Does air pollution explain geographic variation in healthcare expenditure?*

Lieke Beekers
Leiden University

Larissa de Lima Almeida
Leiden University

Rik Rozendaal
Leiden University

February 13, 2026

Abstract

This paper studies the role of air pollution as a driver of geographic variation in healthcare expenditure in the Netherlands. Using microdata and a mover design building on Finkelstein et al. (2016), we decompose the variation into three components: supply, pollution and patient characteristics. Pollution is part of the place effect but affects healthcare through the demand side. We find that pollution explains approximately 5% of variation across municipalities and 11.5% across neighborhoods. Incorporating air pollution in the model increases the share of variation attributed to place while leaving the share that is attributed to supply factors unchanged. Most of the variation in healthcare expenditure is explained by patient characteristics. Finally, we show that moving to a neighborhood with a 10% higher concentration PM2.5 persistently increases healthcare spending by 2%.

Keywords: Healthcare expenditure, air pollution, mover design
JEL Codes: H51, I1, I11

*Results based on calculations by Leiden University using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl

1 Introduction

A growing literature documents substantial regional variation in per capita healthcare spending, health, and life expectancy, and shows that causal effects of place explain a meaningful portion of these differences, independent of sorting (Finkelstein et al., 2016, 2021). These place effects are generally attributed to the supply side of the market for healthcare, i.e., physicians' preferences and access to healthcare. At the same time, a large and growing literature shows that local air pollution has detrimental effects on health and mortality, thereby increasing healthcare expenditure (Deschenes et al., 2017; Deryugina et al., 2019). Given that there exists substantial geographic variation in air quality (e.g., Colmer et al., 2020), this paper asks the question: to what extent does local air pollution explain regional variation in healthcare expenditure?

We answer this question using detailed microdata from *Statistics Netherlands*. Our identification approach follows Finkelstein et al. (2016) and relies on movers, i.e., people who move from one region to another during our sample period. Upon moving, they are faced with a different level healthcare supply but potentially also with a different level of air pollution, which may in turn affect their demand for care. By studying their healthcare spending pre- and post-move, we isolate the place effects of pollution and healthcare supply separately, allowing us to decompose the total variation into three factors: supply, air pollution, and patient characteristics (i.e., sorting).

We find that supply factors explain around 35% of the variation in healthcare spending at the municipality level, while pollution explains around 5%. That yields a place effect of around 40%, implying that the remaining 60% is due to sorting. At the neighborhood level we find that supply factors are much less relevant at around 5% while becomes the dominant driver of the place effect at around 11.5%. Sorting into neighborhoods is the main driver of healthcare spending variation, however, explaining over 80%. We find that, because pollution levels and average healthcare spending are uncorrelated in the cross-section, not accounting for air pollution yield a correct estimate of the supply effect but an underestimate

of the place effect. Finally, we also find that moving to a neighborhood with a 10% higher concentration of fine particulate matter (PM2.5) persistently increases healthcare spending by 2%.

Understanding the drivers of geographic variation in healthcare spending is important because different explanations for geographic variation imply different policy recommendations. If supply factors such as access to care and doctors' preferences and incentives explain a large share of the variation, then targeted policies that address those factors can yield large societal benefits. If, instead, pollution is a key driver, then targeted environmental policies may be more efficient. If neither supply nor pollution explains much variation, and differences are mainly due to sorting by patients, then the policies mentioned above may prove ineffective in reducing healthcare spending, and policies aimed at prevention may be preferable.

We are the first to connect the literature on regional variation in healthcare expenditure to the literature on the health costs of air pollution. In doing so, we make several contributions to both literatures. First, to the literature on regional variation in healthcare spending, we introduce pollution as a potential explanatory factor. This offers a new perspective because it breaks the reasoning that supply effects are only due to place effects and that demand effects are only due to sorting. Pollution affects spending through the demand for healthcare but it local and thus part of the place effect. We show that including pollution in the analysis meaningfully changes the shares of variation that are attributed to place and to sorting, but that the share that is attributed to supply factors is unaffected.

Second, we contribute to the literature that studies the causal effects of air pollution by estimating this relationship using an entirely different identification strategy. Rather than relying on variation in the weather or in policies that cause variation in air pollution, we rely on movers. That is, we identify the health effect of air pollution by studying people who move to a place that has better or worse air quality while controlling for differences in healthcare supply between destination and origin. This approach also allows us to study

the longer term effects of exposure to air pollution, something that is difficult with other identification designs.

A growing literature documents differences in healthcare spending and life expectancy between regions in the United States and decomposes this variation into place effects and differences in patient characteristics that lead to sorting (Deryugina and Molitor, 2021). Finkelstein et al. (2016), for instance, show that 50-60% of variation in healthcare spending is due to the causal effect of place. Finkelstein et al. (2021) show that place has a large causal impact on mortality, and Ding (2026) shows that 45% of variation in mental health treatment utilization is due to place. Salm and Wübker (2020), on the other hand, find that variation in healthcare spending in Germany is overwhelmingly explained by sorting. For the Netherlands, Moura et al. (2019) show that 30% of the variation across provinces is due to place effects. We find a similar place effect at the municipality level that is mostly due to supply factors. Pollution plays a key role at the neighborhood level, but a limited role at the municipality level.

