Pollution, Pooling, and Premiums: Spillovers from Toxic Plant Entry on Corporate Healthcare Costs

Gunjan Seth* Manpreet Singh†

March 15, 2026

Abstract

Employer-sponsored health insurance is a major operating cost, yet little is known about how local industrial activity affects premiums. We show that toxic plant entry reduces premiums for nearby incumbent firms by about 10%, even though premiums are higher cross-sectionally in more polluted areas. The decline is larger when entrants add more workers, smaller when entrants are more toxic, and absent for self-insured incumbent firms and single-insurer markets, consistent with geographic risk pooling. Non-toxic plant entry leads to larger premium declines, while carcinogen reclassification reports increase premiums by 14%, confirming that insurers also price health-risk information. Overall, employer health insurance costs depend not only on firm characteristics, but also on the composition of the local insurance pool and insurers' perceptions of environmental health risk.

*Marshall School of Business, University of Southern California, *Email:* gseth@marshall.usc.edu.
†Scheller College of Business, Georgia Tech, *Email:* msingh92@gatech.edu (corresponding author).

1 Introduction

Non-wage benefits have become critical to corporate strategy and financial management, and employer-sponsored health insurance is the largest and most routinely repriced among them, representing a significant component of firms' operating costs that can meaningfully constrain firm decisions (Ouimet and Tate, 2025).[1] In the United States, employer-sponsored benefits cover nearly 60% of working-age Americans, and employers spend over $800 billion each year to provide coverage for more than 180 million people. Yet despite its scale, we know little about why these costs differ so sharply across firms, and in particular how local industrial activity shapes what firms pay for worker coverage.

Employer health insurance premiums are typically determined in *local* markets and repriced annually to reflect realized medical spending and local market conditions (Aghamolla, Jain, and Thakor, 2025), implying that local industrial activity can transmit quickly into a significant component of firm costs. Communities often face tradeoffs between economic development and health risks. We study whether toxic plant entry changes the health insurance costs of nearby employers. While new toxic plants bring jobs and tax revenue, the standard prediction is that nearby employers should face higher health insurance premiums as environmental exposure increases worker illness and medical utilization.[2]

Although firms near toxic plants with higher emissions pay higher premiums cross-sectionally, we find strikingly opposite results around toxic plant openings. Exploiting

[1] Recent work links health care costs to corporate investment and highlights that these costs impact firm decisions, including capital investment and disproportionately harm low-income workers through higher job separation and unemployment (Tong, 2024; Gao, Ge, Schmidt, and Trillo, 2025). Ouimet and Tate (2025) show that costs vary substantially across firms within the same metropolitan area and industry due to competitive pressure to offer high-quality benefit plans.

[2] More broadly, firms are exposed to geographically local shocks, and firm location is systematically related to corporate outcomes and policies (Almazan, De Motta, and Titman, 2007; Greenstone, Hornbeck, and Moretti, 2010; Giroud and Mueller, 2019; Dougal, Parsons, and Titman, 2022). A substantial literature documents that toxic industrial facilities impose significant health externalities on nearby populations, increasing mortality rates (Currie, Davis, Greenstone, and Walker, 2015), elevating illness (Currie, Zivin, Mullins, and Neidell, 2014), and raising healthcare expenditures (Graff Zivin and Neidell, 2013).

the staggered timing of toxic plant entry in a stacked difference-in-differences design, we find that premiums for incumbent firms within 2 miles of the toxic plant fall by 10%. This represents approximately $411,000 in annual savings per firm or roughly $425 per participant per year. The decline is larger when entrants bring more workers and smaller when entrants are more toxic, consistent with diversification gains being offset by adverse pollution signals.

