figure 9). Appendix Figure 8 shows that the estimated effects of nighttime noise are smaller for mothers living in newer homes, especially those built since 2010. While unobserved factors such as income may contribute to this pattern, the results suggest that construction characteristics may mediate the effect of environmental noise on infant health.

5.2.3 Placebo Exposure and Robustness

Figure 8: Placebo Nighttime Exposure



Notes: The figure reports coefficients and 95% Conley confidence intervals from regressions of the infant health index on actual and placebo nighttime noise exposure. The teal estimate reproduces the main nighttime coefficient from the left panel of Figure 7, which correctly accounts for rail system operating hours when assigning exposure. The orange estimate ignores operating schedules and pools both operating and non-operating hours. The pink estimate separately regresses the infant health index on the placebo exposure measure, which considers only hours when the rail system is not operating. All regressions also include daytime and evening exposure and control for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

None of the rail systems in my sample operate 24 hours a day; each suspends service for between two and seven hours overnight, depending on the system and day of the week. This pattern creates a natural placebo test: during hours when no trains are running, being downwind of an electric rail line should not affect noise exposure and therefore should have no influence on infant health. I implement this test by regressing the infant health index

on (i) the actual nighttime noise exposure, and (ii) the share of non-operating hours during pregnancy when the mother's residence was downwind of an electric rail line. The latter is the placebo exposure measure.[17]

The pink coefficient in Figure 8 shows that being downwind of a rail line during non-operating hours has no association with infant health, as expected. For comparison, the figure also reproduces the nighttime coefficient from Figure 7 in teal and the coefficient on a noise exposure measure that does not account for operating hours and pools across both operating and non-operating hours in orange. The pooled coefficient is attenuated toward zero, consistent with measurement error introduced by failing to distinguish between operating and non-operating hours.

Appendix Tables 5 to 9 report the robustness of the main specification to changes in the sample, downwind variable definition, spatial and temporal fixed effects choices, and standard error calculation, respectively. Appendix Table 10 reports specifications that include mother fixed effects, identifying the effect of noise on infant health by comparing infants born to the same mothers with different in utero noise exposures. Appendix Section E discusses these results.

5.3 The Social Cost of Noise Pollution

The finding that noise pollution harms health at birth suggests that noise pollution may have significant social costs. This section calculates the social cost of noise pollution due to the harms it causes to birth outcomes. Other potential costs of noise, for example due to lost productivity, impacts to adult health (beyond maternal health costs), or disamenity are not considered; the cost estimates presented in this section should thus be interpreted as representing the partial social cost of noise pollution.

Analogously to the infant health index, I define a cost index that accounts for the social cost of adverse birth outcomes for each birth in my sample. I consider four categories of costs: infant and maternal healthcare costs, costs due to elevated infant mortality risk, and reductions to expected lifetime income. Healthcare costs come from tabulations of inpatient and outpatient medical care costs by gestation length summarized in Behrman and Butler (2007). Lifecycle income costs come from Bharadwaj et al. (2018), who use Swedish administrative records to match health at birth to individuals' lifetime labor market histories. Because healthcare and lifetime incomes costs are highly convex in the severity of outcomes, I use estimates of healthcare costs by 4-week bins of gestation length and lifetime

[17] Both specifications also include daytime and evening exposure controls. A similar placebo test cannot be implemented for those hours because, with rare exceptions, all rail systems operate continuously during the day and evening.

Table 3: TSIV Estimates of the Costs of Noise Pollution

Outcome:	Pregnancy-level cost index ($)		Hourly noise (dB)
	Reduced form	TSIV	First stage
	(1)	(2)	(3)
Share of pregnancy downwind	7,816.74*		
	(4,593.03)		
Noise (dB)		2,786.17*	
		(1,648.32)	
Hourly downwind indicator			2.87***
			(0.37)
Hourly F-statistic			60.00
Observations	1,960,153	1,960,153	1,041,209

Notes: Columns (1)–(3) report reduced-form, TSIV, and first-stage estimates, respectively. The dependent variable in columns (1) and (2) is a birth-level cost index combining expected maternal and infant healthcare costs and lifetime income losses based on birth weight and gestation length. In cases of infant mortality, the cost index equals the value of a statistical life (VSL) plus expected healthcare costs. All values are 2025 dollars. Column (1) reports the reduced-form estimate corresponding to Equation (2) with the cost index as the outcome. Column (2) reports the TSIV estimate, obtained by rescaling the reduced-form coefficient in column (1) by the hourly first stage in column (3). All specifications include census block and year by month fixed effects and fourth-order polynomials in temperature, precipitation, and wind speed. Standard errors, clustered at the census tract level, are shown in parentheses. TSIV standard errors in column (2) are estimated using a census-tract block bootstrap with 5,000 replications. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

income costs by 500-gram bins of birth weight. I value changes in mortality risk using the EPA's Value of a Statistical Life (VSL) (EPA, 2014). Appendix Section F further discusses the cost estimates used.

For each birth, the cost index is the sum of the expected costs according to the infant's birth weight and gestation length. If the infant dies, the social cost is the VSL plus expected medical care costs according to the infant's gestation length. Similarly to the infant health index, this approach accounts for covariance between outcomes. It also avoids double-counting: if an infant dies, there is no additional cost due to the infant's expected reduction in lifetime income.

Table 3 summarizes the results of regressing the cost index on my noise exposure measure, fixed effects, and controls, as described in Equation (2). The table also includes TSIV estimates of the expected cost per decibel of increased average in utero noise exposure. Column (1), the reduced form, shows that going form no additional in utero noise exposure to spending the full pregnancy exposed to noise creates an expected social cost of $7,816.74 per infant. Re-scaling this value by the hourly first stage reported in column (3) yields an

expected cost of $2,786.17 per decibel increase in average in utero noise exposure, as reported in column (2).

Assessing first-stage strength is central in instrumental variables analyses and is typically done using the first-stage F-statistic. In this setting, the endogenous regressor, average noise exposure over the pregnancy, is unobserved, making it infeasible to compute a pregnancy-level first-stage F-statistic. I therefore assess instrument strength using two complementary diagnostics. First, the F-statistic in the hourly first-stage regression is 60, indicating a strong relationship between wind direction and contemporaneous noise exposure. While this statistic does not directly characterize instrument strength at the pregnancy level due to aggregation, it establishes a strong relationship between the instrument and noise exposure at the hourly level. Second, I examine the finite-sample distribution of the TSIV estimator using a bootstrap. For ratio estimators such as TSIV, weak identification can manifest as a non-normal sampling distribution of the estimator (e.g., Nelson and Startz, 1990; Staiger and Stock, 1997). Accordingly, comparing the bootstrap distribution of the estimator to a normal reference distribution is informative about weak identification concerns (Zhan, 2018; Andrews and Shapiro, 2024). Appendix Figure 10 shows that the bootstrap distribution is approximately symmetric and closely matches a normal distribution with the same mean and standard deviation. A Kolmogorov-Smirnov test fails to reject normality ($p = 0.17$), providing further evidence against weak identification concerns.

Table 4: The Partial Social Cost of Electric Rail Noise in California

	Annual cost (1)
All CA electric rail	$136,600,000 ($3,662,000, $269,500,000)
Per passenger-trip	$1.11 ($0.03, $2.18)
Per overnight passenger-trip	$14.07 ($0.38, $27.76)

Notes: The table shows estimated costs of electric rail in California. All values are 2025 dollars. Column (1) reports costs per infant exposed to rail noise in utero. Column (2) reports the total annual cost. The first row uses the reduced form estimate from Table 3, scaled by average in-sample exposure. The second row uses an estimate of the marginal contribution of electric rail to overall noise levels in combination with the TSIV estimate in Table 3. The third and fourth columns allocate costs across all electric rail passengers and passengers riding between 11:00 p.m. and 6:59 a.m., respectively. 90% confidence intervals clustered at the census tract level in parentheses.

Table 4 uses the estimates in Table 3 to calculate the social cost of electric rail noise in California attributable to the harm it causes to infant health outcomes. To estimate the total cost of electric rail noise in California, I combine my TSIV estimate of the expected cost of noise per decibel with an estimate of electric rail's marginal contribution to ambient noise levels from the Department of Transportation's National Transportation Noise Map. I calculate the marginal contribution of rail to overall exposure, using a baseline noise level of 45 dB, which corresponds to background noise levels in urban areas and is consistent with the modelling assumptions embedded in the National Transportation Noise Map (Murphy and King, 2014; Moretti and Wheeler, 2025). Appendix F provides further details.

The expected cost of electric rail noise per infant with positive exposure is $4,890. Scaling by the birth rate in exposed areas yields a total annual cost of $137 million.[18] Dividing this cost by the total ridership across all systems yields a per-passenger-trip cost of $1.11. While this externality is small in magnitude, it is large as a share of average fares riders pay, roughly corresponding to between 25% and 50% of average fares across the transit systems I study.

A Pigouvian tax on rail passengers would charge each passenger according to the marginal externality of their trip, fully accounting for heterogeneity in the noise externality according to trip characteristics. I find that the most important dimension of heterogeneity is the time of day at which a trip is taken: Figure 7 shows that the negative externality is driven by elevated noise during the nighttime and early morning. Transit ridership is at its lowest during these hours, comprising about 8% total ridership (Figure 19), yet the loudness of a train is not affected by the number of passengers riding it. Together, these facts imply a per-ride externality of $14.07 for passengers riding trains between 11:00 p.m. and 6:59 a.m. This external cost is large both in comparison to the private costs of alternative transportation modes and as a share of average fares.[19]

The costs reported in Table 4 do not necessarily represent welfare costs. First, several potentially important categories of costs are omitted. Examples include the costs of parents' compensatory behaviors and investments, the costs of additional societal investments in children's education, and the long-run costs of disability and reduced life expectancy. Second, these costs are likely to be a lower bound on willingness to pay to avoid poor health. On the other hand, if healthcare markets are not perfectly competitive, the costs I estimate may include markups.

[18] When calculating annual birth counts, I use data between 2015 and 2019 to obtain an estimate average fertility that reflects recent fertility trends without being biased by the temporary negative shock to fertility during the COVID-19 pandemic. Appendix F discusses this choice.

[19] Note that a $14.07 uniform tax on early morning passengers would not be optimal. The Pigouvian tax would additionally account for trip length, heterogeneity in the noise externality across different sections of track, and population densities along the trip.

5.4 Implications for Transportation Policy

Recent research on optimal transportation policy has consistently found that increasing transit fare subsidies, expanding rail networks, and increasing service frequency are welfare-improving (Parry and Small, 2009; Barwick et al., 2024; Almagro et al., 2024). These studies typically account for time-varying congestion externalities as well as constant air pollution externalities. Does accounting for a time-varying noise pollution externality change optimal transit pricing or frequency? Put another way, is the noise pollution externality I estimate sufficiently large to have quantitative implications for optimal policy?

This subsection assesses the magnitude of the noise pollution externality relative to the road congestion (hereafter, congestion) externality. I focus my comparisons on congestion because congestion externalities dominate marginal external costs of transportation (Parry and Small, 2009). In addition, marginal congestion and noise externalities both vary significantly by time of day, whereas other externalities, such as greenhouse gas emissions, do not.[20] In this section, I focus on the San Francisco Bay Area, because it is the only region in my sample for which I was able to obtain hourly data on public transit ridership for both bus and rail.

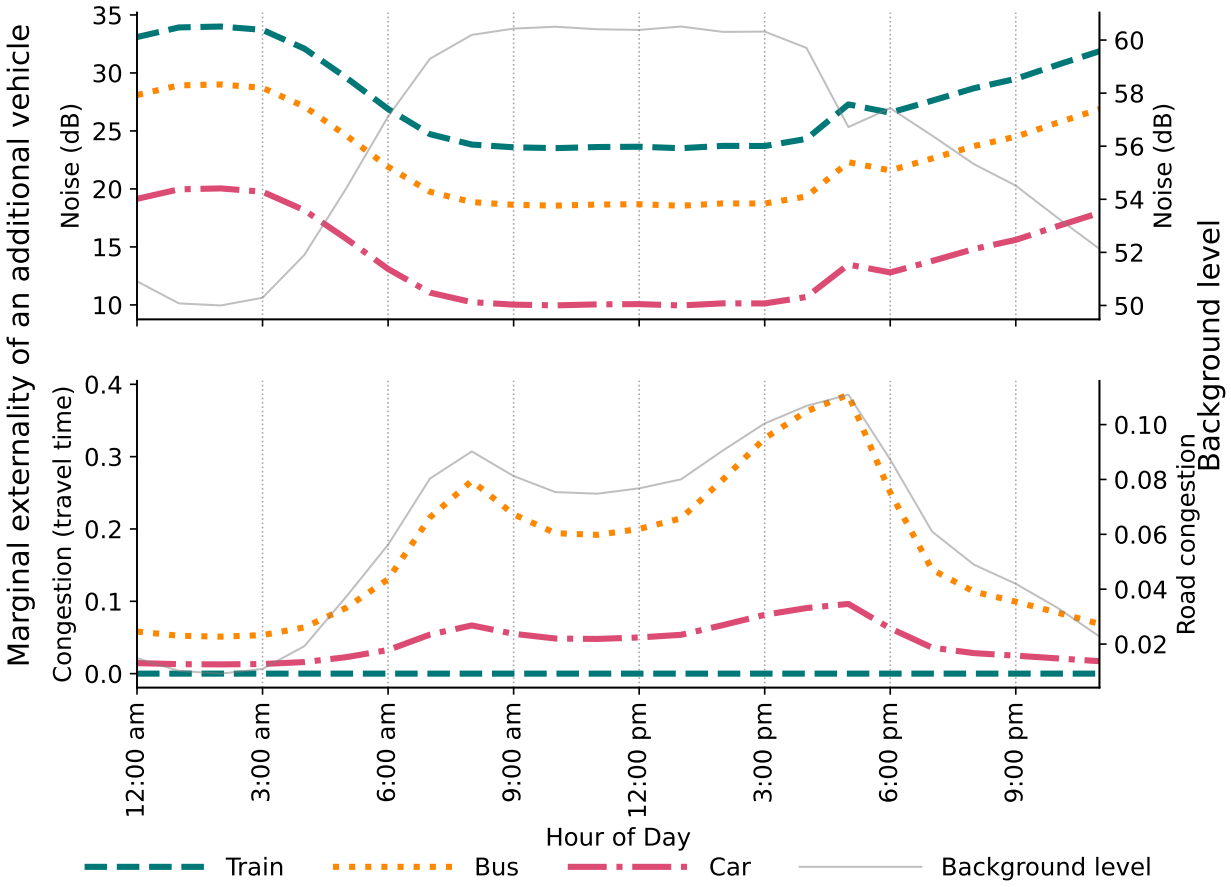
Figure 9 shows the hourly marginal noise and congestion externalities for trains, buses, and cars. Marginal noise externalities are calculated using engineering estimates of the loudness of each vehicle type at a distance of 50 feet and the background noise level, which is calculated from seismic data. Because sound levels are a logarithmic function of sound energy, the marginal contribution of each vehicle type to ambient noise is dependent on the background noise level. For example, if the background noise level is 50 dB, a passing train that is 84 dB loud will cause the ambient noise level to increase to 84 dB, an increase of 34 dB.[21] As a result, the marginal noise increase due to an additional vehicle is largest when background noise is quietest, in the middle of the night. Trains are the loudest, at 84 dB, followed by buses (79 dB), and cars (70 dB).

The bottom panel of Figure 9 shows the hourly marginal congestion externality. Similarly to noise, the marginal congestion externality depends on the background level of congestion. I estimate the relationship between road congestion and travel times using one month of

[20] A third transportation externality that may vary by time of day is air pollution. Studies have found that average levels of primary pollutants vary by hour, including spiking near roads during commute hours (Chavez and Li, 2020; LaDochy and Witiw, 2023). Both secondary pollutant formation and local air pollution dispersal are also influenced by factors such as temperature which vary within the day. However, recent empirical work has found that long-run average air pollution levels matter more for health than within-day spikes, suggesting that the air pollution externality does not significantly vary by the time of day of the emission (Berkouwer and Dean, 2025).

[21] To add sound levels, one must first convert from dB to sound energy, add the energy values, and then convert back to decibels. In this example, $10 \log_{10}(10^{50/10} + 10^{84/10}) = 84.001$ dB.

Figure 9: Marginal Noise and Congestion Externalities by Hour



Notes: The figure shows hourly marginal noise and congestion externalities for trains, buses, and cars. The top panel shows the noise externality, using engineering estimates of the noise produced by each vehicle type at a distance of 50 feet. The right y-axis shows the background noise level, which is calculated using seismic data. While the sound energy produced by each vehicle type is constant, its marginal contribution to ambient noise is dependent on the background level. The bottom panel shows the congestion externality. The background congestion level is show on the right y-axis.

data from the Caltrans Performance Measurement System (PeMS), a network of road-based sensors estimating travel times, vehicle speeds, and road congestion. Appendix Figure 11 shows the marginal congestion externality function I estimate using these data. Following Parry and Small (2009), I assume that buses cause four times as much congestion as passenger cars. Trains have no congestion externality at any time because they operate on separate rights-of-way from buses and cars. The marginal congestion externality is largest at times when roads are congested, during morning and evening rush hours. This is the opposite of the pattern for noise, which is highest at times when the background noise is low, in part due to low road congestion.

Figure 9 illustrates the within-day dynamics of marginal noise and congestion externali-

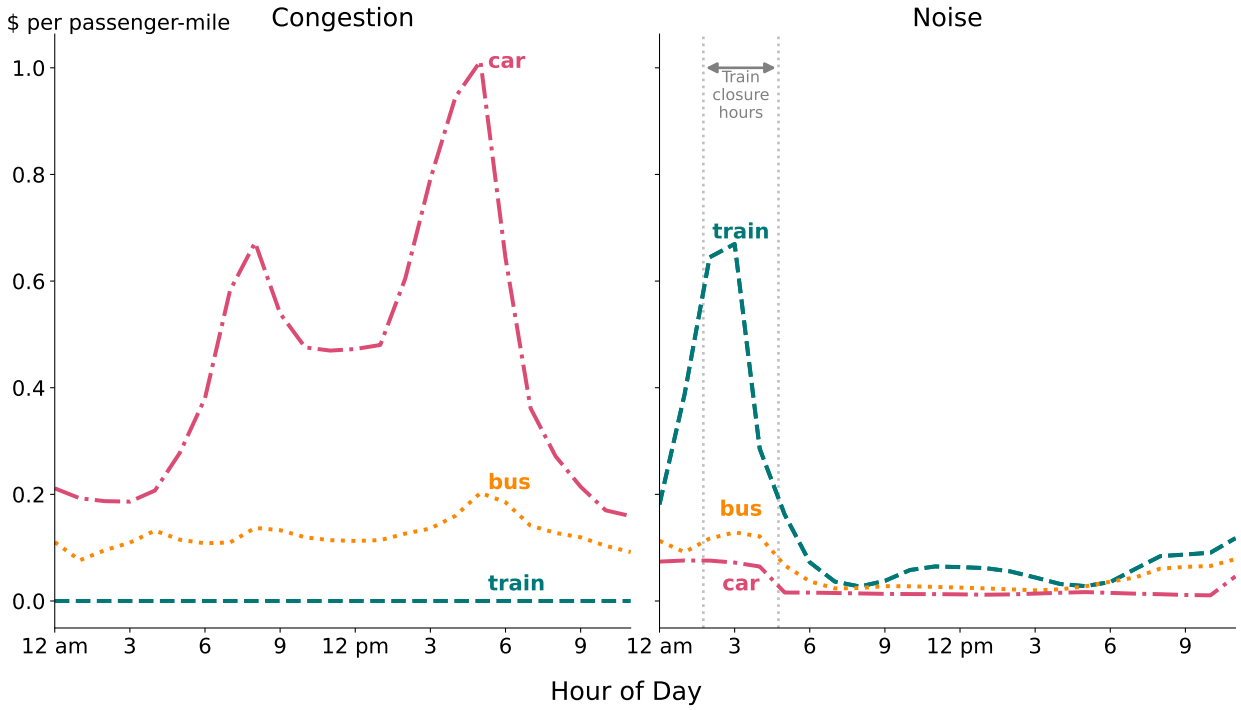
ties per vehicle. Both externalities depend on background noise and congestion, respectively, but in opposite ways: marginal noise externalities are greatest when background noise is low; marginal congestion externalities are greatest when background congestion is high. However, the policy-relevant unit for these externalities is their dollar value per passenger-mile. Calculating these values requires data on several additional variables. Estimating the (partial) noise externality per passenger-mile requires estimates of the marginal damage to infant health of each vehicle and the number of exposed expectant mothers for each mile traveled. Estimating the congestion externality per passenger-mile requires an estimate of the value of time. Both require an estimate of the number of passengers per vehicle to convert from vehicle-mile externalities to passenger-mile externalities.