A large literature documents the adverse health and mortality effects of air pollution. Early studies focused on infants (Chay and Greenstone, 2003; Currie and Neidell, 2005), while later work has focused on entire populations and the elderly. Concerns about people sorting are typically mitigated through the use of weather-related instrumental variables such as wind direction (Deryugina et al., 2019; Anderson, 2020) or temperature inversions (Graff Zivin et al., 2023), through policy interventions Deschenes et al. (2017) or through other exogenous changes in pollution levels (Schlenker and Walker, 2016). Attention has recently shifted towards estimating long-term health effects (Deryugina and Reif, 2023; Lubczyk and Waldinger, 2025) and healthcare costs of air pollution (Basaglia, 2023; Barwick et al., 2026). We add to this literature by introducing a different identification strategy and by studying heterogeneity in the effect of air pollution along previously unexplored dimensions.

This paper continues as follows. Section 2 develops the theoretical framework that guides

our analysis. It builds a simple model of the drivers of healthcare expenditure, decomposing differences into a patient effect, a pollution effect and a supply effect. It also develops our main empirical specification and discusses the implications of leaving out air pollution in an omitted variable analysis. Section 3 introduces the data that we use. Section 4 discusses our results, and Section 5 concludes.

2 Theoretical framework

2.1 Model setup

In studying the drivers of regional variation in healthcare expenditure we build on the work by Finkelstein et al. (2016). In particular, we add air pollution to their theoretical framework to get an insight into how it drives variation in expenditure and to derive our main empirical specifications.

Patient i in year t uses healthcare y_{it}. Like in Finkelstein et al. (2016), we assume that patients differ along three dimensions: health status h_{it}, preference parameter η_i and their geographic location j. We assume that expected utility is $u(y|h_{ijt}, \eta_i) = -\frac{1}{2}(y - h_{ijt})^2 + \eta_i y$. Expected utility is maximized at $y^*_{ijt} = h_{ijt} + \eta_i$. This is the level of care chosen by a patient that is fully informed and faces no price for care. Note that, unlike Finkelstein et al. (2016), we assume that there is a place component to patient health which is due to local pollution.

We assume that the expectation of y^*_{it} depends only on a patient fixed effect, a vector of patient-specific observables x_{it} and observable local pollution P_j: $\mathbb{E}[y^*_{it}|i, j, t, x_{it}, P_j] = \alpha_i + x_{it}\beta + \theta P_j$, where θ is a parameter that regulates the health effect of pollution. We assume here for simplicity that pollution varies by region but not over time, that it enters linearly, increasing healthcare use, that it affects all patients equally, and that it does not interact with patient characteristics x_{it}. These assumptions are quite restrictive but do allow us to demonstrate our point clearly.[1]

[1] We will relax these assumptions in future versions of the paper.

Like Finkelstein et al. (2016), we assume that each patient is matched with a representative physician in their region who determines the patient's care y_{it} to maximize perceived patient utility $\tilde{u}_j(y)$ minus their net private costs of care provision $PC_{jt}(y)$. Hence,

$$y_{ijt} = \arg\max_y \tilde{u}_j(y|h_{it}, \eta_i) - PC_{jt}(y). \tag{1}$$

The difference between perceived and true utility captures differences in beliefs between physicians, such that $\tilde{u}_j(y|h_{it}, \eta_i) = u_j(y|h_{it}, \eta_i) + \lambda_j y$ where λ_j represents physicians' practice style. We follow Finkelstein et al. (2016) and assume that PC_{jt} is linear in y and additively separable in j and t.

Maximizing (1) gives our estimation equation of interest:

$$y_{ijt} = \alpha_i + \gamma_j + \tau_t + x_{it}\beta + \theta P_t + \varepsilon_{ijt}, \tag{2}$$

where $\gamma_j + \tau_t = \lambda_j + PC'_{jt}()$ comes from the supply side of healthcare, $\alpha_i + x_{it}\beta + \theta P_t = h_{ijt} + \eta_i$ comes from the demand side and $\mathbb{E}[\varepsilon_{ijt}|i, j, t, x_{it}, P_t] = 0$.

2.2 Healthcare expenditure decomposition

The key difference between our analysis and Finkelstein et al. (2016) is that we allow for a place component in healthcare demand that is due to local pollution. This implies that we can decompose the variation in healthcare expenditure into three components. These components are (i) a supply-side component attributable to place, (ii) a demand-side component attributable to patients' health and (iii) a demand-side component attributable to place that is due to pollution.

Following the decomposition in Finkelstein et al. (2016), we denote by $\overline{y}_{jt}$ the expectation of y_{ijt} within area j, i.e. $\overline{y}_{jt} = \tilde{y}^*_{jt} + \tau_t + \gamma_j + \theta P_j$, where $\tilde{y}^*_{jt}$ is the expectation of the patient-optimal level of care for a pollution level of 0. Let $\overline{y}_j$ denote the average $\overline{y}_{jt}$ over time: $\overline{y}_j = \tilde{y}^*_j + \gamma_j + \theta P_j$, where $\tilde{y}^*_j$ is the time-average of $\tilde{y}^*_{jt}$. We can then decompose the difference

in average healthcare expenditure between regions j and j' into three components:

$$\overline{y}_j - \overline{y}_{j'} = \overbrace{\underbrace{\tilde{y}^*_j - \tilde{y}^*_{j'}}_{\text{Patient}} + \underbrace{\theta(P_j - P_{j'})}_{\text{Place}}}^{\text{Demand}} + \overbrace{\gamma_j - \gamma_{j'}}^{\text{Supply}}, \tag{3}$$

which yields the following three shares:

$$S_{\text{supply}}(j, j') = \frac{\gamma_j - \gamma_{j'}}{\overline{y}_j - \overline{y}_{j'}}, \tag{4}$$

$$S_{\text{patient}}(j, j') = \frac{\tilde{y}^*_j - \tilde{y}^*_{j'}}{\overline{y}_j - \overline{y}_{j'}}, \tag{5}$$

$$S_{\text{pollution}}(j, j') = \frac{\theta(P_j - P_{j'})}{\overline{y}_j - \overline{y}_{j'}}. \tag{6}$$