This pattern is puzzling precisely because a large literature documents that toxic exposure worsens worker health, raises medical utilization, and increases healthcare costs over time (Currie and Neidell, 2005; Currie et al., 2014; Graff Zivin and Neidell, 2013). The resolution requires recognizing two competing forces operating simultaneously through health insurance markets. The first force (*health risk*) operates as environmental economics predicts: toxic exposure worsens health outcomes and increases medical costs, and insurers incorporate observable emissions from EPA Toxics Release Inventory data as a signal of future health deterioration when setting premiums. The second force (*insurance market pooling*) operates through costly underwriting and information frictions in local insurance markets, which lead insurers to price and pool risk geographically rather than fully relying on firm-specific experience. When insurers cannot perfectly observe individual risk characteristics due to information asymmetries (Akerlof, 1970; Rothschild and Stiglitz, 1976) or find firm-specific underwriting prohibitively costly (Pauly, 1970), they pool risks geographically and set premiums based on group averages rather than individual firm experience (Bundorf et al., 2010).

When new plants enter, they expand the local risk pool, and larger pools improve risk diversification, lowering premiums for all firms priced in the local pool before cumulative health effects from toxic exposure manifest. Consistent with this mechanism, self-insured firms, who bear their own claims risk directly and bypass pooled pricing (Linehan, 2010; U.S. House of Representatives, 2025), show no premium response to nearby plant entry. Premiums also do not fall in markets served by a single insurer, confirming that the effect operates through local insurance market structure rather than through labor market or

economic activity channels.

To isolate the pooling mechanism from health risk effects, we exploit entry by non-toxic plants from the National Establishment Time-Series database. We identify plant entries with at least 20 employees that release zero toxic emissions, i.e., *non-toxic entrants* (non-TRI facilities). Importantly, we exclude any non-toxic entrant if a toxic plant enters within 5 miles at any point during the full estimation window (three years before to three years after the entry year). These non-toxic entrants expand the local pool without introducing observable pollution signals, allowing us to test whether risk diversification alone can lower premiums when firm-specific risk is costly to underwrite. We find that non-toxic plant entry reduces premiums by roughly twice the baseline effect of toxic plant entry on premiums. This difference indicates that observable emissions act as adverse signals that insurers price into premiums under information frictions, attenuating the diversification benefit from pool expansion. The larger effects for non-toxic entrants therefore confirm that our baseline results combine diversification gains with signal-based risk pricing, with the net effect understating the pooling mechanism.

If the pooling channel dominates in the short run, insurers must still be pricing health risk — the diversification effect simply overwhelms it at entry. To confirm this directly, we exploit exogenous changes in perceived carcinogenicity in the National Toxicology Program Reports, following the information shock approach of Gormley and Matsa (2011) and van Binsbergen et al. (2025). These reclassification events shift insurers' beliefs about future health costs without changing local pool size, thereby isolating the risk-pricing channel from any diversification effect. We find that when chemicals are reclassified as carcinogens, premiums increase by 14% for firms within 2 miles of the existing toxic plant, confirming that insurers price risk news upward and validating that the two forces operate in the directions the framework predicts. These patterns together imply that a major component of firms' operating costs is shaped not only by their own characteristics, but by local insurance market structure, pool composition, and the information insurers observe about nearby industrial activity.

Our findings show that employer health insurance costs depend not only on firm characteristics, but also on the composition of the local insurance pool and the information insurers observe about nearby industrial activity. These results speak to three strands of literature. First, we contribute to research on nonwage benefits and corporate outcomes (Ouimet and Tate, 2025; Tong, 2024; Gao et al., 2025; Aghamolla et al., 2025) by showing that local changes in the insured workforce can generate spillovers in employer health insurance premiums across nearby firms through geographic risk pooling. Because premiums are routinely repriced, local shocks to the risk pool can quickly shift a major component of operating costs, with downstream consequences for firm investment, hiring, and location decisions. Second, we contribute to research on local economic shocks and industrial entry (Giroud and Mueller, 2017, 2019; Greenstone et al., 2010) and the health effects of pollution exposure (Currie and Neidell, 2005; Currie et al., 2014; Graff Zivin and Neidell, 2013) by documenting an insurance-market transmission channel through which toxic plant entry affects incumbent firms' operating costs. This channel is distinct from the labor market and supply chain mechanisms studied in prior work and operates in the opposite direction to what direct health cost effects would predict. Third, we contribute to the healthcare economics literature on geographic cost variation and insurer market power (Fisher et al., 2003; Dafny et al., 2012; Cooper et al., 2019; Brot-Goldberg et al., 2024) by showing that local industrial activity affects employer health insurance costs through insurer pooling decisions, and that insurer concentration determines whether diversification gains are passed through to firms.