I use hourly data on 2024 BART rail ridership and Muni bus ridership to estimate the number of passengers per train and bus per hour, respectively. To estimate the number of occupants per car, I use hourly data from the 2022 National Household Travel Survey. I calculate the average number of occupants per car, by hour of trip start time, for households in Pacific region Metropolitan Statistical Areas with populations over 1 million and rail. Appendix Figure 12 plots the number of passengers per hour by mode. Train and bus ridership are both bimodal, with peaks during morning and evening rush hours, and troughs in the single-digit passengers per vehicle in the overnight hours. Average train and bus ridership vary by one and two orders of magnitude, respectively, between their overnight minima and their daytime maxima. In contrast, average passenger car occupancy both varies much less, staying between 1 and 2 occupants per vehicle, and exhibits less pronounced daily peaks.

To calculate the implicit price of noise pollution per passenger-mile, I begin by estimating the additional noise exposure due to a mile of vehicle travel in each hour. I combine the marginal noise estimates reported in the top panel of Figure 9, which correspond to noise at 50 feet from the vehicle, with a simple model of the spatial decay of noise to estimate the increase in average noise levels in 100-meter bands away from the noise source. The number of 100-meter bands, and thus the total exposed population, is determined by the distance at which the noise decay model predicts elevated noise over the background level due to the marginal noise emission. I then multiply these noise increases by the per-dB costs reported in Table 4 to obtain the partial social cost of the marginal noise emission. I calculate the expected number of pregnant women in each 100-meter band away from the average mile of BART line and multiply this value by the per-infant cost estimate to obtain the marginal hourly cost of noise pollution due to an additional vehicle. For congestion, I use the average value of time estimate of \$15 reported by Almagro et al. (2024). Finally, for both noise and congestion, I divide the hourly implicit prices per vehicle by the average number of vehicle

occupants in each hour to obtain the external cost per passenger-mile.

Figure 10: Externalities per Passenger-Mile



Notes: The figure shows externalities, in 2025 dollars, per passenger-mile by hour of day.

Figure 10 displays the externalities per passenger-mile by hour for each mode. The left panel shows congestion externalities and the right panel shows noise externalities. For most of the day, between 6 a.m. and midnight, passenger cars generate the largest externalities, driven by congestion costs, and trains generate the least externalities, consistent with the prior literature. Bus externalities are generally small, always falling below car externalities, since their relatively large marginal contribution to congestion is defrayed by high ridership. However, between 1 a.m. and 5 a.m., train travel generates a largest noise externality, up to $0.80 per passenger-mile during the 4 a.m. hour.

The large overnight noise externality of rail arises for three reasons. First, trains are much louder than buses or cars. This means that the level of the exposure—and thus the social cost of that exposure—is greater at each distance from the rail line, and also that a larger total population is exposed to elevated rail noise. My noise decay model predicts elevated noise up to 800 meters from the source for trains, 500 meters for buses, and 200 meters for cars. Second, rail ridership is at its lowest during the same hours that rail noise is most damaging, leading to a large cost per passenger-mile.

6 The National Burden of Noise Pollution

Extrapolating these results to the US requires two components. The first is an estimate of the cost of noise pollution due to harms to infant health caused by an additional decibel of noise exposure. Under a strong external validity assumption, I use the estimates derived from the California sample for this purpose. The second component is an estimate of average noise exposure. This section uses seismic data and machine learning to produce such an estimate.

I calculate average noise levels for the 975 seismic monitors plotted in figure 2 for 2018, calculating daytime, nighttime, and overall noise separately. I then merge these data with embeddings from the AlphaEarth Embedding Fields (AEF) Model (Brown et al., 2025). The AEF model combines high-resolution satellite imagery from the visible spectrum, hyperspectral satellite imagery, and additional information on climate, vegetation, water availability, elevation, ruggedness, and land use into 64-dimensional embeddings at a 10-meter resolution globally. These embeddings are a geospatial foundation model which can be used to predict any outcome of interest.

I merge seismic noise data with AEF features for the same locations and year, then estimate ridge regression models predicting noise using the 64-dimensional embeddings. Table 5 reports the average out-of-sample R^2 across five cross-validation folds for various noise measures. The first column shows performance using AEF features at their native resolution, which includes information from 1.28 by 1.28 km chips centered at each location. The AEF model achieves R^2 values between 0.44 and 0.51, with strongest performance for daytime median noise and weakest performance for nighttime median noise—a perhaps unsurprising result since AEF contains primarily information from daytime imagery. The second column of table 5 shows average cross-validated out of sample performance for models that include the native AEF embeddings as well as summary statistics of all AEF values within 10 km of each point of interest. Median and nighttime median predictive performance improve modestly, by around two percentage points, while the predictive performance for average noise is unchanged and the predictive performance for daytime median noise falls by five percentage points. Due to the additional computational cost of producing the 10-kilometer summary statistics of the AEF values and their mixed effect of model performance, I proceed using only the native AEF values.

To evaluate the performance of the AEF model compared to other predictions of noise, columns (3) through (5) of Table 5 report predictive performance for several baseline models. Column (3) reports on the predictive performance of the Department of Transportation (DOT)'s National Transportation Noise Map. Column (4) reports on the predictive performance of Mennitt and Fristrup (2016)'s machine learning-based predictive model for median

Table 5: Predictive Model Out of Sample Performance

	This paper		Baselines		
Noise measure	AEF (1)	10 km AEF (2)	DOT (3)	MF (4)	Pop. Density (5)
Panel A: R^2					
Average	0.461	0.461	0.048	-	0.159
Median	0.499	0.519	-	0.414	0.197
Daytime median	0.513	0.464	-	-	0.193
Nighttime median	0.444	0.466	-	-	0.189
Panel B: Mean Absolute Error					
Average	10.262	10.262	13.964	-	13.134
Median	7.196	7.020	-	7.981	10.005
Daytime median	7.426	7.455	-	-	10.499
Nighttime median	6.821	6.624	-	-	8.919
Panel C: Root Mean Squared Error					
Average	12.997	12.997	17.201	-	16.168
Median	9.680	9.489	-	10.323	12.189
Daytime median	9.899	10.381	-	-	12.682
Nighttime median	9.253	9.072	-	-	11.098

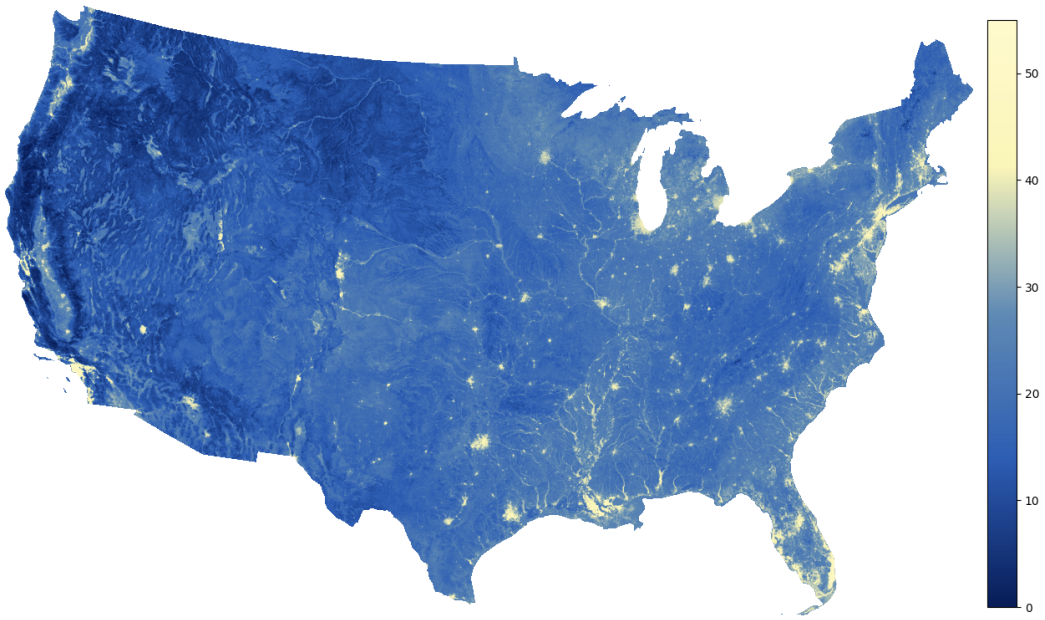
Notes: Table compares the out-of-sample predictive performance (R^2) of models developed in this paper (Columns 1-2) against benchmarks (Columns 3-5). Each entry is the average out-of-sample R^2 from 5-fold cross-validation. Column (1) uses AlphaEarth Foundations (AEF) model features at their native resolution. Column (2) adds 10 km spatial statistics of all AEF features. Column (3) uses the Department of Transportation's (DOT) National Transportation Noise Map. Column (4) uses Mennitt and Fristrup (2016) (MF). Column (5) uses population density.

noise. In both cases, to allow for the possibility that there are systematic difference between microphone-measured and seismically-measured noise, I run ridge regressions of the seismic noise values on the predicted noise values from the two noise maps and report the average out of sample predictive performance across five cross validation folds. I report performance statistics only for the noise measure predicted by each data source: average noise for the DOT map and median noise for the Mennitt and Fristrup (2016) map. The AEF model outperforms the DOT map by an order of magnitude and outperforms the Mennitt and Fristrup (2016) model by 20%.

As a third baseline comparison, Table 5 reports the out of sample predictive performance of a model that uses population density data as its only feature. While the population density mdoel outperforms the DOT map, it only achieves between 30 and 40% of the AEF

model's performance, and half Mennitt and Fristrup (2016)'s model performance. I conclude that the AEF model outperforms existing alternative measures of average noise exposure as well as a simple model based only on population density.

Figure 11: Predicted Median Nighttime Noise, Contiguous US

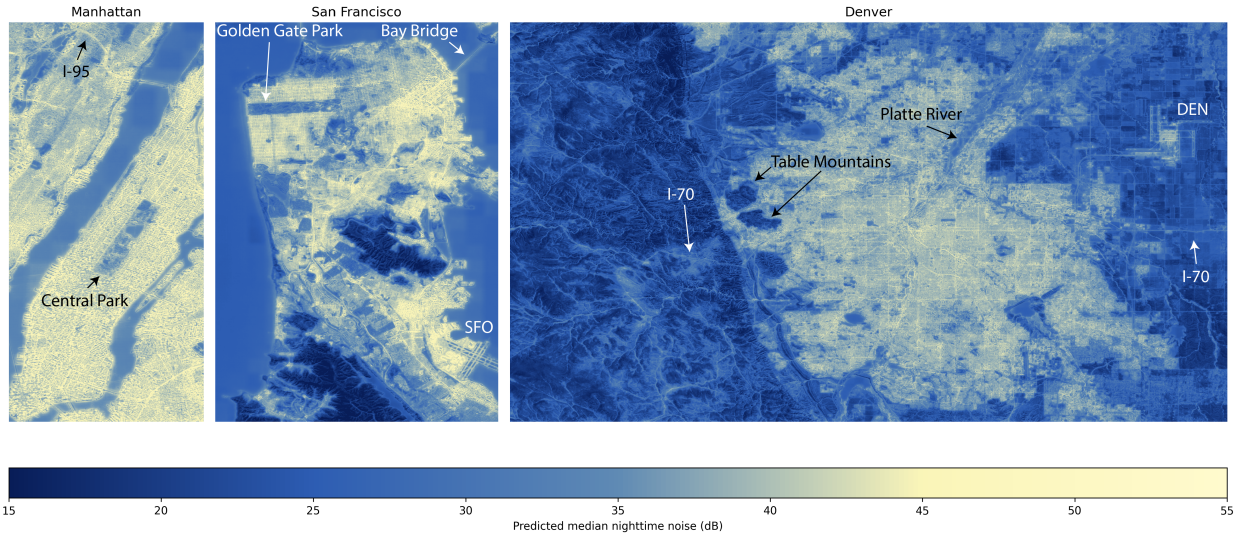


Notes: Figure shows predicted median nighttime noise in 2018 for the contiguous US at 550 m resolution.

Figure 11 plots predicted median nighttime noise levels at a 500 m resolution across the contiguous United States. The most striking and intuitive pattern from the map is that cities are clearly visible as the loudest places in the country. Valleys, such as the San Joaquin Valley in Central California, are predicted to be louder than mountainous and forested areas, likely because sound travels further in open areas. Rivers are also visible in the map, likely because roads and population centers often cluster near rivers, and because the movement of water itself can be noisy.

Figure 12 plots zoomed-in, higher-resolution 10 m predictions for Manhattan, San Francisco, and the Denver metropolitan area. Manhattan, one fo the loudest areas in the country, is almost uniformly loud. Relatively quiet areas include Central Park. The George Washington bridge, which crosses the Hudson river to connect Manhattan to New Jersey, is visible. Noise predictions for San Francisco follow similar patterns: the most densely populated urban areas are loudest, with quieter areas in Golden Gate Park, the Presidio, and San Bruno Mountain. The Bay and Golden Gate Bridges are visible, as are the runways of the San Francisco International Airport. Denver is clearly visible on the national map because of

Figure 12: Predicted Median Nighttime Noise



Notes: Figure shows predicted median nighttime noise for 2018 in Manhattan, San Francisco, and Denver at a 10 m resolution.

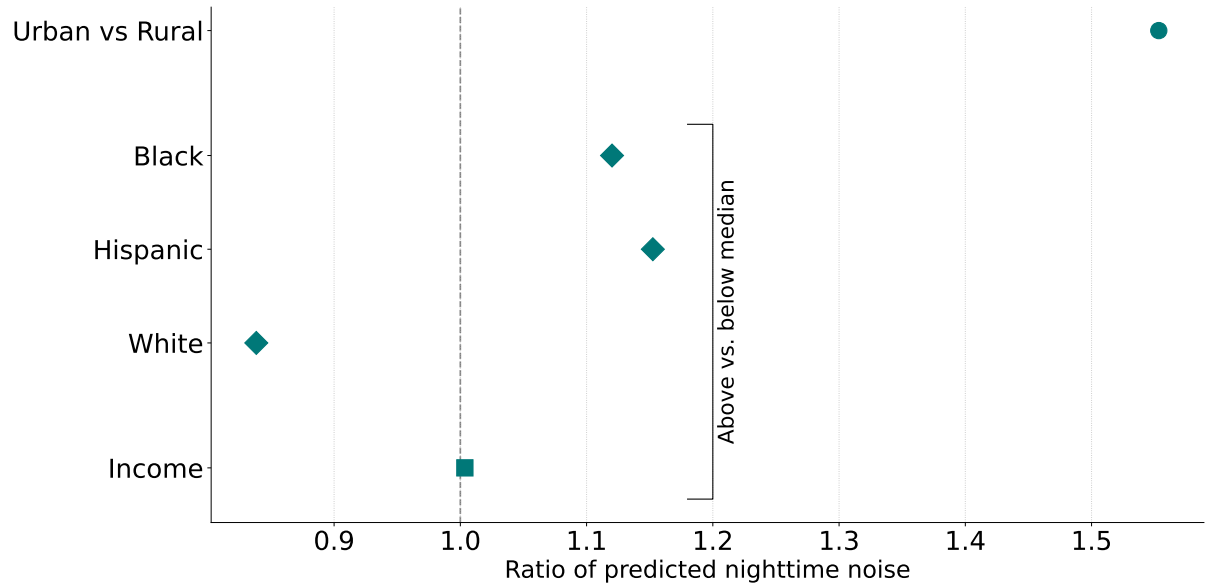
the abrupt change in predicted noise levels as the urban area buts up against the Rocky Mountains to the west. The urban area is generally louder, while the mountainous areas to the west are quieter, with the exception of roads travelling through the mountains. Farmland east of the city is relatively quiet, and the Denver International Airport is clearly visible near the north-east corner of the map.

Figure 13 summarizes differences in nighttime noise exposure by geography and demographics. Urban Census block groups are 55% louder than rural ones, an unsurprising result given that urban residents are more likely to live near major noise sources such as roads, rail lines, and airports. Census block groups with above-median Black and Hispanic populations are also exposed to more noise than Census block groups with below-median Black and Hispanic populations. In contrast, Census block groups with above-median white populations face 83% of the noise exposure of Census block groups with below-median white populations. Noise exposure in Census block groups with above-median per capita annual income is nearly identical to noise exposure in below-median income Census block groups.

I now turn to estimating of the annual national cost of noise pollution due to harms to infant health. To implement this calculation, I combine the per-decibel cost estimated on the California sample with my national noise exposure map. As when calculating the cost of all rail noise, I define excess exposure over 45 dB. To estimate the fetal exposure rate, I use Census tract-level fertility data from the 2015-2019 American Communities Survey.

The annual cost of noise pollution due to harms to infant health in the contiguous United

Figure 13: Inequalities in Predicted Noise Exposure



Notes: Figure shows the ratio of average predicted noise levels for different populations.The first row shows the ratio for fully urban to fully rural Census block groups. About one-third of census block groups are neither fully urban nor fully rural and are dropped from this comparison. The second through fourth rows show ratios for census block groups with above vs. below median share non-Hispanic Black, Hispanic, and non-Hispanic white populations, respectively. The fifth row shows the ratio for census block groups with above vs. below-median income per capita. Dashed vertical line denotes an equal ratio of 1. Race/ethnicity and income data come from the 2015-2019 5-year American Communities Survey. Urban/rural classifications come from the 2020 Census.

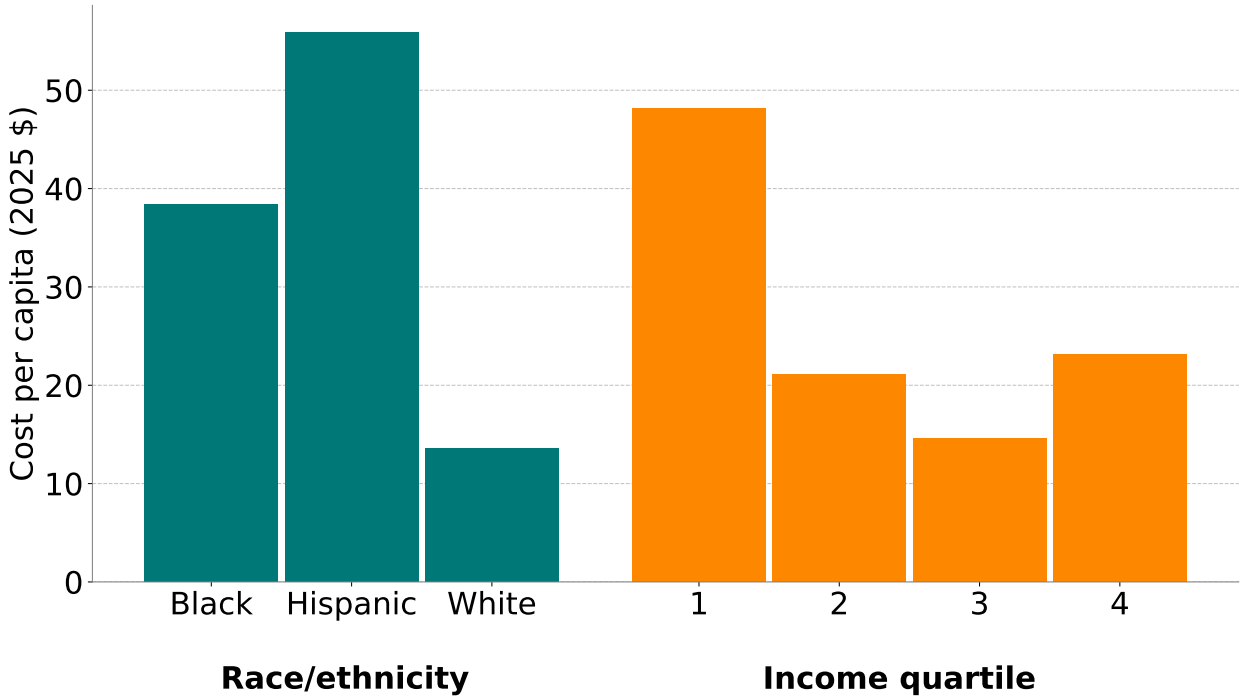
States is $8.4 billion per year. This roughly corresponds to 10,000 excess preterm births (2.7% of the total), 7,500 excess low birth weight births (2.5%), and 630 excess infant deaths (3.1%).[22]

Figure 14 reports costs per capita costs by racial and ethnic groups and quartiles of per capita annual income. Hispanic populations bear the largest annual burden of noise pollution, at $65 per capita, Black populations face a cost of $44 per capita, and White populations face a smaller cost of $16 per capita. The distribution of costs is also uneven across income quartiles, with the first income quartile bearing $55 per capita, more than triple the cost to the third income quartile, which faces a per capita cost of $17.