Comparing our decomposition to the one in Finkelstein et al. (2016), a striking difference stands out. We add pollution to the analysis, which affects the demand side through place. Hence, we can no longer simply interpret supply effects as place effects and demand effects as patient effects. We distinguish between a supply component, a patient component and a pollution component. It is an open question to what extent accounting for pollution will reduce the share of variation that is attributed to the patient component and the supply-side component, respectively.[2]

2.3 Event-study design

Building again on the work by Finkelstein et al. (2016), we develop an event-study representation of (2). We denote by δ^y_i the difference in average healthcare expenditure between mover i's origin and destination region, $\delta_i = \overline{y}_{d(i)} - \overline{y}_{o(i)}$, and by δ^P_i the difference in pollution, $\delta^P_i = P_{d(i)} - P_{o(i)}$. We then rewrite (2) as follows,

$$y_{it} = \alpha_i + \gamma_{o(i)} + \theta P_{o(i)} + I_{r(i,t)>0}\Big(S^i_{\text{supply}}\delta^y_i + \theta\delta^P_i\Big) + \tau_t + x_{it}\beta + \varepsilon_{it}, \tag{7}$$

[2]Note that the three shares above sum to 1 but that individual shares can be negative or greater than 1 as the numerator and denominator can have opposite signs.

where $I_{r(i,t)>0}$ is an indicator variable for year relative to move greater than 0. We then collapse $\alpha_i+\gamma_{o(i)}+\theta P_{o(i)}$ into a single patient fixed effect $\tilde{\alpha}_i$, and parameterize the interactions around the move as flexible functions with relative-year specific coefficients $\mu_{r(i,t)}$:

$$y_{it} = \tilde{\alpha}_i + \mu^y_{r(i,t)}\delta^y_i + \mu^P_{r(i,t)}\delta^P_i + \tau_t + x_{it}\beta + \varepsilon_{it}. \tag{8}$$

Like in Finkelstein et al. (2016), this is the equation that we estimate. Under the working assumptions of our model, the jump around the move in μ^y identifies the weighted average S_{supply}. The jump in μ^P identifies the healthcare expenditure effect of pollution, θ. We then compute the pollution component in our decomposition as the weighted average of the ratio of pollution differences divided by average expenditure differences between destination and origin, multiplied by our estimate of the effect of pollution on healthcare spending. That is, $\hat{S}_{\text{pollution}} = \frac{1}{N}\sum_i \frac{\theta\delta^P_i}{\delta^y_i}$, where N is the total number of movers i in our sample. Subsequently, this allows us to compute $\hat{S}_{\text{patient}} = 1 - \hat{S}_{\text{supply}} - \hat{S}_{\text{pollution}}$.

2.4 Pollution as an omitted variable

Comparing our framework to Finkelstein et al. (2016), the key difference is that we add a variable to the analysis. This subsection aims to structure our thinking on what leaving out air pollution means for the estimates of the supply and patient components that explain regional differences in healthcare spending.

Although we model pollution as a factor that affects healthcare spending through the demand side, i.e., through patients' health, it may be correlated with average healthcare spending in a patient's region. Hence, leaving it out would introduce omitted variable bias into our estimates of the supply effects. Since the patient component is the share of variation that cannot be explained by place, if the place effects are biased, then so is the demand effect. In other words, overestimating one component will lead to an underestimation of the other.

The identification approach explained above relies on estimating place effects $\hat{\gamma}_j$ by study-

ing movers across regions. To formalize our reasoning, let us write down an unrestricted and a restricted version of (8), replacing the model parameters with our estimates of them, leaving out the individual and time fixed effects and the patient characteristics for simplicity, and restricting our attention to the jump at 1 year after the move,

$$y_{it} = \hat{\mu}^{y}_{r(i,t)=1}\delta^{y}_{i} + \hat{\mu}^{P}_{r(i,t)=1}\delta^{P}_{i} + \varepsilon_{it}, \tag{9}$$

$$y_{it,res} = \hat{\mu}^{y}_{r(i,t)=1,res}\delta^{y}_{i} + \varepsilon_{it,res}. \tag{10}$$

The restriction here is that (10) omits pollution differences in estimating $\hat{\mu}^{y}_{1,res}$ (that is, the jump in healthcare spending one year after the move). Following the established literature on omitted variable bias (e.g., Cinelli and Hazlett, 2019), we can define the omitted variable bias as the difference between our restricted estimate and the "true" coefficient: bias $= \hat{\mu}^{y}_{1,res} - \hat{\mu}^{y}_{1}$. Under the assumption that both δ^{y} and δ^{P} are orthogonal to the covariates and fixed effects, the restricted estimate of the jump can then be rewritten as

$$\hat{\mu}^{y}_{1,res} = \hat{\mu}^{y}_{1} + \hat{\mu}^{P}_{1}\frac{\mathrm{Cov}(\delta^{y}_{i}, \delta^{P}_{i})}{\mathrm{Var}(\delta^{y}_{i})}. \tag{11}$$

That is, the bias in our estimate of the jump in healthcare spending after the move if pollution is omitted, which is our estimate of S_{supply}, depends crucially on the covariance between δ^{y} and δ^{P}. If average healthcare expenditure and pollution are uncorrelated, then $\hat{\mu}^{y}_{1,res}$ is an unbiased estimate of S_{supply}. In that case, any share of the regional variation in healthcare expenditure that is explained by pollution goes at the expense of the patient effect. In other words, the supply effect and the demand effect, which consists of patient and pollution, are estimated accurately even if pollution is omitted. However, if average healthcare spending and pollution are correlated, then omitting pollution yields a biased estimate of the supply effect. The place effect in our framework, which consists of supply and pollution, cannot be estimated accurately without accounting for pollution.