The rest of the paper is organized as follows. Section 2 presents institutional details and the conceptual framework that motivates the empirical tests. Section 3 describes the data, Section 4 outlines the empirical strategy, Section 5 presents the baseline results and mechanism tests, and Section 6 concludes.

2 Institutional Details and Conceptual Framework

This section has two parts: Section 2.1 describes the institutional setting for employer-sponsored health insurance, and Section 2.2 presents the conceptual framework that links pooling, pollution signals, and premiums.

2.1 Employer-Sponsored Health Insurance

Employer-sponsored health insurance (ESHI) is a central pillar of the U.S. health system and a major component of non-wage compensation. Employer coverage accounts for roughly 60% of working-age Americans, and employer spending on health benefits is enormous: in 2018, U.S. employers spent about $727 billion on health insurance of workers (Tong, 2024). As of 2024, employer-sponsored health insurance covers about 154 million nonelderly people.[3] These magnitudes underscore that health insurance is not a peripheral perk; it is a firm-wide operating cost that is large, routinely repriced, and embedded in total employee compensation.

Recent national benchmarks underscore the per-worker scale of these costs. In 2025, average annual premiums are about $9,325 for single coverage and $26,993 for family coverage, with workers contributing roughly $6,850 toward the family premiums on average.[4] These levels are large relative to typical wage growth and illustrate how premium changes can shift labor costs even in the short run.

Institutionally, most firms provide coverage through a menu of employer-sponsored plans chosen annually, with premiums paid by employers and employees and with cost sharing through deductibles and copays. Benefits are typically standardized within a firm, so health insurance operates as a firm-wide compensation policy rather than an individualized contract. This standardization matters for our setting: when pricing is

[3] KFF, 2024 Employer Health Benefits Survey, Summary of Findings: https://www.kff.org/report-section/ehbs-2024-2024-employer-health-benefits-survey/.

[4] KFF, 2025 Employer Health Benefits Survey: https://www.kff.org/health-costs/2025-employer-health-benefits-survey/.

pooled at the local market level and benefits are offered broadly within firms, shocks that change the local risk pool or perceived health risk can transmit quickly into employer premiums and, therefore, operating costs. Recent work on non-wage compensation emphasizes this firm-wide nature of benefits and how benefit generosity shapes labor-market sorting and workforce composition.[5]

Finally, ESHI costs matter for corporate outcomes. Tong (2024) shows that increases in employer health care costs reduce capital expenditures and R&D, especially for financially constrained firms, while Gao et al. (2025) document that higher premiums disproportionately harm low-income workers through job separation and earnings losses. In related work, Aghamolla, Jain, and Thakor (2025) highlight that health insurance costs and local market conditions affect firms' benefit expenditures. In their setting, shocks to local health care costs spill over into real outcomes, including higher business bankruptcies, slower employment and establishment growth, and lower wages. These studies motivate our focus on how local environmental shocks shape the pricing of a large and routinely repriced component of firms' operating costs. We next formalize the pricing and pooling mechanisms that link local shocks to premiums.

2.2 Conceptual Framework

In this section we outline a simple conceptual framework that highlights two opposing forces through which plant entry may affect health insurance premiums for nearby incumbent firms. The first force operates through risk diversification within the insurance pool. The second operates through observable pollution signals that may affect insurers' expectations about future health costs.

Small and medium sized firms typically obtain health insurance in pooled local markets rather than through firm specific underwriting. Conducting individualized underwriting requires insurers to collect and process detailed information on workforce health risk, evaluate claims histories, and design customized pricing contracts. These activi-

[5]See Ouimet and Tate (2025) for evidence on standardization and non-wage benefits.

ties involve substantial fixed costs. For small employers the potential gains from such individualized pricing are limited because the insured workforce is small. As a result, the expected benefits of detailed underwriting often do not justify these costs. Insurers therefore commonly pool small firms within local geographic markets and set premiums based on the aggregate risk of the pool rather than the characteristics of each individual firm (Pauly, 1970). In this institutional setting, changes in the composition of the local pool can generate spillovers across participating firms.