While nearly all the costs noise pollution I estimate are borne by urban residents, the distribution of costs across races, ethnicities, and income groups is not driven by the differential demographics of urban compared to rural census tracts. Appendix Figure 13 plots

[22]These values are calculated using the outcome-specific estimates reported in Appendix Table 4 and 2023 data from the National Center for Health Statistics.

Figure 14: Annual Costs of Noise Pollution per Capita Due to Harms to Infant Health



Notes: The figure shows the estimated annual per capita cost of noise pollution due to harms to infant health by race/ethnicity and income quartile. Costs for each group are the sum of tract-level costs (costs per tract multiplied by the group's population in that tract), divided by the group's total national population. Racial and ethnic categories are non-Hispanic Black people, non-Hispanic White people, and Hispanic people of any race. Income quartiles are based on average annual per capita income at the census tract level. All values are in 2025 dollars.

per capita costs in urban census tracts only. While the level of costs is roughly $15 higher across all groups, the qualitative pattern is identical. Rather, cost differences are driven by higher noise exposure for minority and lower-income groups, as well as higher fertility rates among those groups.

7 Conclusion

The history of noise pollution and policy efforts to control it dates to Julius Caesar's Roman Empire, and noise has received theoretical attention from economists since at least Coase (1960). This paper adds modern empirical evidence to this long history. Noise pollution is a costly environmental externality with an unequal distribution across groups. Nighttime noise appears to be especially costly, with implications for the optimal provision and scheduling of rail and air transportation. More broadly, the new noise dataset I build may open opportunities for further research on noise pollution.

I find that in utero noise pollution exposure harms health at birth: a 2 decibel (5 percent) increase in utero noise reduces an infant health index by 4% of a standard deviation, one-third of the Black-White gap in the index. The effects I find on low birth weight rates are roughly 60 percent of the effect of air pollution and are equivalent to a post-tax transfer of $2,500. Urban, Black, and Hispanic residents are disproportionately exposed to noise pollution, which I estimate costs $8.4 billion per year due to the harms it causes to health at birth.

Noise control policy in the United States is limited and blunt. Federal guidelines provide sound pressure level (i.e., decibel) thresholds for aviation and traffic noise, while local noise policy often takes the form of ordinances prohibiting any noise emission above specific thresholds. The results presented here can help inform these thresholds. My findings that noise can harm health up to 5 km away from the source and that the widely used National Transportation Noise Map explains only a small fraction of variation in measured noise suggest that there may be welfare benefits to revising existing policies to reduce noise exposure. The approaches to ambient noise monitoring developed here could help local policymakers enforce existing noise ordinances. This paper's results could also inform cost-benefit analyses of investments reducing noise exposure, or could be used as an input to a Pigouvian tax on noise emissions.

Looking forward, the clean energy transition and technological change are altering soundscapes. These changes might reduce exposures, for example by replacing internal combustion engines with electric motors. On the other hand, key exposures—such as highway noise—may remain unchanged, since rolling noise dominates engine noise above low speeds (Campello-Vicente et al., 2017). New sources of noise pollution, including wind turbines and data centers, are also emerging and proliferating. The results presented here suggest that noise merits attention in policymaking and evaluation—whether as a co-benefit, new cost, or separate consideration.

References

Aizer, Anna and Janet Currie, "The Intergenerational Transmission of Inequality: Maternal Disadvantage and Health at Birth," *Science*, May 2014, *344* (6186), 856–861.

Allroggen, Florian, R John Hansman, Christopher R Knittel, Jing Li, Xibo Wan, and Juju Wang, "Planes Overhead: How Airplain Noise Impacts Home Values," Technical Report, National Bureau of Economic Research 2025.

Almagro, Milena, Felipe Barbieri, Juan Camilo Castillo, Nathaniel G Hickok, and Tobias Salz, "Optimal Urban Transportation Policy: Evidence from Chicago," Technical Report, National Bureau of Economic Research 2024.

Almond, Douglas, "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, August 2006, *114* (4), 672–712.

_ **and Bhashkar Mazumder**, "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy," *American Economic Journal: Applied Economics*, October 2011, *3* (4), 56–85.

_ **, Janet Currie, and Valentina Duque**, "Childhood Circumstances and Adult Outcomes: Act II," *Journal of Economic Literature*, December 2018, *56* (4), 1360–1446.

_ **, Kenneth Y Chay, and David S Lee**, "The Costs of Low Birth Weight," *Quarterly Journal of Economics*, 2005, *120* (3), 3.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, December 2008, *103* (484), 1481–1495.

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

Andrews, Isaiah and Jesse M Shapiro, "Bootstrap Diagnostics for Irregular Estimators," 2024.

Angrist, Joshua D. and Alan B. Krueger, "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 1992, *87* (418), 328–336.

Arana, Miguel, Ricardo San Martin, and Juan C. Salinas, "People Exposed to Traffic Noise in European Agglomerations from Noise Maps. A Critical Review," *Noise Mapping*, September 2014, *1* (1).

Argys, Laura M., Susan L. Averett, and Muzhe Yang, "Residential Noise Exposure and Health: Evidence from Aviation Noise and Birth Outcomes," *Journal of Environmental Economics and Management*, September 2020, *103*, 102343.

Barwick, Panle Jia, Shanjun Li, Andrew Waxman, Jing Wu, and Tianli Xia, "Efficiency and Equity Impacts of Urban Transportation Policies with Equilibrium Sorting," *American Economic Review*, October 2024, *114* (10), 3161–3205.

Basner, Mathias, Alexander Samel, and Ullrich Isermann, "Aircraft Noise Effects on Sleep: Application of the Results of a Large Polysomnographic Field Study," *The Journal of the Acoustical Society of America*, May 2006, *119* (5), 2772–2784.

_ **, Uwe Müller, and Eva-Maria Elmenhorst**, "Single and Combined Effects of Air, Road, and Rail Traffic Noise on Sleep and Recuperation," *Sleep*, January 2011, *34* (1), 11–23.

Behrman, Richard E. and Adrienne Stih Butler, *Preterm Birth: Causes, Consequences, and Prevention*, Washington, D.C.: National Academies Press, April 2007.

Berkouwer, Susanna and Joshua Dean, "Cooking, Health, and Daily Exposure to Pollution Spikes," Technical Report, National Bureau of Economic Research, Cambridge, MA March 2025.

Bharadwaj, Prashant, Petter Lundborg, and Dan-Olof Rooth, "Birth Weight in the Long Run," *Journal of Human Resources*, January 2018, *53* (1), 189–231.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes," *Quarterly Journal of Economics*, 2007, *122* (1), 409–439.

Bonnefoy-Claudet, Sylvette, Fabrice Cotton, and Pierre-Yves Bard, "The Nature of Noise Wavefield and Its Applications for Site Effects Studies: A Literature Review," *Earth-Science Reviews*, December 2006, *79* (3), 205–227.

Brown, Christopher F., Michal R. Kazmierski, Valerie J. Pasquarella, William J. Rucklidge, Masha Samsikova, Chenhui Zhang, Evan Shelhamer, Estefania Lahera, Olivia Wiles, Simon Ilyushchenko, Noel Gorelick, Lihui Lydia Zhang, Sophia Alj, Emily Schechter, Sean Askay, Oliver Guinan, Rebecca Moore, Alexis Boukouvalas, and Pushmeet Kohli, "AlphaEarth Foundations: An Embedding Field Model for Accurate and Efficient Global Mapping from Sparse Label Data," July 2025.

Bureau of Transportation Statistics, "National Transportation Noise Map," November 2022.

Campello-Vicente, Hector, Ramon Peral-Orts, Nuria Campillo-Davo, and Emilio Velasco-Sanchez, "The Effect of Electric Vehicles on Urban Noise Maps," *Applied Acoustics*, January 2017, *116*, 59–64.

Caplan-Auerbach, Jacqueline, Kyla Marczewski, and Gavin Bullock, "Beast Quake (Taylor's Version): Analysis of Seismic Signals Recorded during Two Taylor Swift Concerts in Seattle, July 2023," *GSA Today*, May 2024, *34* (5), 4–10.

Chavez, Mayra and Wen-Whai Li, "Comparison of Modeled-to-Monitored PM2.5 Exposure Concentrations Resulting from Transportation Emissions in a Near-Road Community," *Transportation Research Record*, December 2020, *2674* (12), 130–143.

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

Coase, R. H., "The Problem of Social Cost," *The Journal of Law & Economics*, 1960, *3*, 1–44.

Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, March 2009, *47* (1), 87–122.

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

_ **and** _ , "What Do Economists Have to Say about the Clean Air Act 50 Years after the Establishment of the Environmental Protection Agency?," *Journal of Economic Perspectives*, November 2019, *33* (4), 3–26.

_ **, Joshua Graff Zivin, Katherine Meckel, Matthew Neidell, and Wolfram Schlenker**, "Something in the Water: Contaminated Drinking Water and Infant Health," *Canadian Journal of Economics/Revue canadienne d'économique*, 2013, *46* (3), 791–810.

_ **, Michael Greenstone, and Katherine Meckel**, "Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania," *Science Advances*, December 2017, *3* (12).

_ **, Michael Mueller-Smith, and Maya Rossin-Slater**, "Violence While in Utero: The Impact of Assaults during Pregnancy on Birth Outcomes," *The Review of Economics and Statistics*, May 2022, *104* (3), 525–540.

Dean, Joshua T., "Noise, Cognitive Function, and Worker Productivity," *American Economic Journal: Applied Economics*, October 2024, *16* (4), 322–360.

Denton, Paul, Stewart Fishwick, Victoria Lane, and Debra Daly, "Football Quakes as a Tool for Student Engagement," *Seismological Research Letters*, June 2018, *89* (5), 1902–1907.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif, "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction," *American Economic Review*, December 2019, *109* (12), 4178–4219.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro, "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program," *American Economic Review*, October 2017, *107* (10), 2958–2989.

Du, Xinming and Charles Taylor, "Airlines, Pollution, and Fertility," *SSRN Electronic Journal*, 2024.

Elmenhorst, Eva-Maria, David Elmenhorst, Juergen Wenzel, Julia Quehl, Uwe Mueller, Hartmut Maass, Martin Vejvoda, and Mathias Basner, "Effects of Nocturnal Aircraft Noise on Cognitive Performance in the Following Morning: Dose–Response Relationships in Laboratory and Field," *International Archives of Occupational and Environmental Health*, October 2010, *83* (7), 743–751.

EPA, "Mortality Risk Valuation," https://www.epa.gov/environmental-economics/mortality-risk-valuation April 2014.

_ , "Technical Assistance Document for the Reporting of Daily Air Quality – the Air Quality Index (AQI)," Technical Report EPA 454/B-18-007, Environmental Protetion Agency September 2018.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth, "The Effects of Poor Neonatal Health on Children's Cognitive Development," *American Economic Review*, December 2014, *104* (12), 3921–3955.

Gadal, Cyril, Pauline Delorme, Clément Narteau, Giles F. S. Wiggs, Matthew Baddock, Joanna M. Nield, and Philippe Claudin, "Local Wind Regime Induced by Giant Linear Dunes: Comparison of ERA5-Land Reanalysis with Surface Measurements," *Boundary-Layer Meteorology*, December 2022, *185* (3), 309–332.

Gehring, Ulrike, Lillian Tamburic, Hind Sbihi, Hugh W. Davies, and Michael Brauer, "Impact of Noise and Air Pollution on Pregnancy Outcomes," *Epidemiology*, May 2014, *25* (3), 351.

Godsey, Mike, "West Coast Wind Blog: The Nature of the Beast...A Cut-off Low Kills San Francisco. Southern California and the Pacific Northwest Wind.," September 2023.

Google, "What Are Foundation Models?," 2025.

Goulder, Lawrence H. and Ian W. H. Parry, "Instrument Choice in Environmental Policy," *Review of Environmental Economics and Policy*, July 2008, *2* (2), 152–174.

Gualtieri, Giovanni, "Reliability of ERA5 Reanalysis Data for Wind Resource Assessment: A Comparison against Tall Towers," *Energies*, January 2021, *14* (14), 4169.

Gutenberg, Beno, *Die seismische Bodenunruhe*, Stürtz, 1911.

Hener, Timo, "Noise Pollution and Violent Crime," *Journal of Public Economics*, November 2022, *215*, 104748.

Hepworth, Peter, "Accuracy Implications of Computerized Noise Predictions for Environmental Noise Mapping," *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, December 2006, *2006* (6), 1604–1609.

Herrnstadt, Evan, Anthony Heyes, Erich Muehlegger, and Soodeh Saberian, "Air Pollution and Criminal Activity: Microgeographic Evidence from Chicago," *American Economic Journal: Applied Economics*, October 2021, *13* (4), 70–100.

Hogan, Dylan and Wolfram Schlenker, "Non-Linear Relationships between Daily Temperature Extremes and US Agricultural Yields Uncovered by Global Gridded Meteorological Datasets," *Nature Communications*, May 2024, *15* (1), 4638.

Hornikx, M., M. Dohmen, K. Conen, T. van Hooff, and B. Blocken, "The Wind Effect on Sound Propagation over Urban Areas: Predictions for Generic Urban Sections," *Building and Environment*, October 2018, *144*, 519–531.

Hoynes, Hilary, Doug Miller, and David Simon, "Income, the Earned Income Tax Credit, and Infant Health," *American Economic Journal: Economic Policy*, February 2015, *7* (1), 172–211.

Hu, Shishan, Scott Fruin, Kathleen Kozawa, Steve Mara, Suzanne E. Paulson, and Arthur M. Winer, "A Wide Area of Air Pollutant Impact Downwind of a Freeway during Pre-Sunrise Hours," *Atmospheric Environment*, May 2009, *43* (16), 2541–2549.

Hudda, Neelakshi, Tim Gould, Kris Hartin, Timothy V. Larson, and Scott A. Fruin, "Emissions from an International Airport Increase Particle Number Concentrations 4-Fold at 10 Km Downwind," *Environmental Science & Technology*, June 2014, *48* (12), 6628–6635.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker, "Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy*, June 2017, *125* (3), 848–902.

Jephcote, Calvin, Sierra N. Clark, Anna L. Hansell, Nigel Jones, Yingxin Chen, Claire Blackmore, Katie Eminson, Megan Evans, Xiangpu Gong, Kathryn Adams, Georgia Rodgers, Benjamin Fenech, and John Gulliver, "Spatial Assessment of the Attributable Burden of Disease Due to Transportation Noise in England," *Environment International*, August 2023, *178*, 107966.

Johnson, Allan Chester, Paul Robinson Coleman-Norton, and Frank Card Bourne, *Ancient Roman Statutes: A Translation with Introduction, Commentary, Glossary, and Index*, The Lawbook Exchange, Ltd., 2003.

Johnson, Hans, "California's Plunging Birth Rates," January 2023.

Kearney, Melissa S. and Phillip B. Levine, "The US COVID-19 Baby Bust and Rebound," *Journal of Population Economics*, October 2023, *36* (4), 2145–2168.

_ , _ , and Luke Pardue, "The Puzzle of Falling US Birth Rates since the Great Recession," *Journal of Economic Perspectives*, February 2022, *36* (1), 151–176.

Keiser, David A and Joseph S Shapiro, "Consequences of the Clean Water Act and the Demand for Water Quality*," *The Quarterly Journal of Economics*, February 2019, *134* (1), 349–396.

Keiser, David A. and Joseph S. Shapiro, "US Water Pollution Regulation over the Past Half Century: Burning Waters to Crystal Springs?," *Journal of Economic Perspectives*, November 2019, *33* (4), 51–75.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kriebel, Anthony R., "Refraction and Attenuation of Sound by Wind and Thermal Profiles over a Ground Plane," *The Journal of the Acoustical Society of America*, January 1972, *51* (1A), 19–23.

LaDochy, Stephen and Michael Witiw, "California Weather and Air Pollution," in Stephen LaDochy and Michael Witiw, eds., *Fire and Rain: California's Changing Weather and Climate*, Cham: Springer Nature Switzerland, 2023, pp. 185–196.

Lecocq, Thomas, Stephen P. Hicks, Koen Van Noten, Kasper van Wijk, Paula Koelemeijer, Raphael S. M. De Plaen, Frédérick Massin, Gregor Hillers, Robert E. Anthony, Maria-Theresia Apoloner, Mario Arroyo-Solórzano, Jelle D. Assink, Pinar Büyükakpınar, Andrea Cannata, Flavio Cannavo, Sebastian Carrasco, Corentin Caudron, Esteban J. Chaves, David G. Cornwell, David Craig, Olivier F. C. den

Ouden, Jordi Diaz, Stefanie Donner, Christos P. Evangelidis, Läslo Evers, Benoit Fauville, Gonzalo A. Fernandez, Dimitrios Giannopoulos, Steven J. Gibbons, Társilo Girona, Bogdan Grecu, Marc Grunberg, György Hetényi, Anna Horleston, Adolfo Inza, Jessica C. E. Irving, Mohammadreza Jamalreyhani, Alan Kafka, Mathijs R. Koymans, Celeste R. Labedz, Eric Larose, Nathaniel J. Lindsey, Mika McKinnon, Tobias Megies, Meghan S. Miller, William Minarik, Louis Moresi, Víctor H. Márquez-Ramírez, Martin Möllhoff, Ian M. Nesbitt, Shankho Niyogi, Javier Ojeda, Adrien Oth, Simon Proud, Jay Pulli, Lise Retailleau, Annukka E. Rintamäki, Claudio Satriano, Martha K. Savage, Shahar Shani-Kadmiel, Reinoud Sleeman, Efthimios Sokos, Klaus Stammler, Alexander E. Stott, Shiba Subedi, Mathilde B. Sørensen, Taka'aki Taira, Mar Tapia, Fatih Turhan, Ben van der Pluijm, Mark Vanstone, Jerome Vergne, Tommi A. T. Vuorinen, Tristram Warren, Joachim Wassermann, and Han Xiao, "Global Quieting of High-Frequency Seismic Noise Due to COVID-19 Pandemic Lockdown Measures," *Science*, September 2020, *369* (6509), 1338–1343.

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, May 2009, *99* (3), 1006–1026.

McNamara, Daniel E. and Raymond P. Buland, "Ambient Noise Levels in the Continental United States," *Bulletin of the Seismological Society of America*, August 2004, *94* (4), 1517–1527.

Medved, Sašo, "Building Acoustics and Noise Control in Buildings," in Sašo Medved, ed., *Building Physics: Heat, Ventilation, Moisture, Light, Sound, Fire, and Urban Microclimate*, Cham: Springer International Publishing, 2022, pp. 331–406.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw, "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *The American Economic Review*, 1994, *84* (4), 753–771.