3 Data

3.1 Healthcare costs

For our healthcare costs data we draw on Dutch administrative data from *Statistics Netherlands*, covering the period $2012 - 2022$. Our data contain information on annual health care expenditures for all Dutch residents covered by the mandatory basic health insurance scheme, effectively capturing the entire population since participation in this insurance system is compulsory. Recorded expenditures include reimbursed health care costs and mandatory deductible payments.[3] Expenditures can be disaggregated into major categories of care, including general practitioner services, hospital care, and pharmaceutical care. Our current analysis includes the entire Dutch population of individuals who are older than 65.[4]

3.2 Residential information

In each year from 2012 to 2022, we observe an identification number which is unique to a particular address of residence at the neighborhood level. These addresses consist of all available addresses in the Geographical Basis Registration, the Municipal Basis Registration, and the Housing Register. We link this identification number to individuals' neighborhood (and municipality) of residence. We use the definition of neighborhoods as established in 2023. We define movers as individuals who have relocated from another neighborhood (municipality).

3.3 Air pollution

We use fine particulate matter (PM2.5) as a measure of air pollution. PM2.5 refers to inhalable particles with a diameter of less than 0.0025 millimeters. Due to their small size, these particles can penetrate deep into people's lungs and even into their bloodstream,

[3]Each individual insured under the basic scheme is obligated to pay a fixed annual amount out-of-pocked before the coverage applies.

[4]We observe the entire population and we will add analyses on the younger generations in a later version of the paper.

causing respiratory and cardiovascular failure.[5] This makes PM2.5 particularly harmful for public health, especially among those with pre-existing health conditions. Fine particulate matter is emitted directly in combustion and forms in the atmosphere in chemical reactions of other pollutants, such as sulfur dioxide and nitrogen oxides. Its concentration is correlated with the concentrations of other pollutants (e.g., Basaglia, 2023). Our results should thus not be interpreted as the isolated effect of PM2.5 but as the effect of air pollution more broadly, measured through the concentration of fine particulate matter.

Our data on PM2.5 comes from the Dutch National Institute for Public Health and the Environment (RIVM). We rely on annual average pollution concentration maps with a resolution of 1km×1km. These maps are produced using the reanalysis model OPS (Operational Priority Substances), which is developed and run by RIVM and affiliated scientists. OPS takes data on emissions, weather and land use as inputs and models dispersion, transport and chemical reactions. It is calibrated using pollution monitor data.[6]

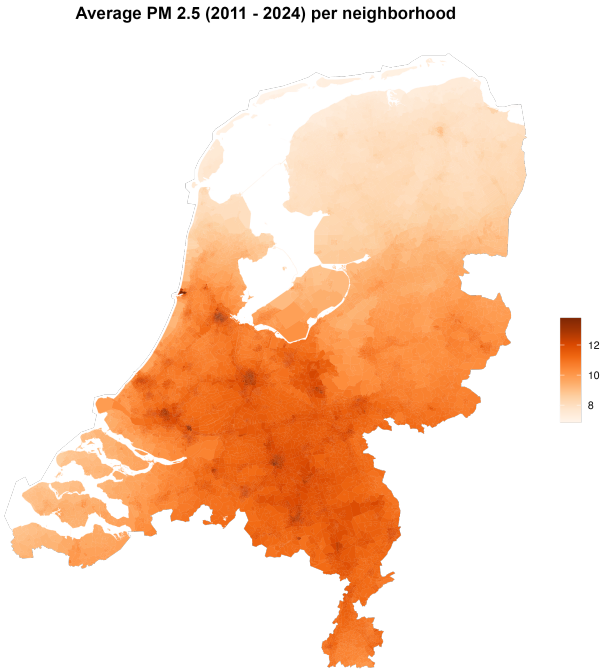
We then combine the PM2.5 concentration maps from RIVM with the regional classification of neighborhoods from Statistics Netherlands to compute weighted average PM2.5 concentrations by neighborhood.[7] We then take the time-average of pollution by neighborhood and match this to people's neighborhood of residence in the Statistics Netherlands secure microdata environment. Figure 1 maps the concentration of PM2.5 in the Netherlands between 2012 and 2024 at the neighborhood level.

[5]For more information on PM2.5, see, for instance, `https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm`.

[6]The documentation of the OPS model is available here: `https://www.rivm.nl/media/ops/OPS-model.pdf`. The resulting maps can be found here: `https://www.rivm.nl/gcn-gdn-kaarten/concentratiekaarten/downloaden/archief`.

[7]The regional classifications are available here: `https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/cbs-gebiedsindelingen`. We use the classification for 2023. Neighborhood averages are weighted by the share of the neighborhood area that overlaps with a given 1km×1km cell on the pollution map.