Plant entry can affect premiums through a diversification channel. Consider a local insurance pool with L covered workers. Let σ^2 denote the variance of individual worker health expenditures. If worker health risks are independent across individuals, the variance of average health costs in the pool equals σ^2/L. Suppose a new plant employing L_{new} workers enters the market and its employees join the same insurance pool. The pool size increases to $L + L_{\text{new}}$, reducing the variance of average costs to $\sigma^2/(L + L_{\text{new}})$. If insurers are risk averse and price contracts based on both expected costs and uncertainty about realized claims, a reduction in variance lowers the risk premium embedded in insurance prices. This diversification benefit lowers premiums for all firms participating in the pool.

Plant entry may also affect premiums through an informational channel related to environmental exposure. Under information asymmetry (Akerlof, 1970; Rothschild and Stiglitz, 1976), insurers cannot directly observe the health status of individual workers or the average health risk of each firm's workforce. However, insurers may observe publicly available environmental information that is correlated with future health outcomes. In particular, emissions reported in the EPA's Toxics Release Inventory provide signals about local pollution exposure. These emissions may affect insurers' expectations about future medical costs even before the claims are realized. When a new plant enters the local market, insurers may update their beliefs about expected health expenditures using information about the entrant's emissions intensity. Entrants with high emissions raise the pool's average pollution signal and may increase the expected cost component of

premiums.

The overall premium response therefore depends on the relative strength of these two forces. When a new plant adds a large number of employees to the local insurance pool, the diversification effect may dominate, causing premiums to decline as the pool expands. For entrants with high emissions, the adverse pollution signal may partially offset or even outweigh the diversification benefit. For non-toxic entrants that report no emissions, both forces operate in the same direction because pool expansion reduces risk while the absence of emissions lowers the pollution signal. In Appendix A we formalize these mechanisms in a simple pooling model and derive a set of testable predictions for insurance pricing.

Prediction 1 (Diversification Effect). Toxic plant entry reduces premiums for nearby fully insured incumbents. The reduction is larger when entrant employment is higher and smaller when observable emissions are higher.

Prediction 2 (Insurance Market Mechanism). Premium effects of plant entry are present for fully insured firms but absent for self insured firms that opt out of pooled pricing.

Prediction 3 (Non-Toxic Entry). Non-toxic entrants generate larger premium declines than toxic entrants because pool expansion is not offset by an adverse emissions signal.

Prediction 4 (Pollution Signal Effect). Information shocks that increase perceived pollution risk but do not change the size of the local workforce or the insurance pool, such as carcinogen reclassifications, increase premiums for nearby incumbent firms.

3 Data

Our empirical analysis combines employer-sponsored health insurance records from Form 5500 filings with toxic plant entry and emissions data from the EPA Toxics Release Inventory (TRI). These datasets allow us to link firms' health insurance costs to nearby industrial plant openings. We link these sources geographically using geocoded establishment-level locations and match firms across datasets using Employer Identification Numbers

(EINs) and Dun & Bradstreet (D&B) numbers.

3.1 Health Insurance Premiums: Form 5500

Our primary outcome variable is the annual health insurance premium paid by each employer, drawn from the U.S. Department of Labor's Form 5500 filings. All private-sector employers sponsoring benefit plans with 100 or more participants are required to file Form 5500, and smaller firms may file the abbreviated Form 5500-SF. We use data spanning 1999–2022. For the standard Form 5500, we merge the main form with Schedule A, which reports insurance carrier payments, using the plan acknowledgment identifier (`ack_id`) for filings after 2008 and the EIN–plan number combination for earlier years. We retain only health insurance plans. For filings post 2008, plans are identified using the welfare benefit health indicator (`wlfr_bnft_health_ind`). For earlier filings that lack this indicator, we identify plans using benefit codes and plan name keywords (e.g., "health," "medical," "PPO," "HMO"). For Form 5500-SF filers, which do not file Schedule A separately, we use the total income amount reported for plans coded as health benefits (benefit code 4A). We collapse all health plans to the firm–year level, summing total premiums paid and total number of participants. Our outcome variable is the log of total health insurance premium paid per participant.