Mennitt, Daniel J. and Kurt M. Fristrup, "Influential Factors and Spatiotemporal Patterns of Environmental Sound Levels in the Contiguous United States," *Noise Control Engineering Journal*, May 2016, *64* (3), 342–353.

Moretti, Enrico and Harrison Wheeler, "The Traffic Noise Externality: Costs, Incidence and Policy Implications," 2025.

Muñoz-Sabater, Joaquín, Emanuel Dutra, Anna Agustí-Panareda, Clément Albergel, Gabriele Arduini, Gianpaolo Balsamo, Souhail Boussetta, Margarita Choulga, Shaun Harrigan, Hans Hersbach, Brecht Martens, Diego G. Miralles, María Piles, Nemesio J. Rodríguez-Fernández, Ervin Zsoter, Carlo Buontempo, and Jean-Noël Thépaut, "ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications," *Earth System Science Data*, September 2021, *13* (9), 4349–4383.

Münzel, Thomas, Andreas Daiber, Nicole Engelmann, Martin Röösli, Marin Kuntic, and Jamie L. Banks, "Noise Causes Cardiovascular Disease: It's Time to Act," *Journal of Exposure Science & Environmental Epidemiology*, December 2024, pp. 1–10.

Munzel, Thomas, Tommaso Gori, Wolfgang Babisch, and Mathias Basner, "Cardiovascular Effects of Environmental Noise Exposure," *European Heart Journal*, April 2014, *35* (13), 829–836.

Murphy, Enda and Eoin A. King, "Environmental Noise Pollution," in "Environmental Noise Pollution," Elsevier, 2014, pp. 1–7.

National Academies of Sciences, Engineering, and Medicine, *How Weather Affects the Noise You Hear from Highways*, Washington, D.C.: The National Academies Press, August 2018.

National Academy of Engineering, *Technology for a Quieter America*, Washington, D.C.: National Academies Press, September 2010.

National Archives, "Constitution of the United States—A History," November 2015.

Nelson, Charles R. and Richard Startz, "The Distribution of the Instrumental Variables Estimator and Its T-Ratio When the Instrument Is a Poor One," *The Journal of Business*, 1990, *63* (1), S125–S140.

Nijland, H. A. and G. P. Van Wee, "Traffic Noise in Europe: A Comparison of Calculation Methods, Noise Indices and Noise Standards for Road and Railroad Traffic in Europe," *Transport Reviews*, September 2005, *25* (5), 591–612.

Nobles, Jenna, Alison Gemmill, Sungsik Hwang, and Florencia Torche, "Fertility in a Pandemic: Evidence from California," *Population and Development Review*, 2024, *50* (S1), 101–128.

of Transportation Bureau of Transportation Statistics, U.S. Department, "National Transportation Noise Map Documentation," 2020.

Parry, Ian W. H and Kenneth A Small, "Should Urban Transit Subsidies Be Reduced?," *American Economic Review*, May 2009, *99* (3), 700–724.

Ramon, Jaume, Llorenç Lledó, Verónica Torralba, Albert Soret, and Francisco J. Doblas-Reyes, "What Global Reanalysis Best Represents Near-Surface Winds?," *Quarterly Journal of the Royal Meteorological Society*, 2019, *145* (724), 3236–3251.

Riahi, Nima and Peter Gerstoft, "The Seismic Traffic Footprint: Tracking Trains, Aircraft, and Cars Seismically," *Geophysical Research Letters*, 2015, *42* (8), 2674–2681.

Ross, Charles, "Acoustic Shadows in the Civil War," in "Acoustical Society of America," Vol. 136 October 1998.

Roswall, Nina, Andrei Pyko, Mikael Ögren, Anna Oudin, Annika Rosengren, Anton Lager, Aslak H. Poulsen, Charlotta Eriksson, David Segersson, Debora Rizzuto, Eva M. Andersson, Gunn Marit Aasvang, öm Gunnar Engstr, ørgensen Jeanette T. J, Jenny Selander, Jesper H. Christensen, Jesse Thacher, Karin Leander, Kim Overvad, Kristina Eneroth, Kristoffer Mattisson, ard Lars Barreg, Leo Stockfelt, Maria Albin, Matthias Ketzel, Mette K. Simonsen, Marten Spanne, Nielsen Ole Raaschou, Patrik K. E. Magnusson, Pekka Tiittanen, Peter Molnar, Petter Ljungman, Timo Lanki, Youn-Hee Lim, Zorana J. Andersen, Göran Pershagen, and Mette Sorensen, "Long-Term Exposure to Transportation Noise and Risk of Incident Stroke: A Pooled Study of Nine Scandinavian Cohorts," *Environmental Health Perspectives*, 2021, *129* (10), 107002.

Salomons, Erik M., *Computational Atmospheric Acoustics*, Dordrecht: Springer Netherlands, 2001.

Schlenker, Wolfram and W. Reed Walker, "Airports, Air Pollution, and Contemporaneous Health," *The Review of Economic Studies*, April 2016, *83* (2), 768–809.

Shen, Siyuan, Chi Li, Aaron van Donkelaar, Nathan Jacobs, Chenguang Wang, and Randall V. Martin, "Enhancing Global Estimation of Fine Particulate Matter Concentrations by Including Geophysical a Priori Information in Deep Learning," *ACS ES&T Air*, May 2024, *1* (5), 332–345.

Smith, Rachel B., Sean D. Beevers, John Gulliver, David Dajnak, Daniela Fecht, Marta Blangiardo, Margaret Douglass, Anna L. Hansell, H. Ross Anderson, Frank J. Kelly, and Mireille B. Toledano, "Impacts of Air Pollution and Noise on Risk of Preterm Birth and Stillbirth in London," *Environment International*, January 2020, *134*, 105290.

Sperry, Benjamin R, "Why Is It so Loud Today?," Technical Report, Illinois Department of Transportation 2023.

Staiger, Douglas and James H. Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 1997, *65* (3), 557–586.

Suarez, Gerardo, Torild van Eck, Domenico Giardini, Tim Ahern, Rhett Butler, and Seiji Tsuboi, "The International Federation of Digital Seismograph Networks (FDSN): An Integrated System of Seismological Observatories," *IEEE Systems Journal*, September 2008, *2* (3), 431–438.

Thacher, Jesse D., Aslak H. Poulsen, Ole Raaschou-Nielsen, Ulla A. Hvidtfeldt, Jørgen Brandt, Jesper H. Christensen, Jibran Khan, Gregor Levin, Thomas Münzel, and Mette Sørensen, "Exposure to Transportation Noise and Risk for Cardiovascular Disease in a Nationwide Cohort Study from Denmark," *Environmental Research*, August 2022, *211*, 113106.

Tran, Kathy V., Joan A. Casey, Lara J. Cushing, and Rachel Morello-Frosch, "Residential Proximity to Oil and Gas Development and Birth Outcomes in California: A Retrospective Cohort Study of 2006–2015 Births," *Environmental Health Perspectives*, June 2020, *128* (6), 067001.

van Blokland, Gijsjan and Bert Peeters, "Modeling the Noise Emission of Road Vehicles and Results of Recent Experiments," in "Internoise" Ottawa 2009.

van Donkelaar, Aaron, Randall V. Martin, Michael Brauer, Ralph Kahn, Robert Levy, Carolyn Verduzco, and Paul J. Villeneuve, "Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application," *Environmental Health Perspectives*, June 2010, *118* (6), 847–855.

Wiener, Francis M. and David N. Keast, "Experimental Study of the Propagation of Sound over Ground," *The Journal of the Acoustical Society of America*, June 1959, *31* (6), 724–733.

Willshire, W. L., "Long Range Downwind Propagation of Low-Frequency Sound," Technical Report NASA-TM-86409 April 1985.

Wisconsin Department of Transportation, "Why Is It So Loud Today?," Technical Report, Wisconsin Department of Transportation 2019.

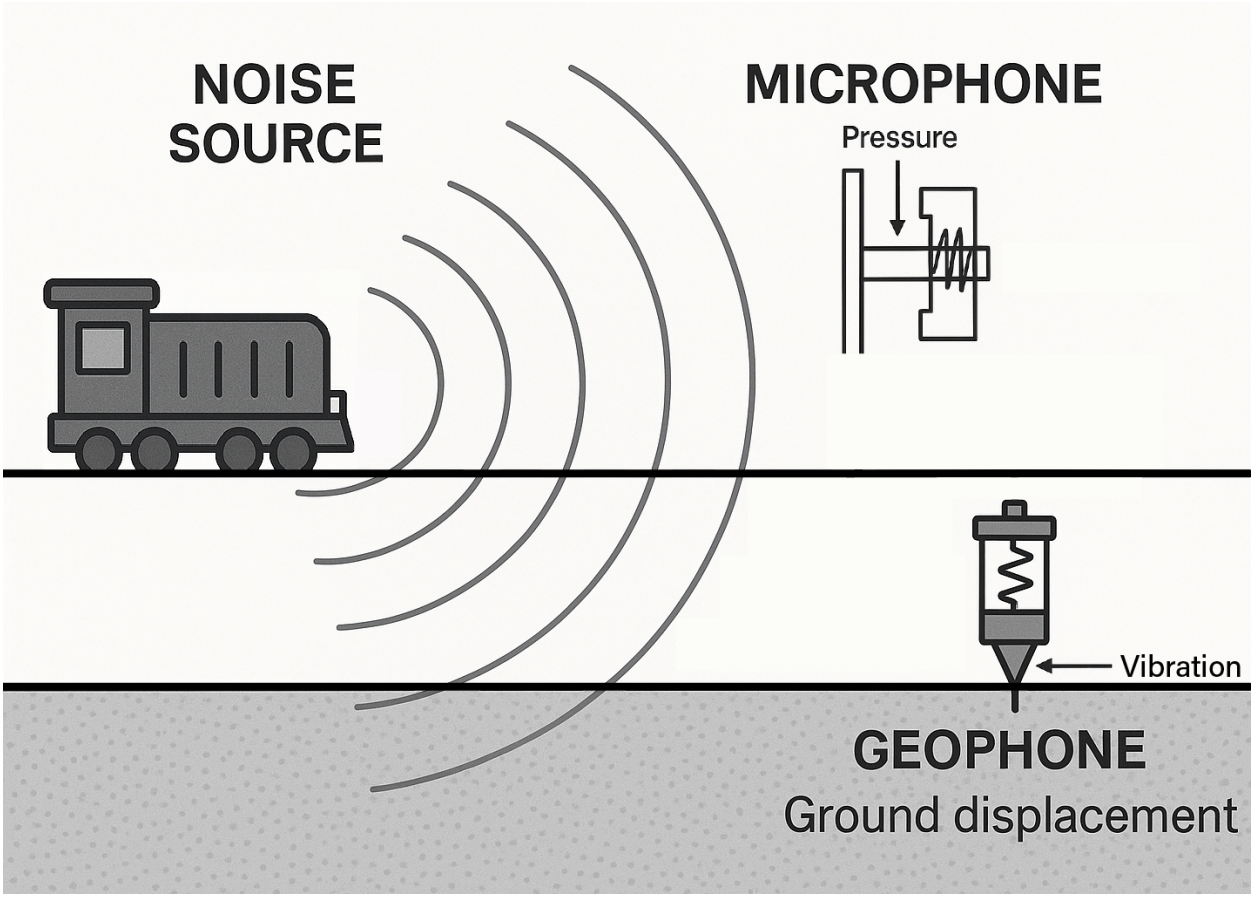
Zhan, Zhaoguo, "Detecting Weak Identification by Bootstrap," 2018.

Zhao, Qingyuan, Jingshu Wang, Wes Spiller, Jack Bowden, and Dylan S. Small, "Two-Sample Instrumental Variable Analyses Using Heterogeneous Samples," *Statistical Science*, May 2019, *34* (2).

Zhu, Yifang, Thomas Kuhn, Paul Mayo, and William C. Hinds, "Comparison of Daytime and Nighttime Concentration Profiles and Size Distributions of Ultrafine Particles near a Major Highway," *Environmental Science & Technology*, April 2006, *40* (8), 2531–2536.

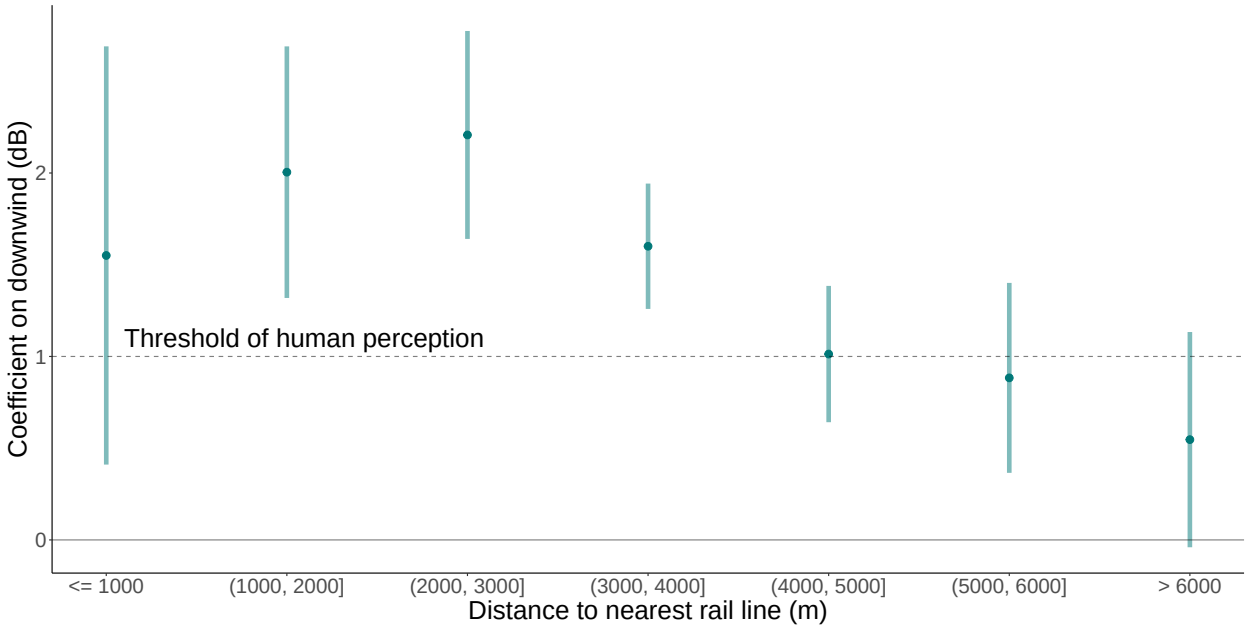
A Additional Tables and Figures

Appendix Figure 1: Schematic Showing Microphones and Seismic Noise Measurement



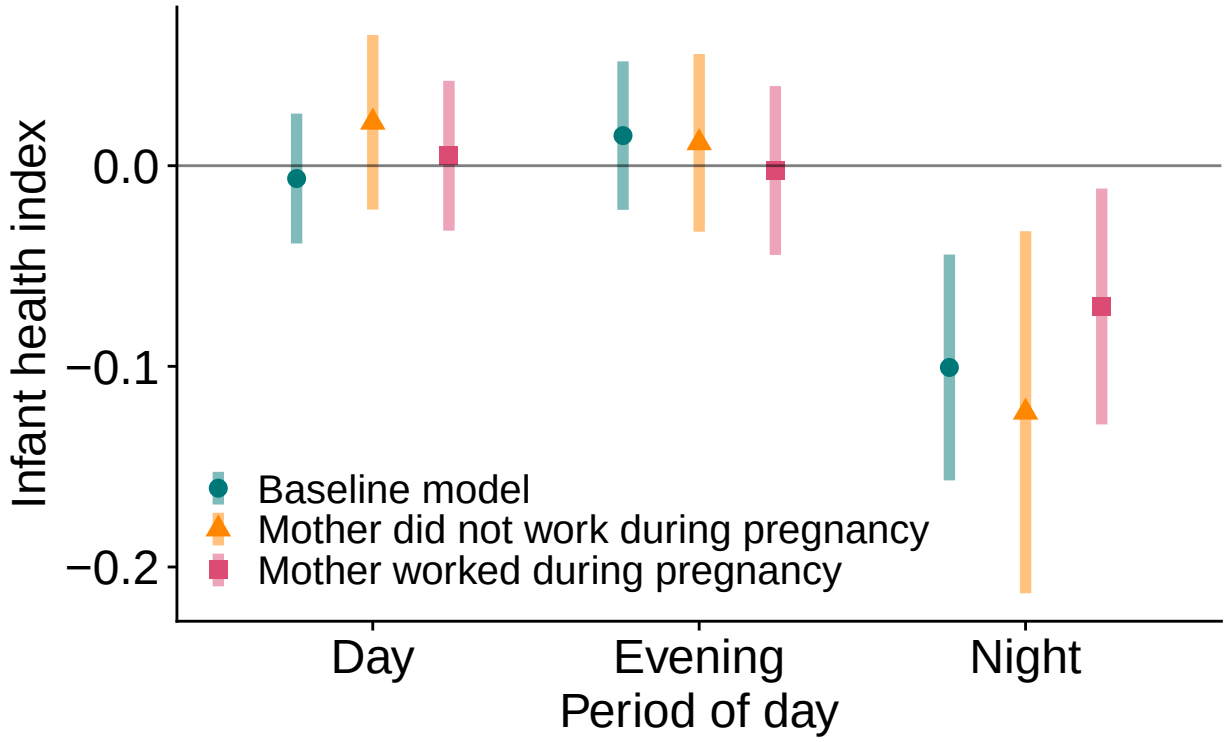
Notes: The figure shows a stylized example of noise measurement using a microphone and a seismometer (a geophone). Sound waves from a noise source such as a train are emitted in all directions. Some of these waves come into contact with a microphone, which measures changes in atmospheric pressure due to the waves traveling through the atmosphere. Other waves come into contact with the ground, causing the ground to vibrate. The geophone measures the vibrations caused by these waves travelling through the ground.

Appendix Figure 2: Impact of Downwind Status on Noise by Distance to Electric Rail



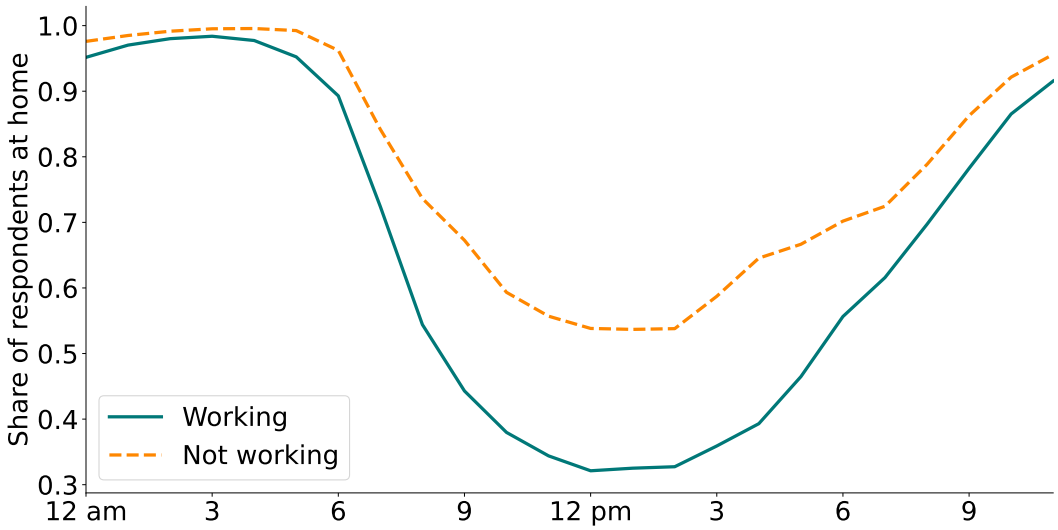
Notes: The figure shows the coefficient on an additional hour downwind by distance to the nearest electric rail line. Each coefficient corresponds to a separate regression estimated on a sample of seismometers within the distance bin indicated by the x-axis. The error bars report 95% confidence intervals clustered by seismometer. All regressions include fourth order polynomials in temperature, precipitation, and wind speed, as well as monitor and year by month fixed effects. The dashed horizontal line at 1 dB corresponds to the approximate threshold of human perception for a change in sound pressure level.