Figure 1: Regional variation in air pollution in the Netherlands



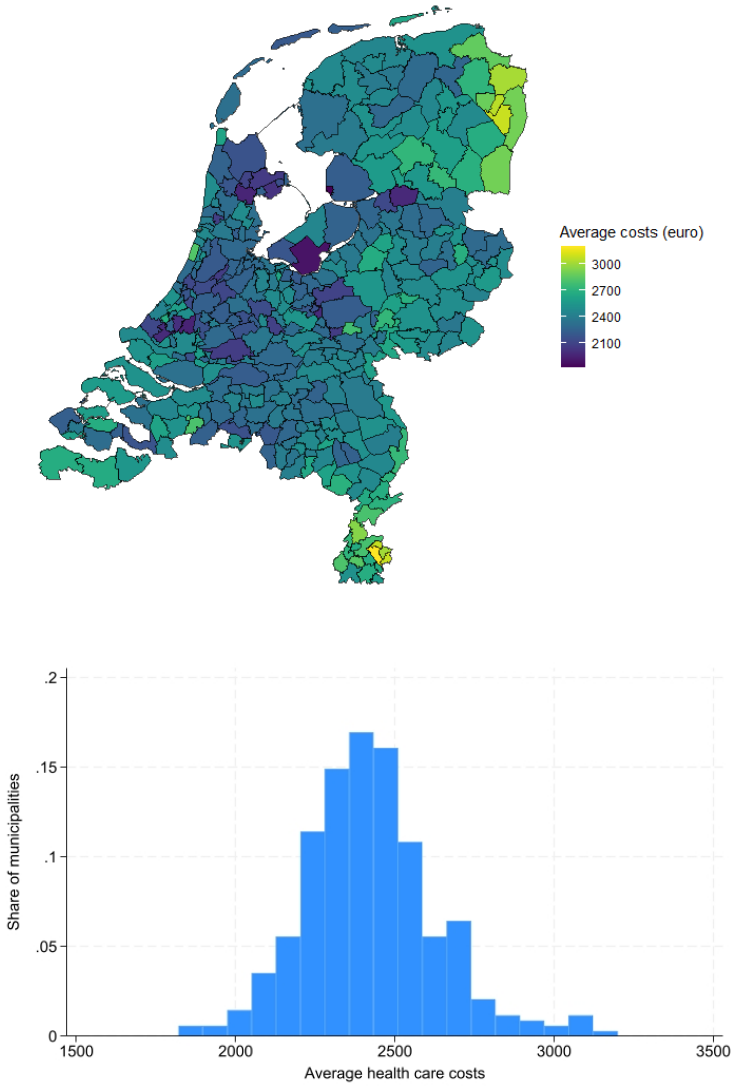
Notes. This map documents the average level of fine particular matter (PM2.5), between 2011 and 2024, at the neighborhood level. Darker (lighter) red areas correspond to higher (lower) levels of PM2.5.

3.4 Sample restrictions and descriptive statistics

When we compute the average health care costs at the regional level (i.e., municipalities and neighborhoods), we omit movers in their move year. All average health care costs are computed by first averaging health care costs across individuals in each year and then taking a simple average across years. In other words, we use the entire Dutch population (i.e., stayers and movers) to calculate the average health care costs across regions. When we estimate our event-study as represented in equation (8), we only include one-time movers in our sample. We define a one-time movers as an individual who has moved exactly once throughout $2012 - 2022$. For our event-study, we are left with 843,591 one-time movers at the neighborhood level, and 381,560 one-time movers at the municipality level.

Figure 2 shows the distribution of average health care costs across municipalities in our final sample. There are substantial differences in average spending across municipalities,

Figure 2: Distribution of health care costs across municipalities



Notes. The map shows the regional variation in health care costs across municipalities. The histogram shows the distribution of average health care costs across municipalities. The sample include all movers and non-movers, but excludes the health care costs during the year of move for movers.

where municipalities in the North (Groningen) and South (Limburg) spend on average approximately 3000 euros per year on health care, as compared to approximately 2200 euros in the West (Randstad).

The upper panel in Figure 3 shows the distribution of δ_i^y, which denotes the difference in average log health care expenditures between a one-time mover's destination and origin,

where a move is defined at the municipality level. Similar to the US, these differences are distributed symmetrically around 0 (Finkelstein et al., 2016). Interestingly, the same pattern is observed for the distribution of δ_i^P, which denotes the difference in log air pollution levels between a mover's destination and origin. Both statistics indicate that equally many individuals move to areas with higher pollution or health-care costs as to areas with lower levels.