Appendix Figure 3: Effect of Train Noise On Infant Health by Time of Day and Maternal Work Status



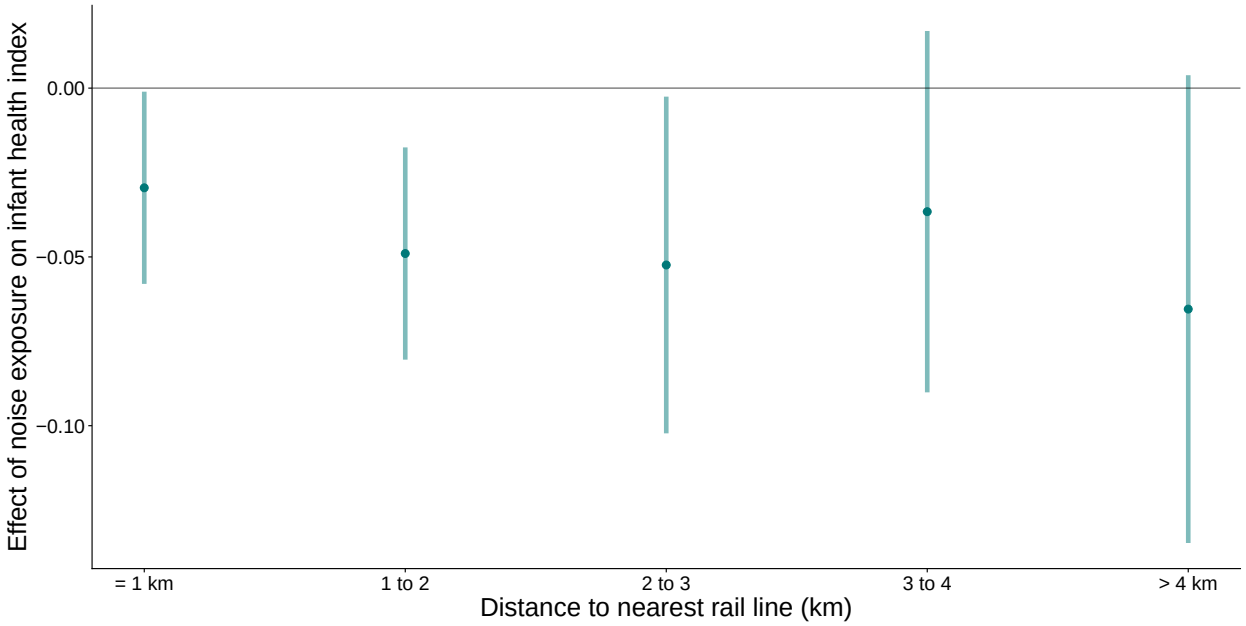
Notes: The figure shows coefficients and 95% Conley confidence intervals from regressions of the infant health index on the share of hours exposed to elevated noise by period of the day, allowing for heterogeneity by maternal employment status. The baseline model shown in teal replicates the results in the top left panel of figure 7. All regressions control for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

Appendix Figure 4: Share of Urban CA Women at Home by Hour of Day and Work Status



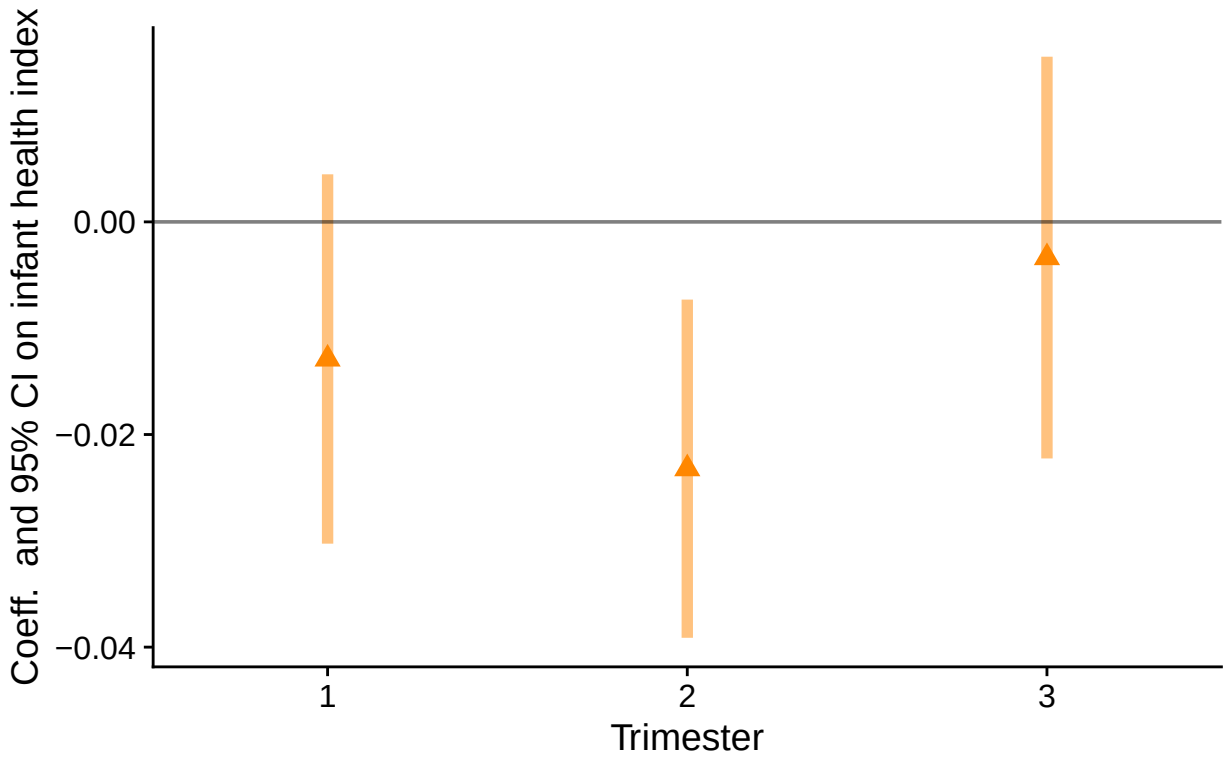
Notes: The figure shows the share of women aged 15-50 living in California metropolitan areas at home by hour of the day and work status. Data come from the American Time Use Survey, 2003-2024. Sleep and personal activities coded as at home if not otherwise specified.

Appendix Figure 5: Heterogeneity by Distance to Electric Rail



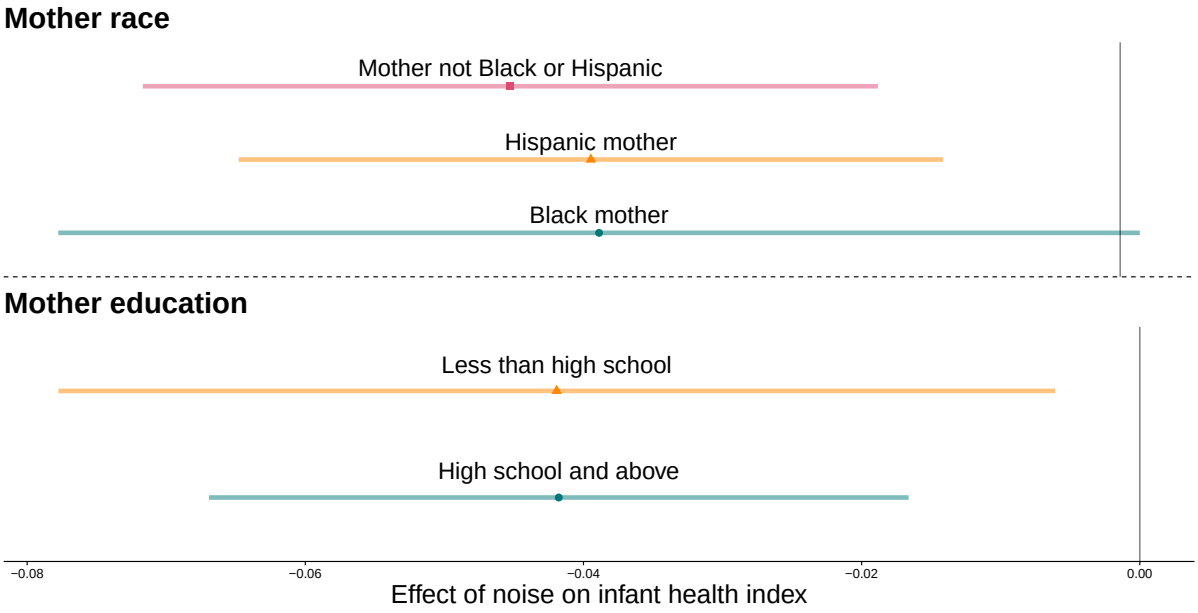
Notes: The figure plots estimated effects of noise exposure on the infant health index by one km distance bins to the nearest electric rail line. The regression includes controls for mother characteristics, weather, $PM_{2.5}$, year by month, and census block fixed effects. 95% Conley confidence intervals with a 2.7 km bandwidth are shown.

Appendix Figure 6: Heterogeneity by Trimester of Exposure



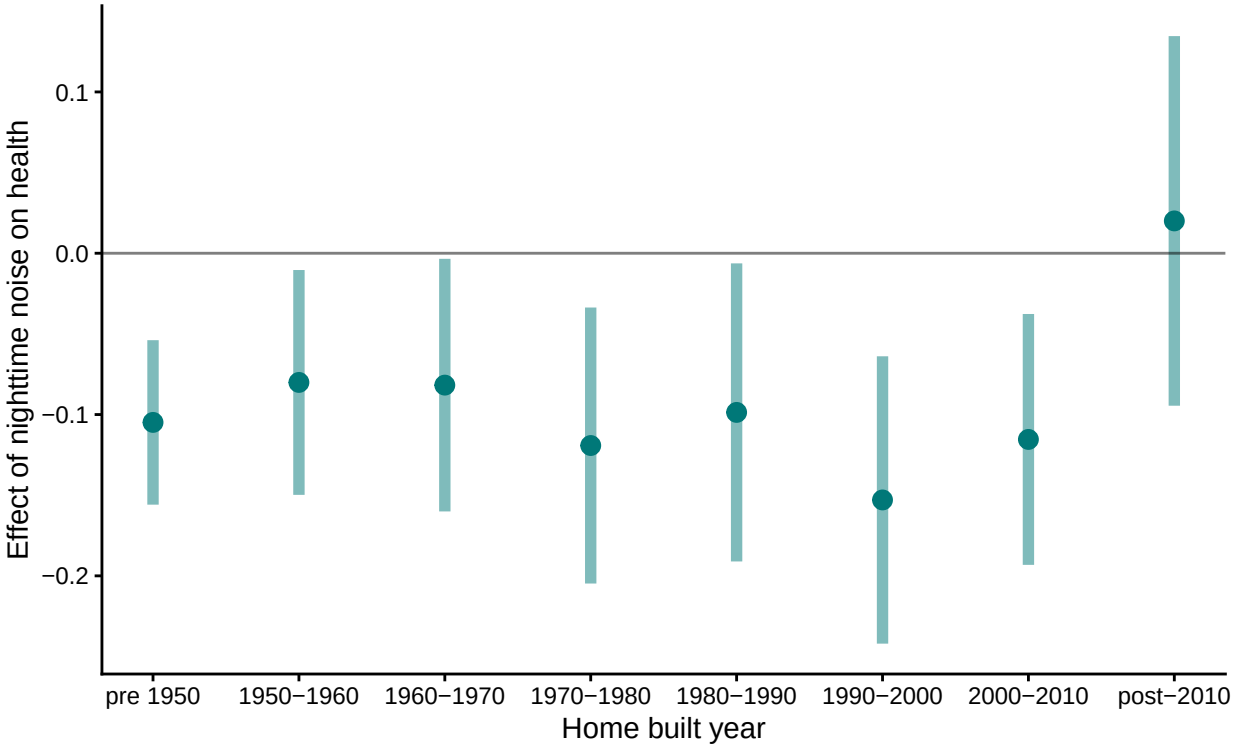
Notes: The figure shows coefficients and 95% Conley confidence intervals from regressions of the infant health index on the share of hours exposed to elevated noise during each trimester of pregnancy. All regressions control for mother characteristics, weather, $PM_{2.5}$, year by month fixed effects, and census block fixed effects.

Appendix Figure 7: Heterogeneity by Maternal Race and Education



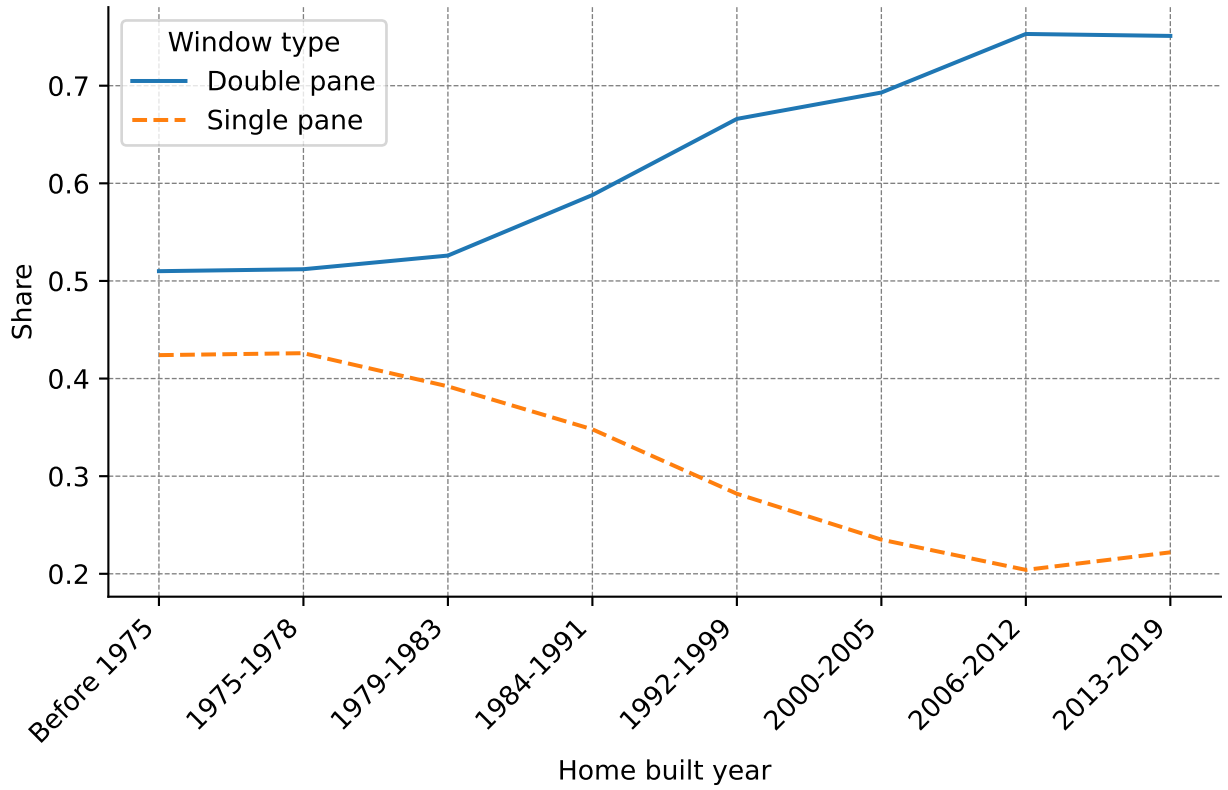
Notes: The figure shows estimated effect of noise exposure on the infant health index by trimester of exposure. The first trimester is defined as the initial 13 weeks of the pregnancy; the second is the 14th through 26th weeks; the third is the 27th through 39th. The regression includes controls for mother characteristics, weather, $PM_{2.5}$, year by month, and census block fixed effects. 95% Conley confidence intervals with a 2.7 km bandwidth are shown.

Appendix Figure 8: Heterogeneity by home construction year



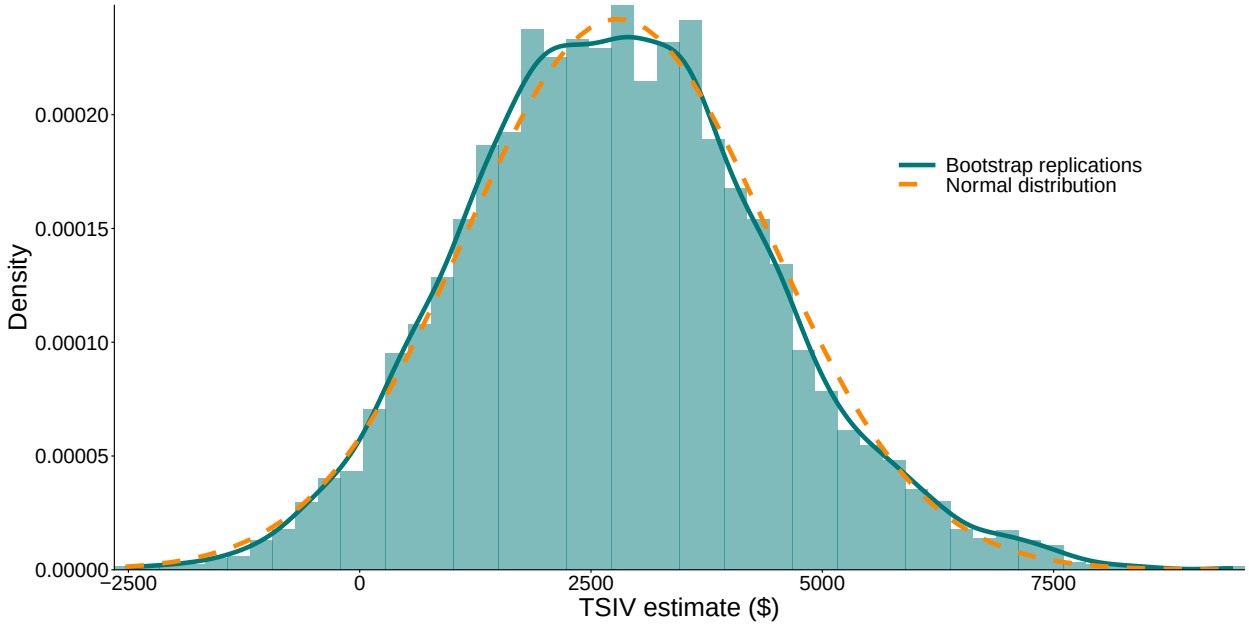
The figure reports estimated effects of nighttime noise exposure on the infant health index by the age of the mother's home. Regressions control for daytime and evening noise exposure, maternal characteristics, weather, $PM_{2.5}$, year by month, and census block fixed effects. 95% Conley confidence intervals with a 2.7 km bandwidth are shown. Home age data are from CoreLogic.

Appendix Figure 9: Window Type by Home Built Year



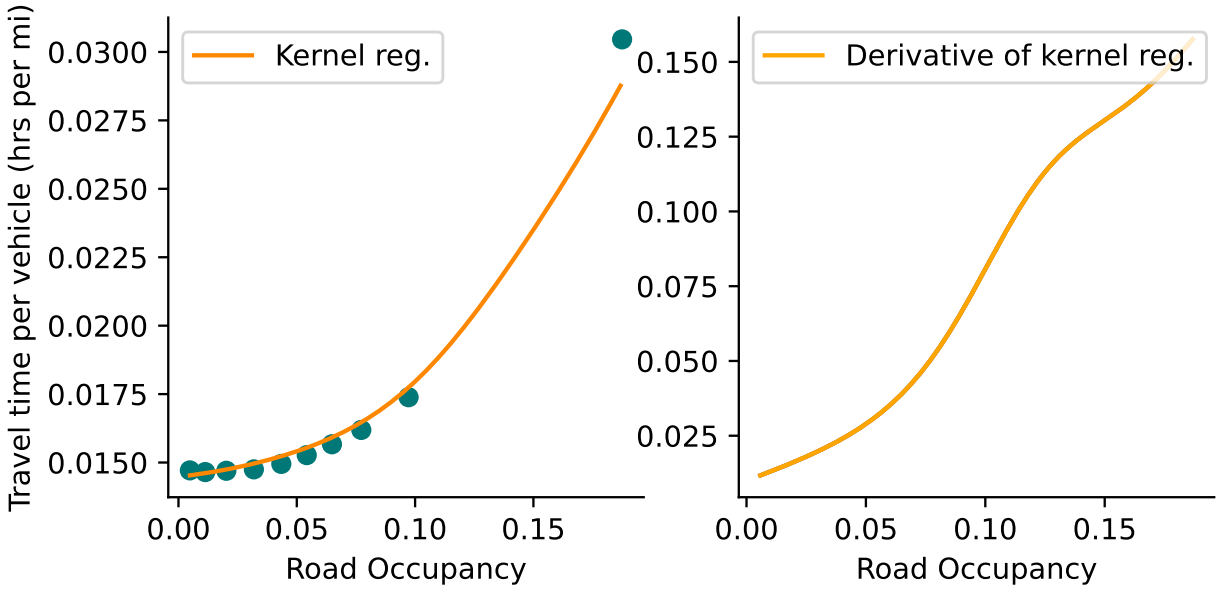
Notes: The figure shows shares of homes with all or mostly double pane and all or mostly single pane windows by year built. The shares do not sum to one because some homes have a mix of single, double, and triple-pane windows. Data come from the California Energy Commission's Residential Appliance Saturation Study.

Appendix Figure 10: Distribution of Bootstrapped TSIV Estimates

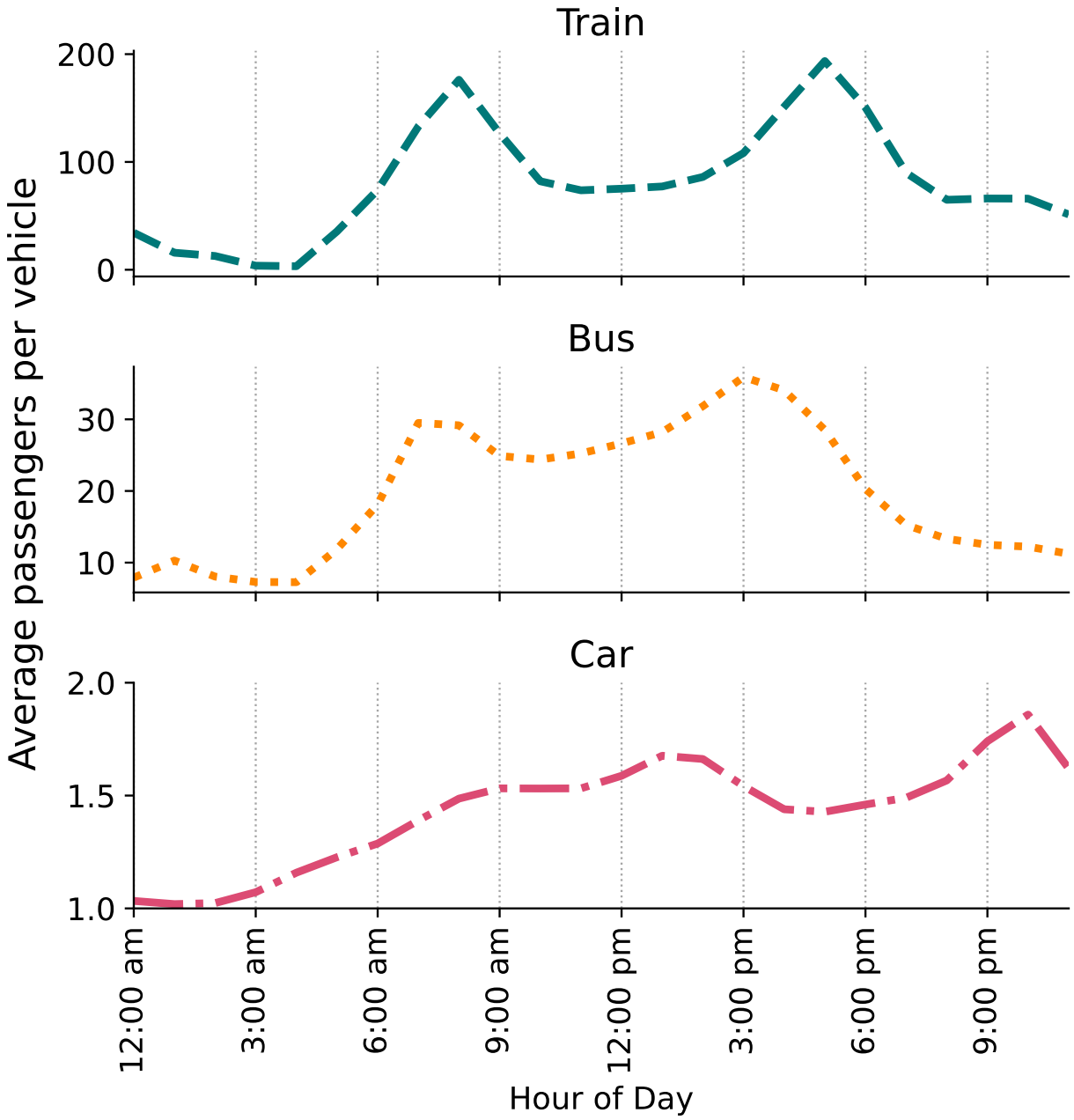


Notes: The figure shows the distribution of TSIV parameters from column (2) of Table 3 over 5,000 replications. The solid teal line is a kernel density estimate of the histogram. The dashed orange line is a normal distribution with the same mean and standard deviation as the bootstrap estimates. The distribution is approximately symmetric and well-approximated by the normal reference distribution. A Kolmogorov–Smirnov test fails to reject normality (test statistic = 0.015, $p = 0.17$).

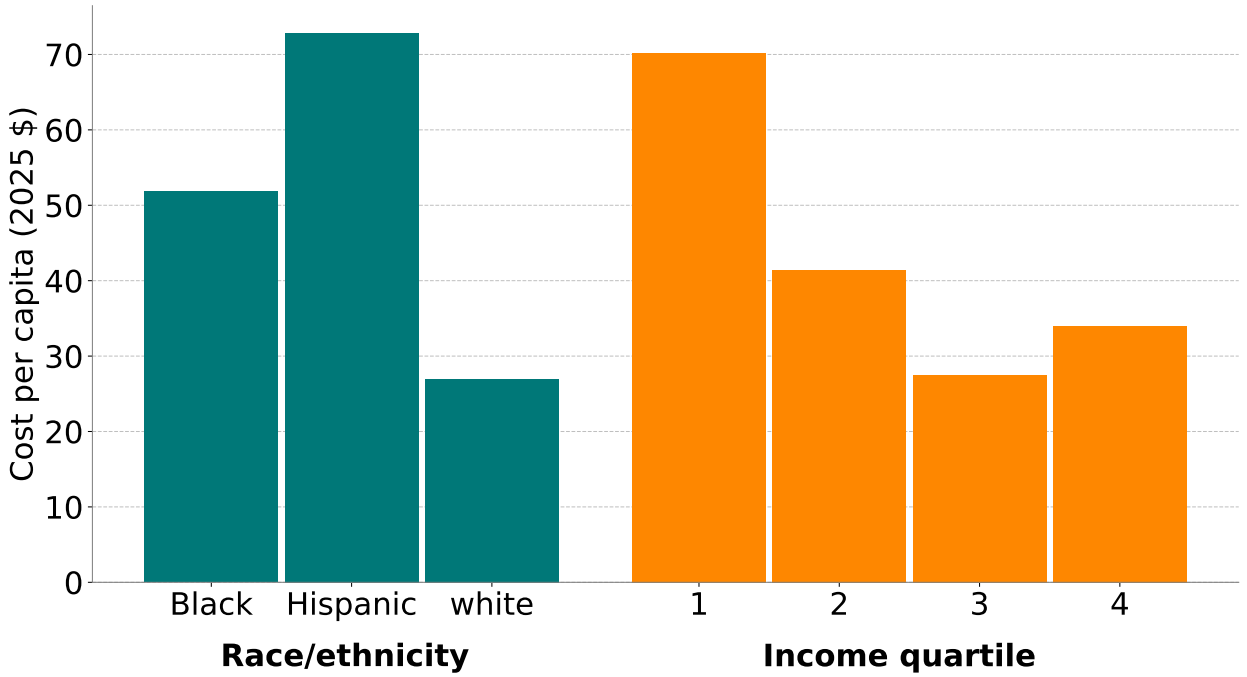
Appendix Figure 11: Fitted Congestion Externality Functions



Appendix Figure 12: Average Passengers Per Hour by Mode



Appendix Figure 13: Annual Costs of Noise Pollution per Capita, Urban Census Tracts Only



Notes: The figure shows the estimated annual per capita cost of noise pollution due to harms to infant health by race/ethnicity and income quartile for urban census tracts only. Costs for each group are the sum of tract-level costs (costs per tract multiplied by the group's population in that tract), divided by the group's total national population. Racial and ethnic categories are non-Hispanic Black people, non-Hispanic White people, and Hispanic people of any race. Income quartiles are based on average annual per capita income at the census tract level. All values are in 2025 dollars.

Appendix Table 1: Seismometer Reporting Date and Duration Summary

Network	No. of stations	Earliest start	Latest end	Avg. reporting days
NCEDC	28	1932-01-01	2023-09-07	1390
RaspberryShake	278	2017-07-18	2023-02-15	163
SCEDC	25	1976-01-01	2023-07-26	1297

Notes: The table summarizes reporting dates and durations for seismometers used to validate the research design. "NCEDC" is the Northern California Earthquake Data Center; "SCEDC" is the Southern California Earthquake Data Center. Both NCEDC and SCEDC are professionally-managed seismometer networks used for research and earthquake measurement. RaspberryShake is a distributed network of seismometers managed by independent researchers, educational institutions, and hobbyists.

Appendix Table 2: Summary Statistics

	(1) In sample	(2) Out of sample	(3) Standardized difference
A. Birth outcomes			
Birth weight (grams)	3,300	3,320	-0.034
	(566)	(561)	
Share low birth weight (< 2500 grams)	0.0690	0.0649	0.017
	(0.253)	(0.246)	
Share very low birth weight (< 1500 grams)	0.0106	0.00991	0.007
	(0.103)	(0.0990)	
Infant mortality rate per 1,000 births	1.03	1.14	-0.003
	(32.1)	(33.8)	
Gestation length (weeks)	39.1	39.1	0.004
	(2.14)	(2.14)	
Share preterm (< 37 weeks)	0.0991	0.0988	0.001
	(0.299)	(0.298)	
Share very preterm (< 34 weeks)	0.0263	0.0260	0.002
	(0.160)	(0.159)	
B. Maternal characteristics			
Age	29.1	28.7	0.066
	(6.30)	(6.23)	
Share teen mother	0.0704	0.0744	-0.015
	(0.256)	(0.262)	
Share Black	0.0850	0.0416	0.195
	(0.279)	(0.200)	
Share Hispanic	0.460	0.484	-0.048
	(0.498)	(0.500)	
Years of education	13.1	13.0	0.034
	(2.66)	(2.51)	
Share less than high school education	0.224	0.210	0.034
	(0.417)	(0.407)	
N	2,661,687	8,534,945	

Notes: Table shows summary statistics for key outcomes (panel A) and maternal characteristics (panel B). Column (1) reports statistics for births in the main analysis sample, within five km of an electric rail line. Column (2) reports statics for births outside the main analysis sample. Both columns report the mean and standard deviation below it in parentheses. Column (3) reports the difference between column (1) and column (2) divided by the standard deviation of the variable in the full sample. The full sample consists of the universe of births in California between 1999 and 2021.

Appendix Table 3: Regressions of Mother Characteristics on Noise Exposure Measure

	Demographic index (1)	Asian (2)	Black (3)	Hispanic (4)	White (5)	<20 yrs (6)	Dropout (7)	Birth order (8)	Home price ($1000s) (9)	Year home built (10)
Noise exposure	0.009	0.000	-0.003	0.025**	-0.007	0.006	0.002	-0.007	33.424	-0.047
	(0.008)	(0.009)	(0.006)	(0.012)	(0.011)	(0.007)	(0.017)	(0.017)	(46.701)	(0.434)
Dep. var. mean	0.000	0.170	0.087	0.460	0.200	0.071	0.220	2.100	760	2,000
Observations	1,960,153	1,960,153	1,960,153	1,960,153	1,960,153	1,960,153	1,960,153	1,960,153	523,275	1,080,047

Notes: The table shows the results of regressing fixed maternal and infant characteristics on the share of pregnancy spent downwind of an electric rail line, as described by equation (3). All specifications include census block and year by month fixed effects and controls for fourth-order polynomials in temperature, precipitation, and wind speed polynomials. "Black" and "Hispanic" are indicators for whether the mother is Black and Hispanic, respectively. "<20 yrs" is an indicator for whether the mother is less than 20 years old. "Dropout" is an indicator for whether the mother's highest level of educational attainment is less than high school. "Female infant" is an indicator for whether the infant is female. "Demographic index" is an inverse covariance-weighted average of standardized versions of the variables in columns (2) through (7). Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 4: The Effect of Noise on Components of the Infant Health Index

	Birth weight (g)	Low birth weight	Very low birth weight	Preterm
	(1)	(2)	(3)	(4)
Noise exposure	-14.1511*	0.0050	0.0004	0.0070
	(-1.9192)	(1.5745)	(0.3880)	(1.6125)
Dep. var. mean	3,300	0.0690	0.0110	0.0940
Effect relative to mean	-0.43%	7.2%	3.4%	7.4%
Observations	2,670,439	2,670,439	2,670,439	2,670,439
R^2	0.07	0.04	0.04	0.04
	Very preterm	Infant died	10 min. Apgar	5 min. Apgar
	(1)	(2)	(3)	(4)
Noise exposure	0.0017	0.0005*	-0.1072**	-0.0468*
	(0.8931)	(1.8233)	(-2.4862)	(-1.8572)
Dep. var. mean	0.0250	0.0010	8.9000	8.9000
Effect relative to mean	6.8%	48%	-1.2%	-0.53%
Observations	2,670,439	2,670,439	1,960,611	1,960,153
R^2	0.04	0.03	0.07	0.06

Notes: The table shows the results of estimating equation 2. Each column corresponds to a regression with a different component of the infant health index as the dependent variable: an indicator for low birth weight (< 2500 grams), an indicator for very low birth weight (< 1500 grams), an indicator for preterm birth (pregnancy shorter than 37 weeks), an indicator for very preterm birth (pregnancy shorter than 34 weeks), an indicator for infant death, 10-minute Apgar score, 5-minute Apgar score, and birth weight in grams, respectively. All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley t-statistics calculated with a bandwith of 2.7 km are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 5: Robustness to Sample Restrictions

	Infant health index		
	(1)	(2)	(3)
Noise exposure	-0.042***	-0.043***	-0.039***
	(0.013)	(0.013)	(0.012)
Sample	Full	Between rail and highway	< 4 km from rail
Observations	1,960,153	1,427,874	1,587,527
R^2	0.06	0.06	0.06

Notes: The table shows robustness to variations in the estimation sample. Column (1) repeats the main results from Table 2, column (4). Column (2) includes only locations between electric rail lines and highways when there is a highway within five km of an electric rail line. Column (3) keeps only observations less than four km from the nearest electric rail line (the main sample includes observations up to five km from the nearest line). All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 6: Robustness to Downwind Variable Choice

	Infant health index			
	(1)	(2)	(3)	(4)
Noise exposure	-0.042***	-0.025	-0.031**	-0.031
	(0.013)	(0.018)	(0.013)	(0.020)
Downwind measure	Weighted average	Closest only	Union	Range
Observations	1,966,024	1,966,024	1,966,024	1,966,023
R^2	0.06	0.06	0.06	0.06

Notes: Column (1) reproduces the preferred specification. Column (2) considers only the closest rail line to the mother's address. Column (3) takes the union of interior angles between rays connecting the mother's address to rail segment end points. Column (4) takes the union of all 90-degree quadrants, rather than their segment length and inverse-squared distance weighted average, as in column (1). All measures consider only rail lines within 5 km of the mother's address. All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 7: Robustness to Spatial Fixed Effects Specification

	Infant health index			
	(1)	(2)	(3)	(4)
Noise exposure	-0.042***	-0.027***	-0.035***	-0.027***
	(0.013)	(0.009)	(0.010)	(0.009)
Year by month fixed effects	✓	✓	✓	✓
Census block fixed effects	✓			
1 km grid cell fixed effects		✓		
Census block group fixed effects			✓	
Census tract fixed effects				✓
Observations	1,960,153	1,960,098	1,960,153	1,960,153

Notes: All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 8: Robustness to Temporal Fixed Effects Specification

	Infant health index				
	(1)	(2)	(3)	(4)	(5)
Noise exposure	-0.042***	-0.044***	-0.032**	-0.023*	-0.030
	(0.013)	(0.014)	(0.013)	(0.013)	(0.018)
Year by month fixed effects	✓	✓	✓	✓	
Census block fixed effects	✓	✓	✓	✓	✓
Month-Census tract fixed effects		✓	✓	✓	
Census tract fixed effects			✓		
Year-Census tract fixed effects				✓	
Year-Month-Census tract fixed effects					✓
Year × Census tract time trends			✓		
Observations	1,960,153	1,959,701	1,959,701	1,959,155	1,894,312

Notes: All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 9: Robustness to Standard Error Calculation Method

	Infant health index				
	(1)	(2)	(3)	(4)	(5)
Noise exposure	-0.042***	-0.042***	-0.042**	-0.042***	-0.042***
	(0.013)	(0.014)	(0.017)	(0.008)	(0.011)
Standard error calculation	Conley (2.7 km)	Conley (5 km)	Conley (10 km)	Census tract	Census tract by year
Observations	1,966,024	1,966,024	1,966,024	1,966,024	1,966,024
R^2	0.06	0.06	0.06	0.06	0.06

Notes: All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month and census block fixed effects. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 10: Mother Fixed Effects

	Infant health index		
	(1)	(2)	(3)
Noise exposure	-0.023**	-0.042**	-0.037
	(0.010)	(0.017)	(0.024)
Year by month fixed effects	✓	✓	✓
Mother fixed effects	✓	✓	✓
Census tract fixed effects		✓	
Census block fixed effects			✓
Observations	931,868	931,849	921,836

Notes: All specifications include controls for mother race, age, and education; temperature, precipitation, and wind speed; air quality; and year by month fixed effects. The sample used to estimate these regressions includes only mothers who had mulitple births during the sample. Conley standard errors with a 2.7 km bandwidth are shown in parentheses. Asterisks denote p-values: * < 0.10, ** < 0.05, ***: < 0.01.

Appendix Table 11: Alternative Functional Forms for Baseline Performance

Noise measure	This paper		Baselines		
	AEF	10 km AEF	DOT	MF	Pop. Density
	(1)	(2)	(3)	(4)	(5)
Average	0.461	0.461	0.048	-	0.364
Median	0.499	0.519	-	0.439	0.382
Daytime median	0.513	0.464	-	-	0.387
Nighttime median	0.444	0.466	-	-	0.33

Notes: Table compares the out-of-sample predictive performance (R^2) of models developed in this paper (Columns 1-2) against benchmarks (Columns 3-5). Whereas the main text table (Table 5) uses the baselines directly, here the baseline model also include deciles of the prediction values. Each entry is the average out-of-sample R^2 from 5-fold cross-validation. Column (1) uses AlphaEarth Foundations (AEF) model features at their native resolution. Column (2) adds 10 km spatial statistics of all AEF features. Column (3) uses the Department of Transportation's (DOT) National Transportation Noise Map. Column (4) uses Mennitt and Fristrup (2016) (MF). Column (5) uses population density.