4 Results

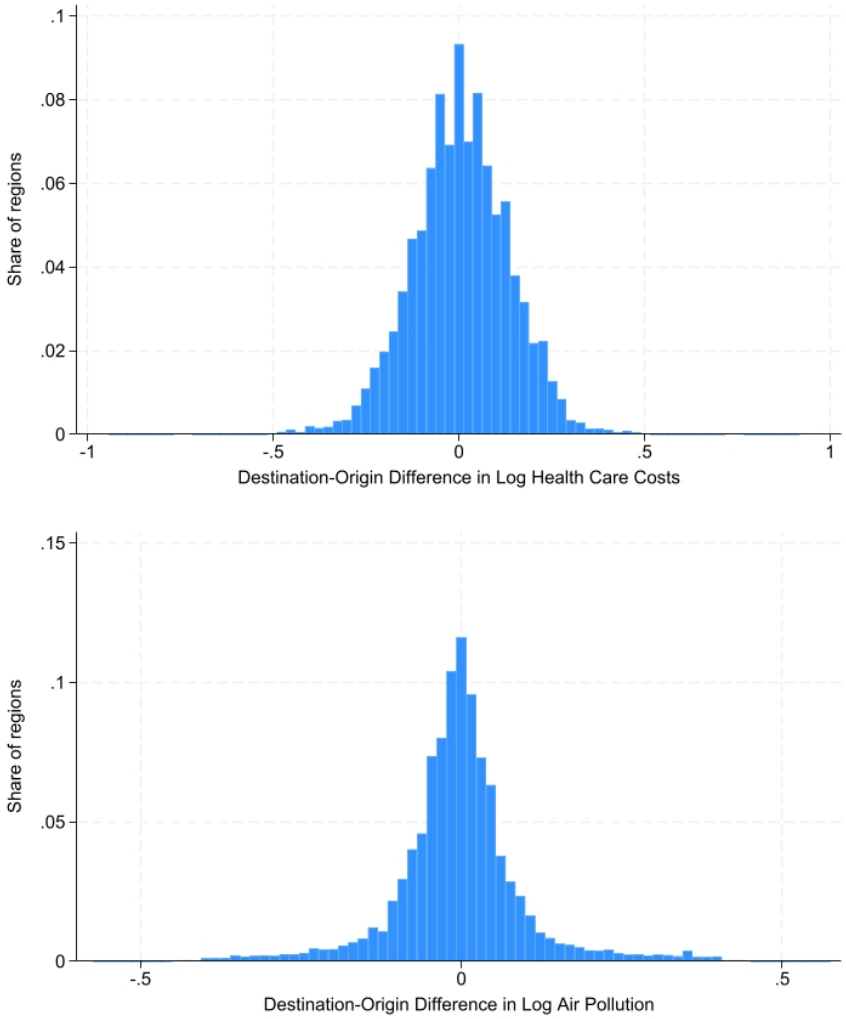
4.1 Supply effects

Figure 4 shows our results of the event study design for healthcare spending, i.e. our estimates of μ^y in (8), which tell us what share of the total variation in healthcare expenditure is due to supply factors. Several things stand out. First, we find no meaningful difference between the estimates when we exclude air pollution (in blue) and when we include it (in red), which indicates that including air pollution does not meaningfully alter our estimate of the supply effect. This finding indicates that the omitted variable bias discussed in Section 2.4 is negligible.

Second, the figure shows a sharp discontinuous jump at the time of the move, which aligns with the event-study results as documented for the US (Finkelstein et al., 2016) and the Netherlands (Moura et al., 2019). Both at the municipality level and at the neighborhood level the effect roughly stabilizes after year 1, indicating a persistent change healthcare spending. There is a slight downward trend in our coefficient estimates before the move, but the last 5 years pre-move are not significantly different from 0.

Third, we find that supply factors explain much more variation at the municipality level than at the neighborhood level. Our estimate of the supply effect at the municipality level is 30 to 40%, whereas it is only around 5% at the neighborhood level. This is not surprising as healthcare supply factors are likely to vary mostly by hospital, while we can expect them

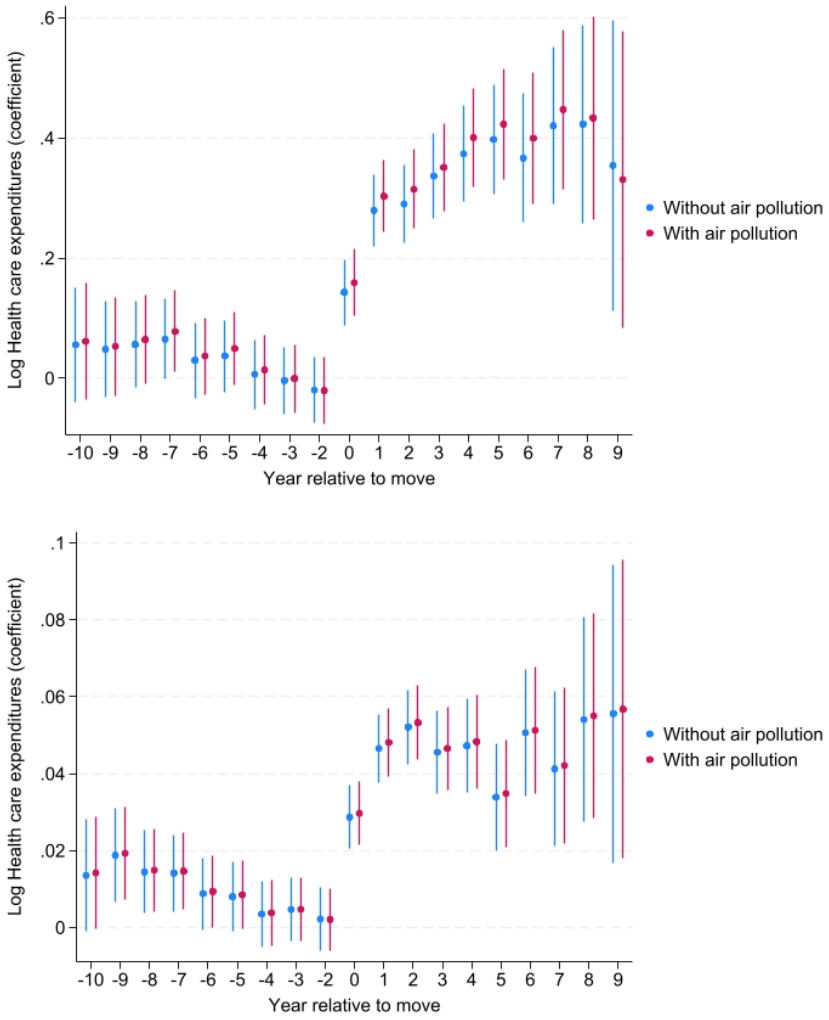
Figure 3: Distribution of Destination-Origin Difference in Log Health Care Costs



Notes. The upper panel in this figure presents the distribution, among all one-time movers, of the difference δ_i^y in average log health care expenditures between their origin and destination municipalities. The lower panel in this figure presents the distribution, among all one-time movers, of the difference δ_i^P in average log air pollution between their origin and destination municipalities. The sample comprises all movers and includes $3,747,611$ individual-years.

to be similar across neighborhoods that are near the same hospital.