B Additional Background Information

B.1 Noise Maps

Noise maps, which estimate average noise levels at specific locations, are commonly used to assess noise exposure. Noise maps are created using numerical models of noise transport given information about noise sources, such as road and rail traffic, as well as the local built environment, topography, and climate (Murphy and King, 2014).

While noise maps may be useful for broadly assessing population-level noise exposures, they have several limitations. Direct noise measurements are only used for validating noise models, if they are used at all (Murphy and King, 2014). The specific modelling software used can also generate significant differences across noise maps conditional on input data, ranging from differences of 2 dB(A) (a small but perceptible difference) to 15 dB(A) (a large difference, similar to the difference in the noise levels of a vacuum cleaner and a lawn mower) (Hepworth, 2006; Nijland and Van Wee, 2005). To a lesser extent, the spatial interpolation method used to generate continuous noise maps from the discrete model outputs can also generate variation across maps created using the same input data (Murphy and King, 2014). In one analysis comparing European noise maps, Arana et al. (2014) find significant differences in estimated exposure across maps calculated with different input data and methodologies, concluding that "the uncertainty associated with the results is such that [...] no conclusions can be drawn concerning the effectiveness of remedial measures."

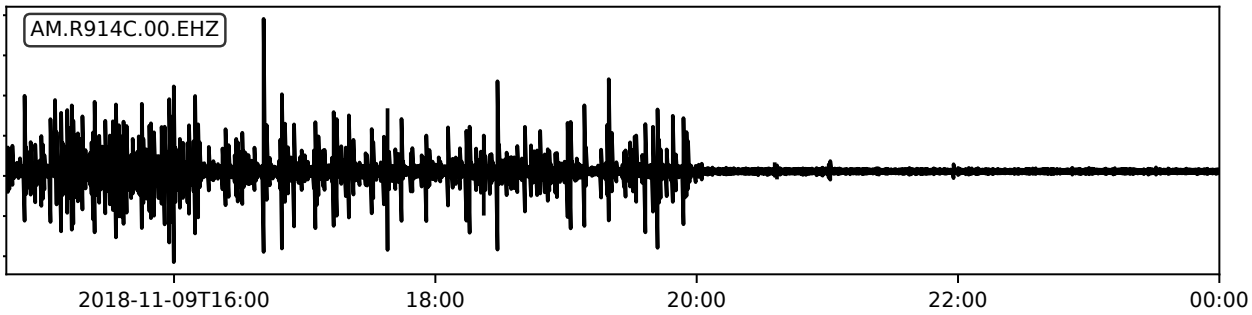
Measurement issues are further compounded by simplifying assumptions modelers make due to data availability or computational complexity. For example, the US National Transportation Noise Map does not account for weather and atmospheric effects, assumes that all ground cover is "acoustically soft," (i.e. is vegetation and not asphalt), and does not account for changes in sound propagation due to natural and man-made features such as hills and buildings (Bureau of Transportation Statistics, 2020). The model is not validated using direct noise measurements, reporting only that "noise levels are evaluated by subject matter experts for confirmation that levels are within a reasonable order of magnitude" (Bureau of Transportation Statistics, 2020). For these reasons, the Web page hosting the National Transportation Noise Map reads, "[the National Transportation Noise Map] is intended for the tracking of trends and should **not** be used to evaluate noise levels in individual locations and/or at specific times" (Bureau of Transportation Statistics, 2020, formatting in original) While it may be possible to sign the bias of particular components of the measurement error, it is difficult to characterize its overall magnitude and direction, or to assess how the error might vary across locations.

C Data Appendix

C.1 Noise Data

C.1.1 Seismic data processing

Appendix Figure 14: Raw seismic waveform data



Notes: Raw seismic waveform data for RaspberryShake monitor R914C, located at 51.74774775° N, 1.251424032° W, Oxford, United Kingdom, on November 9, 2018. The monitor recorded 100 samples per second over this period of just over 9 hours, resulting in about 3.24 million observations for this period alone.

To process the raw seismic data, I follow the methods in Lecocq et al. (2020), which reflect best practices for seismic noise analysis. Seismometers typically measure displacement across multiple different dimensions, such as the 3 physical dimensions, as well as higher moments of displacement, such as acceleration. I work with vertical acceleration, which is most commonly available and widely used due to its precision, sensitivity at a wide range of frequencies, and comparability with standard models of seismic noise (McNamara and Buland, 2004).

The first step in processing the raw seismic data is computing an acceleration power spectral density (PSD) for each 1-hour period in the data. This yields an empirical distribution of the power of the seismic acceleration signal in specific frequency bands. I then convert from the acceleration PSD, A_{dB}, to the displacement spectral power, D_{pow}, using a Fourier transform for each frequency f:

$$D_{pow}(f) = \frac{10^{\frac{A_{dB}(f)}{10}}}{(2\pi f)^2} \tag{6}$$

Then, using Parseval's identity, I obtain the root mean squared time-domain displacement, $d(t)$, band-passed for a specific frequency range $[f_{min}, f_{max}]$:

$$d(t) = \sqrt{\int_{f_{min}}^{f_{max}} D_{pow}(f) df}. \tag{7}$$

I focus on the frequency range $[1.0, 20.0]$, reflecting the range of frequencies corresponding to anthropogenic noise sources documented in the seismology literature (Bonnefoy-Claudet et al., 2006). To investigate this choice's potential impact on my research design and first stage results, I ran my first stage regression (equation 4) for noise measured in the following frequency ranges: .

C.1.2 Relationship between seismic noise and audible noise

In the acoustics literature, sound levels are typically measured in decibels (dB). Given a sound pressure p, the sound level L_p in decibels is given by:

$$L_p = 20 \log_{10} \left(\frac{p}{p_0} \right), \tag{8}$$

where p_0 is the reference pressure, which in acoustics is typically set at 20 micropascals, pressure of the smallest sound detectable by the human ear. When p_0 is set at 20 micropascals, the decibel is the logarithm of the ratio the pressure of a sound to the pressure of the lowest audible sound to a healthy human ear.[23]

Working with decibels is more convenient that working directly with measures of pressure because of the enormous range of pressures that human ears can detect. For example, the maximum noise level humans can hear without pain is about 140 dB (Murphy and King, 2014). This corresponds to a sound pressure level of about 200 Pascals (200 million micropascals), a value that is 7 orders of magnitude above the hearing threshold value of 20 micropascals. Moreover, the ear responds to sound pressure logarithmically: an increase in 10 dB corresponds to an approximate doubling in perceived loudness (Murphy and King, 2014).

How are sound pressure p and intensity I related to quantities measured by seismometers? Consider an idealized scenario in which a sound wave with pressure p comes into contact with the ground. According to Newton's Second Law, the acceleration of the ground is equal to the ratio of the force exerted by the wave to the mass of the ground. Thus, ground acceleration is proportional to sound pressure, which is proportional to displacement d, as defined in equation (7). I thus define the sound level in decibels as measured by displacement

[23] The formula for decibels in acoustics is sometimes written in its unsimplified form as $L_p = 10 \log_{10} \left(\frac{p}{p_0} \right)^2$. This notation clarifies that the quantities being compares are sound intensities, which are proportional to the square of a sound wave's pressure. Multiplying the log-ratio of intensities by 10 converts from bels to decibels.

analogously to definition of L_p in equation (8):

$$L_d = 20 \log_{10} \left(\frac{d}{d_0} \right) \tag{9}$$

I can use equation (9) to convert changes in displacement to changes in decibels. For example, the increase in noise due to being downwind of an electric rail line is

$$\begin{aligned} L_d^D - L_d^N &= 20 \log_{10} \left(\frac{d^D}{d_0} \right) - 20 \log_{10} \left(\frac{d^N}{d_0} \right) \\ &= 20 \left[\log_{10} \left(d^D \right) - \log_{10} \left(d^N \right) \right] \\ &= 20\beta \end{aligned}$$

for some reference value d_0, where in the last line we substituted in equation (4), where the superscripts D and N correspond to downwind and not downwind, respectively. Thus, converting $\hat{\beta}$ from equation (4) from log-displacement to decibels simply involves re-scaling the coefficient. Note that this conversion does not depend on a choice of reference value d_0.

In practice, the extent to which a seismometer measures audible sound depends on the sensitivity of the seismometer, the loudness and frequency of the sound, the distance between the sound source and the monitor, and the density of the earth's surface, among other factors. There are many reasons why audible and seismic noise data may differ. These include measurement error in either instrument, calibration differences between the instruments, and differences in noise sources at the location of the noise monitor and the location of the seismometer. Seismometers likely better capture low-frequency audible noise sources, which due to their lower wavelength travel farther and are less attenuated when moving from the atmosphere to the ground.

C.2 Infant Health Data

I obtain data on health at birth from the universe of births in California between 1999 and 2021. I obtain these data from California vital statistics files, which I accessed from the California Department of Public Health via a research data request. I use the Birth Statistical Master File (BSMF) for births between 1999 and 2017, and the California Comprehensive Births Master File (CCMBF) for births between 2018 and 2021. I harmonize variables across BSMF and CCBMF to create a single database of containing measures of infant health such as gestation length, birth weight, and Apgar scores, rich parental characteristics, and mothers' home addresses.

Prior to analyzing the data, I take several basic data cleaning steps. I drop births

where the birth date or mother's address was not recorded. Since 2007, the California vital statistics files have recorded obstetric estimates of gestation, which are the attending physician's estimate of gestation length based on prenatal care history. Obstetric estimates of gestation length are generally viewed as more reliable than estimated gestation based on the mother's reported last menstrual period (see Behrman and Butler, 2007, chp. 2, for a review). I thus use the obstetric estimate of gestation length as my preferred estimate when available. Following Tran et al. (2020), I also drop implausible birth weights below 500 grams and above 5500 grams, and implausible gestation lengths shorter than 22 weeks or longer than 44 weeks.

I use standard definitions of preterm birth and low birth weight. Preterm birth is birth before 37 weeks of completed gestation; low birth weight is birth weight less than 2,500 grams, both following WHO definitions (Behrman and Butler, 2007).

C.2.1 Infant health index construction

To construct the infant health index, I follow the procedure detailed in Anderson (2008):

1. For each outcome $\{y_{i1}, y_{i2}, \ldots, y_{iK}\}$, switch signs as necessary so that the positive direction always indicates a "better" outcome.

2. Standardize all outcomes by demeaning them and dividing them by their standard deviations: $\tilde{y}_{ik} = \frac{y_{ik} - \overline{y}_k}{\sigma_k}$.

3. To create the index, $\overline{s}_i$, take a weighted average of the $\tilde{y}_{ik}$s, where the weights are given by the inverse of the covariance matrix of the transformed outcomes. This weighting scheme maximizes the amount of information in the index. The weights used are reported in appendix table 12.

C.2.2 Maternal characteristics

For all covariates, I replace missing values with their own factor level, thus maximizing my sample size.

Mother's age. Mother's age is recorded identically in BSMF and CCMBF. I top code mother's age at 60 and define mothers younger than age 20 as teen mothers.

Mother's education. BSMF tracks mother's years of education whereas CCMBF tracks the mother's educational attainment. I harmonize these variables to create two variables: total years of education completed, and highest level of education attained. To assign years

Appendix Table 12: Weights for Infant Health Index

Variable	Weight
1(Birth weight $<$ 1500 g)	0.10
1(Birth weight $<$ 2500 g)	0.08
1(Gestation $<$ 34 weeks)	0.05
1(Gestation $<$ 37 weeks)	0.14
1(Infant died)	0.24
10-minute Apgar score	0.19
5-minute Apgar score	0.08
Birth weight (g)	0.11

Notes: Table shows the weight assigned to each outcome when calculating the infant health index. Weight are determined by the inverse covariance matrix of the transformed outcomes, as described in appendix section C.2.1.

of education to CCMBF observations, I map the educational attainment levels reported in the CCMBF to years of education following the mapping described in table 13. I bottom code and top code all years of education variables to be between 8 and 17 years to match the coarsest coding in the BSMF. I also use the mapping in table 13 to map years of education from BSMF to educational attainment levels.

Mother's race and ethnicity. CCMBF includes a field that tracks both race and ethnicity. The possible values of this field are non-Hispanic white, non-Hispanic Black, non-Hispanic American Indian/Alaskan Native, non-Hispanic Hawaiian/Pacific Islander, non-Hispanic other, two or more races, any Hispanic single race, and not stated/unknown. Between 2000 and 2017, BSMF includes separate fields tracking race and ethnicity. One field tracks single- and multi-race status (white only, Black only, American Indian/Alaskan native only, Asian only, Hawaiian/Pacific Islander only, other, two or more races, and not stated/unknown), the other tracks ethnicity (Hispanic or not Hispanic, including detail on national origin if Hispanic). I harmonize the 2000-2017 BSMF coding with the CCMBF coding, creating a single field that tracks both race and ethnicity. In 1999, the BSMF did not include a field that combines multi-race codes; I do not attempt to harmonize the raw race fields with the rest of the dataset. This means I do not use race data for observations from 1999.

Appendix Table 13: Mapping of Educational Attainment Levels to Years of Education

Educational attainment level	Years of education
Less than high school	< 12
High school graduate / GED	12
Some college	13
Associate's degree	14
Bachelor's degree	16
Master's degree	17
Professional degree	17
Doctorate	17

Notes: Table shows the years of education assigned to each educational attainment level in the CCMBF. Graduate degrees (master's, professional, and doctorate) are top-coded at 17 to match BSMF coding. CCMBF includes years of education for individuals who have fewer than 12 years of education.

C.3 Other data sources

C.3.1 Rail lines and highways

I obtain data on the location of California electric rail lines from OpenStreetMap (OSM). I scraped these data from OSM using Overpass Turbo in February 2023. I manually verified OSM's accuracy by comparing the downloaded data to official rail system maps and satellite imagery. I augmented the OSM data by manually gathering data on the opening dates and operating hours of each electrified rail segment in California. I obtained opening dates from press releases and news articles. I assume that operational schedules did not significantly change over my sample period and obtain operating hours information from rail schedules that were current as of April 2024. Note that all the electrified rail lines in California belong to mass transit agencies with an exclusive right of way, so operational schedules should capture all traffic on the lines, except for after-hours maintenance work. I allow for different operating schedules on weekdays, Saturdays, and Sundays. I do not account for holiday schedules. The Bay Area Rapid Transit (BART), Los Angeles County Metropolitan Transit Authority (LACMTA), San Diego Metropolitan Transit System (SDMTS), and Valley Transit Authority (VTA) all have less than one hour of variation in operating hours across lines within their transit system. I thus assign the same operating hours to all lines within each system. On the other hand, the San Francisco Muni Metro and Sacramento Rapid Transit systems have significant variation in operational hours across lines. I thus assign operating hours at the line level for these systems.

To identify rail segments that are co-located with highways, I create 50 m buffers around each highway and rail line segment, and label any rail line segment overlapping more than

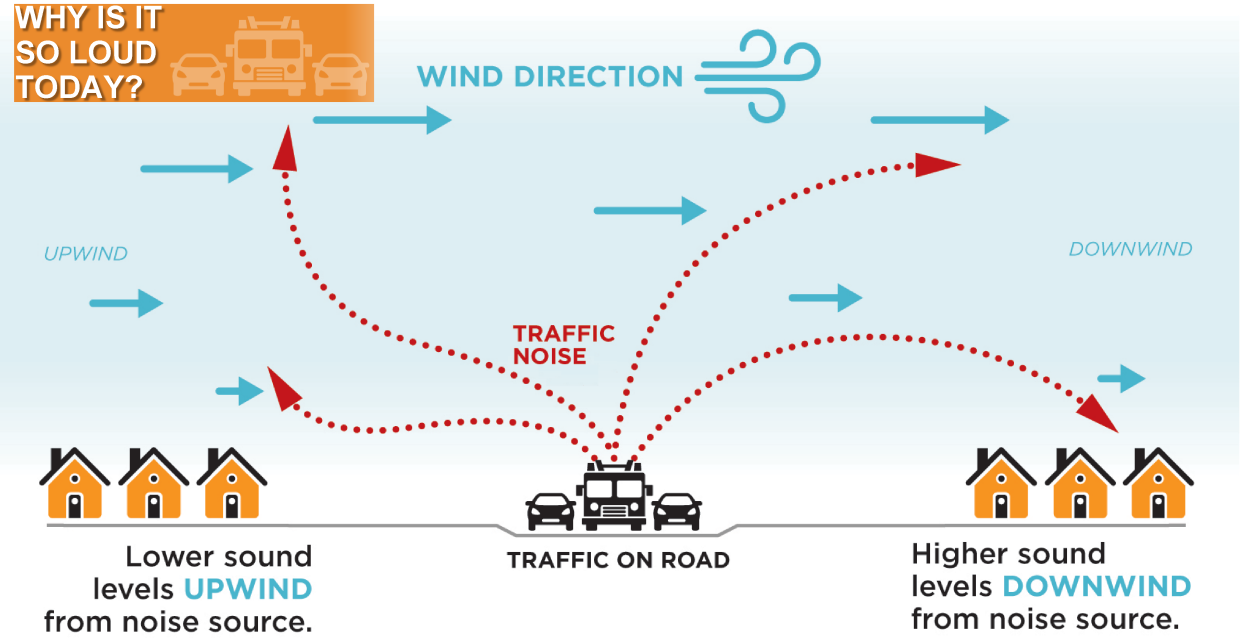
10% of its area with any buffered highway segment as coincident with a highway. I identify highways using interstate and non-interstate highways as well as highways in the Department of Defense's Strategic Highway Netowork in the National Highway System shapefile.

C.3.2 Weather data

D Empirical Framework: Further Details

D.1 Upwind and Downwind Noise Propagation

Appendix Figure 15: Illustration of Downwind and Upwind Noise Propagation

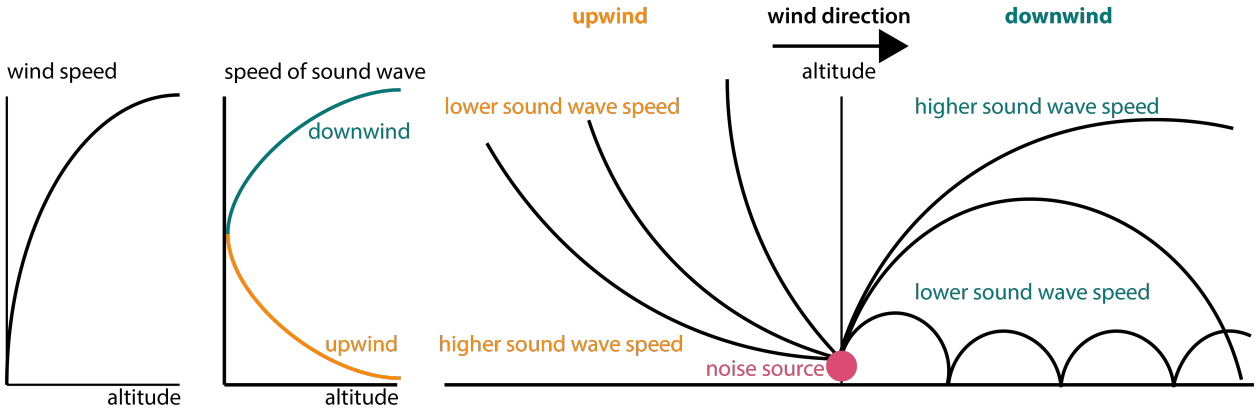


Note: The figure shows differential bending (refraction) of sound waves upwind and downwind of a noise source. Source: Illinois Department of Transportatin (Sperry, 2023).

Appendix Figure 15 reproduces a diagram from a fact sheet released by the Illinois Department of Transportation title "Why is it so loud today?" The fact sheet proposes changes in the wind direction as a leading explanation for hourly, daily, and seasonal changes in noise exposure.

Figure 16 further illustrates the physical mechanism determining the differential propagation of noise upwind and downwind of its source. The far left panel shows the positive vertical gradient in wind speed: wind speeds are higher at higher altitudes. As a sound wave traveling in the downwind direction gains altitude, the increasing wind speed causes it to

Appendix Figure 16: Physical Mechanism Causing Upwind and Downwind Noise Differences



Notes: The figure illustrates how the wind direction affects noise propagation. The far left panel shows the relationship between altitude and wind speed. The second panel shows how the wind speed affects the speed of sound waves, reducing their speed as they gain altitude for sound waves traveling upwind and increasing their speed for downwind waves. The third panel shows how the changing wave speeds cause them to bend (refract), leading to attenuated noise upwind and amplified noise downwind of the source.

move faster, such that the speed of the wave is lower at lower altitudes and higher at higher altitudes. Conversely, a sound wave traveling in the upwind direction will slow down as it gains altitude, as the increasing headwind slows it down. Sound waves bend toward regions where they travel more slowly, causing upwind sound waves to bend away from the ground and downwind sound waves to bend toward it. This mechanism leads to noise levels being attenuated upwind of a noise source and amplified downwind of a noise source (Salomons, 2001; National Academies of Sciences, Engineering, and Medicine, 2018).

D.2 Downwind assignment details

Consider an individual located at some point p at time t. To determine whether they are exposed to elevated electric rail noise due to the wind direction, I must transform the meteorological wind direction θ_{pt} into a binary variable equaling 1 if the point p is downwind of any rail segment $s \in S_{pt}$, and 0 otherwise:

$$\text{Downwind}_{pt} = \begin{cases} 1 \text{ if there exists } s \in S_{pt} \text{ such that } \theta_{pt} \in \Theta_{ps} \\ 0 \text{ otherwise} \end{cases} . \tag{10}$$

Implementing equation 10 proceeds in two steps. First, we must define the set S_{pt} of rail segments to consider. Then, for each segment $s \in S_{pt}$, we must determine the range of wind angles Θ_{ps} for which the point p is downwind of s.

The first step amounts to deciding which rail segments out of all possible rail segments

to consider when determining whether the point p is downwind. I specify the set of such rail segments S_{pt} by applying two criteria. First, I restrict the set of potential rail segments to those within some radius R of the point p. Second, I restrict to only segments that are in operation (i.e., have trains running on them) at time t. A segment may not be in operation at time t if the segment is not yet open (i.e., a rail system expands over the course of the sample, and at time t, trains are not yet running on segment s), or if the segment does not operate at certain times (e.g., if the rail service is shut down on holidays, or in the middle of the night). This yields the following definition of S_{pt}:

$$S_{pt} = \{s : d(s,p) \leq R \text{ and } s \text{ is in operation at time } t\},$$

where $d(\cdot)$ is the Euclidean distance metric.

In the second step, for each rail segment $s \in S$, I define $\Theta_{ps} \subset [0°, 360°) \subset \mathbb{R}$ as the range of angles falling within the interior angle between the line segments s_1p and ps_2, where s_1 and s_2 are the endpoints of the rail segment s. Figure **??** illustrates the calculation of Θ_{ps} for a single rail segment s and for a full set of segments S_{pt}.

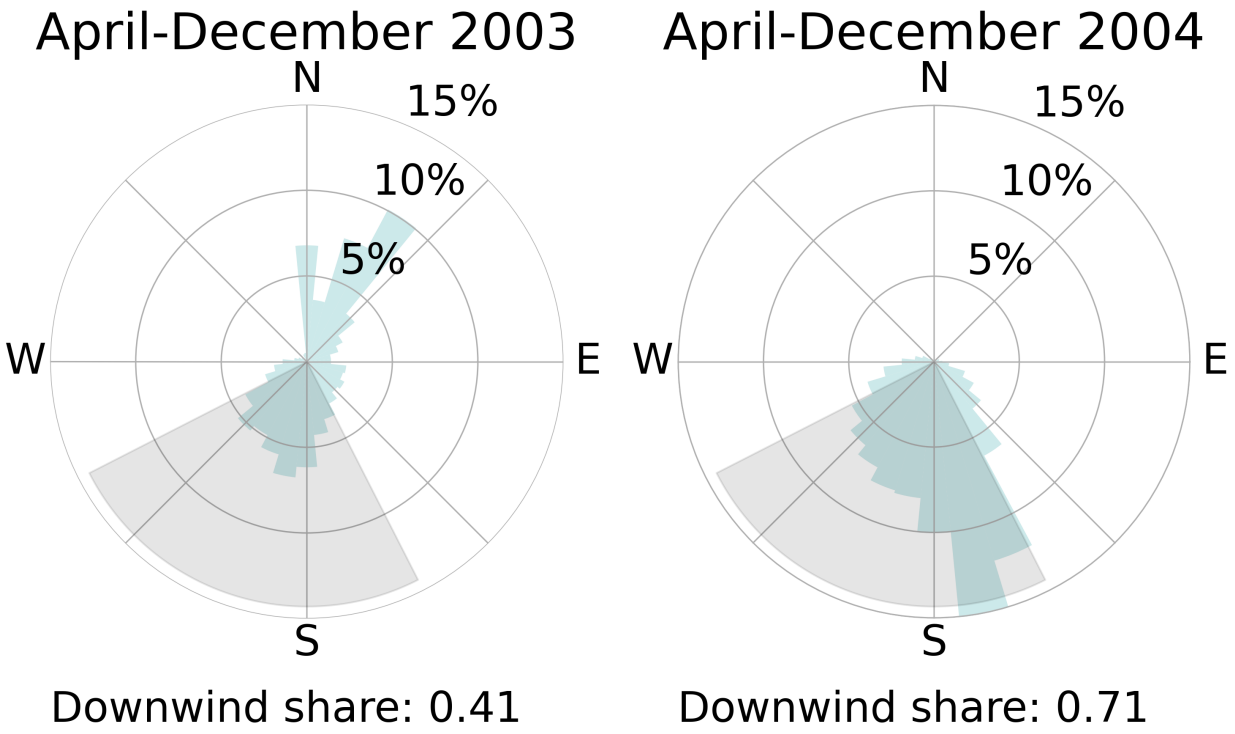
Figure **??** shows the results of applying equation 10 for a hypothetical hour in which the wind is blowing from the west in all locations in the sample $\theta_{pt} = 270° \ \forall p$ and all rail lines are operating. The points plotted correspond to hypothetical addresses, spaced evenly along a 1 square kilometer grid. Each point is within 5 kilometers of at least on electric rail segment. As intended, points to the East of rail segments are classified as downwind, while all other points are classified as not downwind.

D.3 Identifying variation

The research design leverages changes in electric rail noise exposure due to within-census block variation in downwind status across time, conditional on year by month fixed effects. Appendix Figure 17 illustrates this variation for a single location in Pasadena, California, located just north of a LACMTA electric rail line. The left panel describes the distribution of the wind direction and downwind status for a hypothetical nine-month pregnancy between April 1 and December 31, 2023. The right panel shows the distribution of the wind direction for a different hypothetical pregnancy in the same location one year later. There is significant variation the distribution of wind directions across the two periods, resulting in a 30 percentage point difference in the share of the hypothetical pregnancies spent downwind of the electric rail line—and thus potentially exposed to elevated noise.

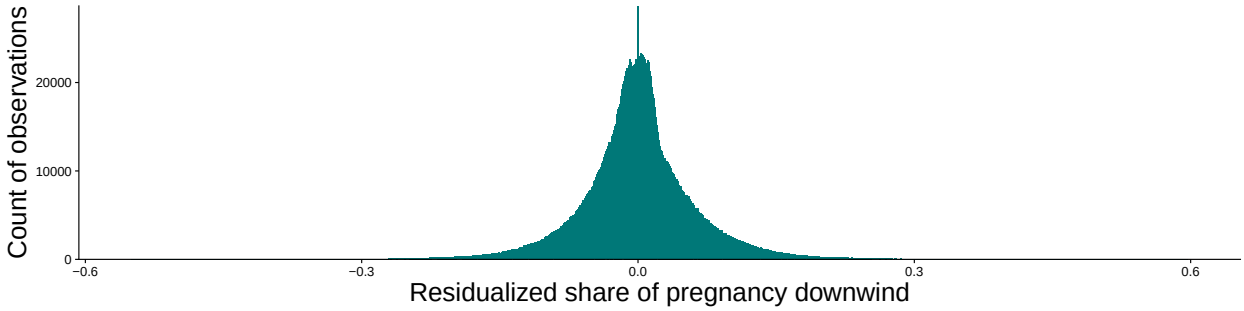
Appendix Figure 18 displays the remaining variation in downwind share after residualizing on census block and year by month fixed effects. The 5th-95th percentile range of the

Appendix Figure 17: Example Variation in Wind Direction Within Location Across Time



Notes: The figure shows the distribution of the wind direction for two nine-month periods at the same location in Pasadena, California. The shaded gray quadrants are the set of wind directions for which the location is downwind of a LACMTA electric rail line. The total share of hours during each nine-month period spend downwind of the electric rail line is shown under each wind rose.

Appendix Figure 18: Residualized Variation in Downwind Shares, Full Sample



Notes: The figure shows the distrubution of the share of pregnancy time spent downwind of an electric rail line for the full sample used to estimate the effect of noise on infant health after residualizing on census block and year by month fixed effects.

distribution is -0.10 to 0.11. The minimum of the distribution is -0.54 and the maximum is 0.60.

What generates variation in downwind status, conditional on census block and year by month fixed effects? Variation in wind direction at this scale will primarily be driven by transient weather systems. For example, winter storms generate changes to dominant wind patterns, with season-to-season variation in their timing, duration, and intensity. Another disruptor to dominant wind patterns is a "cut-off low," a low-pressure system that is detached from the main flow of the jet stream and can linger in a single area for a week or more, causing disruptions to normal wind patterns (e.g., Godsey, 2023). These unpredictable weather patterns cause local variation in wind direction within the same seasons across years, generating variation in downwind noise exposure. Of course, they are also correlated with other weather conditions, such as temperature and precipitation, which may also affect infant health, motivating the use of weather controls in the analysis.

E Results: Sensitivity Analyses

Appendix Table 6 reports results using different approaches to calculating the downwind variable. Column (1) reproduces the preferred specification. Column (2) considers only the closest rail line to the mother's address. Column (3) takes the union of interior angles between rays connecting the mother's address to rail segment end points. Column (4) takes the union of all 90-degree quadrants, rather than their segment length and inverse-squared distance weighted average, as in column (1). All measures consider only rail lines within 5 km of the mother's address. While all point estimates are negative and similar in magnitude, only columns (1) and (3) are statistically significant at conventional levels. This may reflect measurement error introduced by the downwind measures in columns (2) and (4): the measure in column (2) will tend to make exclusion errors, as it considers only a single rail segment, while the measure in column (3) will tend to make inclusion errors, since the unadjusted 90-degree quadrants may include wind directions that are not in fact downwind of an electric rail line.

Appendix Table 7 reports results using different sets of spatial fixed effects. Column (1) reproduces the preferred specification. Columns (2) through (4) 1 km grid cell, census block group, and census tract fixed effects, respectively. The effect of noise on infant health remains statistically precise and of reasonably similar magnitude across these specifications.

Appendix Table 8 reports results using different temporal fixed effects. Column (1) reproduces the preferred specification, which uses year by month fixed effects. Column (2) year fixed effects only, thus not controlling for seasonality in birth outcomes that could

be correlated with downwind exposure. Column (3) includes year and month fixed effects separately. Column (4) includes year by month fixed effects and year by county time trends. Agaubm the estimated effect of noise on infant health remains statistically precise and similar in magnitude across these specifications.

Appendix Table 9 reports the preferred specification using different standard error calculation methods. Columns (1) through (3) use Conley standard erros with 2.7, 5, and 10 km bandwidths, respectively. As expected, the standard error grows as the bandwidth increases, but the estimated effect remains statisically precise at conventional levels for all specifications. Columns (4) and (5) cluster standard errors by census tract and census tract by year, respectively. These standard errors are marginally smaller than Conley standard errors.

Appendix Table 10 reports results for specifications using mother fixed effects. Column (1) includes mother fixed effects but no spatial fixed effect. Much of the variation used to estimate this coefficient will thus come from mothers who moved residential addresses between pregnancies, potentially leading to large changes in exposure. The estimated coefficient is negative, statistically precise, and attenuated relative to the preferred specification. Column (2) adds a census tract fixed effect, which restricts variation to within census tracts. The coefficient is nearly identical to the preferred specification, and is precise at the 5% level (relative to the 1% level for the preferred specification). Finally, column (3) add a census block fixed effect, which mirrors the spatial fixed effect in the preferred specification. While the point estimate remains negative, the standard error is large and the coefficient is not statistically distinguishable from zero in this highly demanding specification.

F Cost Calculation Details

This appendix describes the details of the cost estimates presented in Section 5.3. Biological maturity for life outside the uterus is the primary determinant of infant health at birth, affecting an infant's short- and long-term mortality and morbidity risk, and thus the costs of poor infant health (Behrman and Butler, 2007). Gestational age is viewed as the best proxy for biological maturity, so when possible I focus on the impact of noise on the risk of preterm birth (a measure of gestational age) when calculating costs.

The costs of poor infant health are highly convex in the severity of outcomes. For example, infants born slightly preterm, with between 32 and 36 weeks of gestation, spend on average 5 days after birth in a hospital, incurring average inpatient costs of $18,178, whereas infants born extremely preterm, with under 28 weeks of gestation, spend 76 days

in the hospital, costing $302,455 (Behrman and Butler, 2007).[24] I find that noise increases the risk of moderately severe adverse birth outcomes, such as the likelihood of preterm birth and low birth weight, but does not affect the risk more severe outcome such as very preterm birth and very low birth weight (Figure 5). Using cost estimates that average across the full distribution of adverse infant health outcomes would thus overstate the social cost of noise pollution. To avoid this, I use cost estimates pertaining to the specific outcomes I observe effects for wherever possible.

Table 4 reports 95% confidence intervals capturing the sampling variation in my point estimates. Income estimates also account for the sampling variation in the estimates reported by Bharadwaj et al. (2018). I use the delta method to calculate the standard error of the product of my estimate and that of Bharadwaj et al. (2018). Because I do not have access to Bharadwaj et al. (2018)'s original data, I cannot calculate the covariance between our estimates, and thus assume they are independent.

I consider several measures of noise exposure. First, I calculate the costs due to excess exposure driven by the meteorological conditions I leverage in my research design. To do so, I simply calculate the average share of pregnancy spent exposed to elevated electric rail noise due to being downwind of an electric rail line in my main analysis sample. To measure exposure corresponding to a 1 dB increase in average in utero noise, I use a 50% exposure share, reflecting the result reported in Table **??** that it is on average 2 dB louder per hour spent downwind.

Calculating the total external costs of California electric rail lines due to noise pollution requires two additional steps. First, I need an estimate of the marginal contribution of electric rail to average noise levels. I obtain this estimate from the 2020 US National Transportation Noise Map. While the National Transportation Noise Map has significant measurement error and likely understates total noise exposure (see Appendix B.1), it is useful in this context because it separately estimates the contributions of trains, airplanes, and cars to total noise exposure. Estimates of the total cost of noise using the National Transportation Noise Map are thus likely to be lower bounds. I calculate the marginal contribution of train noise to total noise exposure by differencing the rail noise raster from the airplane and car noise raster. I then subtract 45 dB from this raster, which is the minimum value reported in the National Transportation Noise Map. This step assumes that average noise levels would be 45 dB in the absence of rail noise—a reasonable assumption given that average noise levels in residential noise levels are typically between 40 and 50 dB (Murphy and King, 2014). Second, I calculate the average annual birth rate in each 500 meter grid cell that is exposed to rail

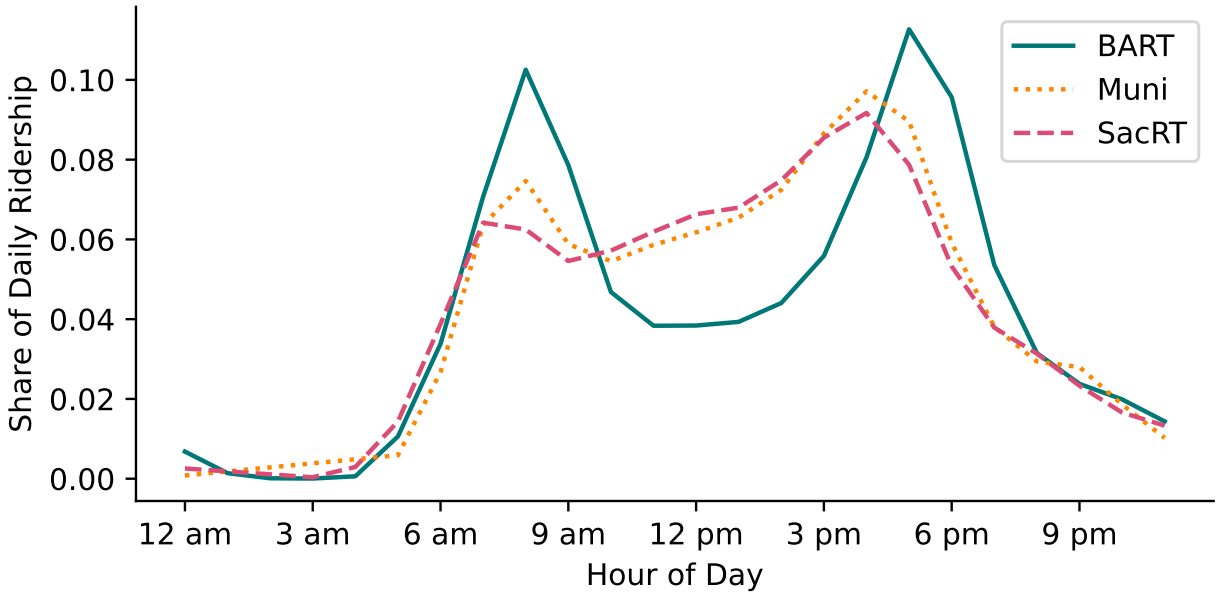
[24]The values in Behrman and Butler (2007) have been adjusted to 2025 dollars using the US Bureau of Labor Statistics' Consumer Price Index.

noise. I use data on births between 2015 and 2019. I choose this period for several reasons. First, birth rates in California and nationally have persistently declined in recent decades (Kearney et al., 2022; Johnson, 2023). Using births data for my full sample starting in 1999 would thus overstate contemporary birth rates. On the other hand, my data also cover the beginning of the COVID-19 pandemic (2020-21), during which birth rates fell, before later rebounding (Kearney and Levine, 2023; Nobles et al., 2024). Using 2020-21 births data to estimate birth rates could thus understate post-pandemic fertility. Finally, I use five years of data, rather than a single year, to obtain a more stable estimate of average contemporary birth rates.

To calculate the cost per overnight passenger, I must count the annual number of electric rail riders during those hours. I submitted public records requests for hourly ridership data to the six rail systems in my sample (BART, LACMTA, Muni Metro, SacRT, SDMTS, and VTA) for 2024, the most recent year available at the time of the requests. Of these, BART, Muni, and SacRT responded with hourly ridership data. To estimate hourly ridership for the other systems, I calculated the share of all BART and Muni Metro passengers between 11:00 p.m. and 6:59 a.m. Approximately 7% of BART, 7% of SacRT riders, and 8% of Muni Metro passengers take the train during those hours. I assume that overnight ridership constitutes 8% of total ridership for the three systems I do not directly observe hourly ridership data for, and multiply 2024 ridership totals, which are publicly available, by this percentage to estimate the number of overnight riders in 2024 in each system. In total, I estimate that there 9,710,544 rides between 11:00 p.m.-6:59 a.m in 2024. Note that this counts total rides, not unique passengers.

Figure 19 plots hourly ridership shares for BART, Muni, and SacRT. All three systems have lowest ridership in the overnight hours and peak during commute hours. BART follows a more extreme daily pattern, with large peaks in the 8 a.m. and 5 p.m. hours and a significant reduction in ridership in between. Muni and SacRT have local peaks in the 8 a.m. and 7 a.m. hours, respectively, then gradually increase hourly ridership until peaking in the 4 p.m. hour.

Appendix Figure 19: Hourly Ridership by Transit System



Notes: The figure plots the share of total ridership by transit system using 2024 ridership data from BART, Muni, and SacRT. Data were obtained via public records requests to each transit agency.