Figure 4: Event study results: Health care costs coefficients



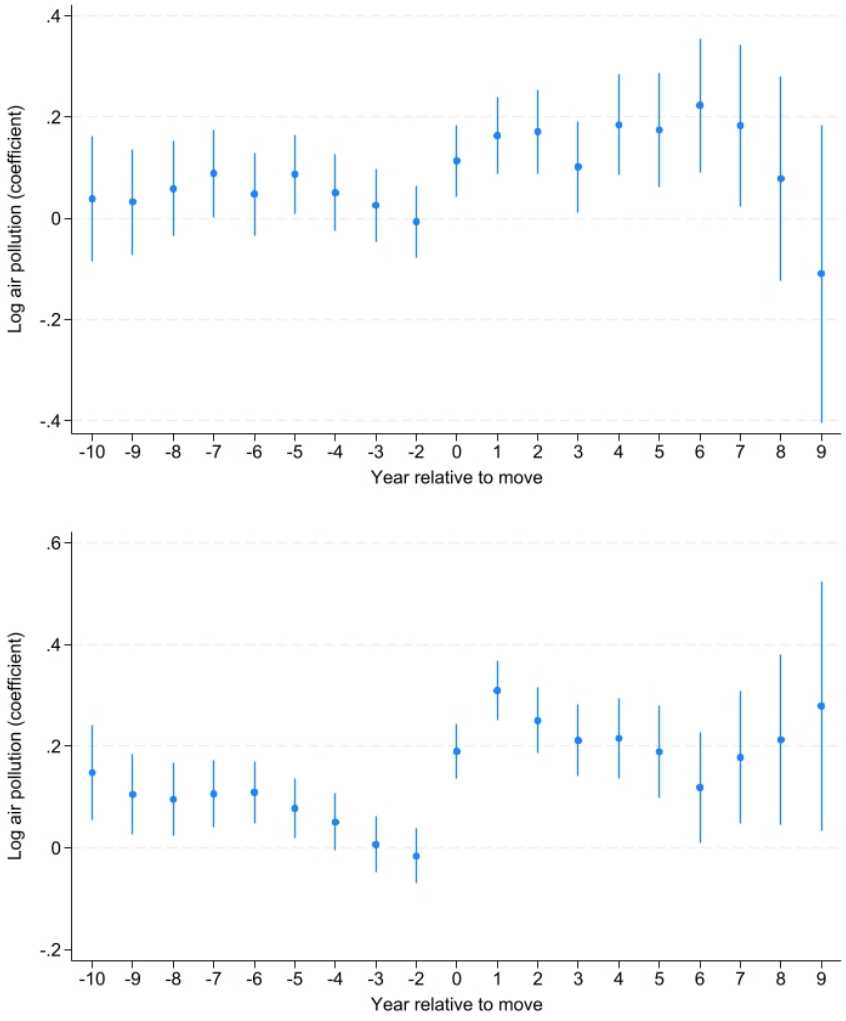
Notes. These figures show the estimated coefficients $\hat{\mu}^{y}_{r(i,t)}$ as represented in equation (8). We normalize the coefficient $\mu^{y}_{r(i,-1)}$ to zero. The blue estimates correspond to the estimated coefficients when $\mu^{P}_{r(i,t)}$ are set to zero, and the red estimates correspond to the estimated coefficients when $\mu^{P}_{r(i,t)}$ are allowed to be non-zero. The top (bottom) figure shows the estimated coefficients when we consider municipalities (neighborhoods) as the geographical division.

4.2 Pollution effects

Figure 5 shows our results of the event study design for exposure to air pollution, i.e., our estimated coefficients of $\mu^{P}_{r(i,t)}$ in equation (8). These results inform us about the value of θ as represented in equation (6). Similarly to the effects of exposure on health care costs

as presented in Figure 4, we again document a discontinuous jump of approximately 0.2, at the time of the move, which is more prominent for our analysis at the neighborhood level as opposed to the municipality level. These results suggest that moving to areas with a 1% higher level of air pollution leads to a 0.2% increase in health care costs, and this effect appears to persist over time.

Figure 5: Event study: Pollution coefficients



Notes. This figure shows the estimated coefficients $\hat{\mu}^{P}_{r(i,t)}$ as represented in equation (8) We normalize the coefficient $\hat{\mu}^{P}_{r(i,-1)}$ to zero. The top (bottom) figure shows the estimated coefficients when we consider municipalities (neighborhoods) as the geographical division.

4.3 Decomposition

Combining our estimates on the supply effect and the pollution effect with the average ratio $\frac{\delta_i^P}{\delta_i^y}$ across movers allows us to complete our decomposition of the geographic variation in healthcare expenditure in the Netherlands. We do this decomposition separately for municipalities and neighborhoods. First, for municipalities we find that supply factors explain 30 to 40% of variation. We estimate θ, the health effect of air pollution, to be around 0.2. The average of $\frac{\delta_i^P}{\delta_i^y}$ across movers at the municipality level is 0.2497, which yields a pollution share of 5%.[8] Summing supply and pollution gives a place effect of 35 to 45%. The remaining 55 to 65% of variation in healthcare spending across Dutch municipalities is then due to sorting by patients with different characteristics. Splitting the variation into supply and demand yields shares of 30-40% and 60-70%, respectively.

At the neighborhood level we find quite a different picture. Supply explains only around 5% of the variation between neighborhoods. Our estimate of θ is roughly the same as in our municipality-level analysis, so 0.2. The average ratio $\frac{\delta_i^P}{\delta_i^y}$ at the neighborhood level is 0.5730, which yields a pollution share around 11.5%. That is more than double our estimate of the pollution share at the municipality level, reflecting the fact that air pollution varies strongly between neighborhoods within municipalities. The total place effect, which consists of supply and pollution, is around 16.5%. While supply factors are the dominant driver of place effects at the municipality level, pollution is the dominant force at the neighborhood level. We attribute the remaining 83.5% to differences in patient characteristics, indicating that sorting effects are much stronger at the neighborhood level than at the more aggregate level.

[8] That is, $0.2 \times 0.2497 \approx 0.05$.

5 Conclusion

This paper studies local air pollution as a potential driver of geographic variation in healthcare expenditure. We show theoretically that pollution complicates the decomposition of variation into local supply factors and patient level demand factors. Pollution is part of the place effect but affects healthcare through the demand side. We thus decompose total expenditure into three components: supply, pollution and patient characteristics. Using Dutch administrative microdata and building on the mover design in Finkelstein et al. (2016), we estimate these effects at the municipality and neighborhood level.

Our findings provide three key insights. First, although incorporating air pollution meaningfully shifts the shares of variation attributed to place and to patients, it does not materially affect the estimated contribution of supply factors. This suggests that omitted variable bias arising from leaving out pollution is limited because of the weak correlation between pollution levels and average healthcare spending across regions. Second, we document a clear and persistent causal effect of pollution on healthcare use: moving to an area with 10% higher PM2.5 increases individual healthcare spending by roughly 2%. This effect is comparable across the municipality and neighborhood levels, and it highlights the importance of environmental conditions in shaping healthcare demand. Third, the relative importance of pollution differs substantially by spatial resolution. At the municipality level, pollution explains around 5% of geographic variation, whereas at the neighborhood level, where spatial variation in pollution is larger, it explains more than 11%. As a result, pollution becomes the dominant component of place-based differences at the neighborhood level, surpassing supply factors. Across both levels of aggregation, however, patient sorting remains the largest contributor to overall variation in healthcare expenditure.

Taken together, our results show that air quality is a meaningful contributor to regional differences in healthcare spending. Policies aimed at reducing pollution may therefore yield broad health and healthcare spending benefits beyond the immediate environmental gains. At the same time, because patient characteristics remain the main driver of varia-

tion, improving prevention and targeting high-risk populations remain essential for reducing disparities in healthcare expenditures. Future work should investigate heterogeneity in the effects of pollution and explore the role of air pollution in explaining regional variation in life expectancy.

References

Anderson, Michael L (2020) "As the wind blows: The effects of long-term exposure to air pollution on mortality," *Journal of the European Economic Association*, 18 (4), 1886–1927.

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur (2026) "The Health Care Cost of Air Pollution: Evidence from the World's Largest Payment Network," *Review of Economics and Statistics*, 108 (1), 110–128.

Basaglia, Piero (2023) "The spatial distribution of pollution reduction benefits: Quasi-experimental evidence from England," *Available at SSRN 4678105*.

Chay, Kenneth Y and Michael Greenstone (2003) "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession," *The Quarterly Journal of Economics*, 118 (3), 1121–1167.

Cinelli, Carlos and Chad Hazlett (2019) "Making Sense of Sensitivity: Extending Omitted Variable Bias," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82 (1), 39–67, 10.1111/rssb.12348.

Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis (2020) "Disparities in PM2. 5 air pollution in the United States," *Science*, 369 (6503), 575–578.

Currie, Janet and Matthew Neidell (2005) "Air pollution and infant health: what can we learn from California's recent experience?" *The Quarterly Journal of Economics*, 120 (3), 1003–1030.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif (2019) "The mortality and medical costs of air pollution: Evidence from changes in wind direction," *American Economic Review*, 109 (12), 4178–4219.

Deryugina, Tatyana and David Molitor (2021) "The causal effects of place on health and longevity," *Journal of Economic Perspectives*, 35 (4), 147–170.

Deryugina, Tatyana and Julian Reif (2023) "The long-run effect of air pollution on survival,"Technical report, National Bureau of Economic Research.

Deschenes, Olivier, Michael Greenstone, and Joseph S Shapiro (2017) "Defensive investments and the demand for air quality: Evidence from the NOx budget program," *American Economic Review*, 107 (10), 2958–2989.

Ding, Hui (2026) "Geographic variation in mental health treatment utilization: evidence from migration," *American Economic Journal: Economic Policy*, 18 (1), 35–68.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams (2016) "Sources of Geographic Variation in Health Care: Evidence From Patient Migration*," *The Quarterly Journal of Economics*, 131 (4), 1681–1726, 10.1093/qje/qjw023.

——— (2021) "Place-based drivers of mortality: Evidence from migration," *American Economic Review*, 111 (8), 2697–2735.

Graff Zivin, Joshua, Matthew Neidell, Nicholas J Sanders, and Gregor Singer (2023) "When externalities collide: Influenza and pollution," *American Economic Journal: Applied Economics*, 15 (2), 320–351.

Lubczyk, Moritz and Maria Waldinger (2025) "The Long-Term Effects of Air Pollution on Health and Labor Market Outcomes: Evidence from Socialist East Germany,"Technical report, CESifo Working Paper.

Moura, Ana, Martin Salm, Rudy Douven, and Minke Remmerswaal (2019) "Causes of regional variation in Dutch healthcare expenditures: Evidence from movers," *Health Economics*, 28 (9), 1088–1098.

Salm, Martin and Ansgar Wübker (2020) "Sources of regional variation in healthcare utilization in Germany," *Journal of Health Economics*, 69, 102271.

Schlenker, Wolfram and W Reed Walker (2016) "Airports, air pollution, and contemporaneous health," *The Review of Economic Studies*, 83 (2), 768–809.