For our analysis, we match insurance records to firms' geographic information by linking EINs to D&B numbers using the DUNS–EIN cross reference file. For EINs without a direct cross reference match, we perform fuzzy name and zip code matching against the Mergent Intellect (formerly D&B) establishment database. Geocoded establishment level latitude and longitude coordinates are sourced from Mergent Intellect's company details file, which provides precise location information for each D&B establishment. To ensure clean identification of firm location, we restrict the sample to single location firms, defined as those for which the EIN maps to exactly one D&B establishment number, yielding approximately 58,000 firms with validated geographic information.

3.2 Toxic Plant Entry: EPA Toxics Release Inventory

Data on toxic industrial plant openings come from the EPA's Toxics Release Inventory. The TRI requires facilities that manufacture, process, or use listed toxic chemicals above threshold quantities to report annual releases. We use TRI facility records spanning 1987–2022 and define a plant's entry year as the first year it appears in the TRI panel. We supplement TRI facility records with D&B numbers when available (approximately 47% of TRI facilities report a D&B number directly; an additional 31% can be linked via the parent company D&B field) and use Mergent Intellect to obtain facility-level latitude and longitude coordinates, sales, and employment data.

To construct our treatment and control groups, we compute pairwise geodesic distances between every toxic plant opening and all single-location incumbent firms using the `geodist` command in Stata. We define treated incumbents as firms located within a given radius (1 to 3 miles) of a plant opening. We use a 25-mile radius to define the broader pool from which control firms are drawn, building a balanced panel spanning three years before and three years after each plant opening. Our main specification uses a 2-mile treatment radius, consistent with the evidence in prior work that health and pollution externalities attenuate sharply beyond 2 miles (Currie et al., 2014; Gormley and Matsa, 2011).

In addition to identifying plant entry, TRI emissions data allow us to construct a measure of entrant toxicity. Entrant toxicity is measured using the EPA's Risk-Screening Environmental Indicators (RSEI) Cancer Hazard score, a toxicity-weighted metric constructed from TRI-reported chemical releases that reflects the potential carcinogenic hazard associated with a facility's emissions.

3.3 Summary Statistics

Our baseline estimation sample links Form 5500 health insurance records to the toxic plant entry dataset. The final sample for the main difference-in-differences analysis

merges 7,995 toxic plant openings to 35,354 incumbent firm–year observations within 5-miles of the toxic plant.

Table 1 reports summary statistics for incumbent firms and toxic plant entrants. Panel A shows that healthcare premiums vary substantially across incumbent firms in the estimation sample. The average annual healthcare premium per participant in our sample is about $4,200, and the average plan has around 1090 participants. Roughly 19% of firms in our sample are self insured. The number of insurance companies serving firms rises with geographic scope, from about 3.3 within 1 mile to roughly 13 in the 3 mile and 3 to 5 mile rings. Average sales among incumbent firms range from about $60 to $80 million across distance rings.

The sample appears broadly comparable across rings. The firms located closer to toxic plants do not appear systematically different from firms in the 3 to 5 mile control ring in terms of insurance market characteristics. This comparability supports the empirical design by suggesting that treated and control firms are drawn from broadly similar local environments. At the same time, there is substantial dispersion in healthcare premiums, plan participants, and insurer counts, which provides meaningful variation for the analysis.

Panel B reports summary statistics for toxic plant entrants in the entry year. Entrants display substantial heterogeneity in both emissions and economic size. Mean TRI pounds equal about 35,182, while the median is 1,130, indicating a highly skewed emissions distribution. Similar variation can be observed in the RSEI cancer hazard measure. Toxic plant entrants have mean employment of about 99 workers and mean sales of about $44 million. On average, entrant plants have about 0.25 incumbent firms within 1 mile, 0.73 within 2 miles, 1.61 within 3 miles, and 1.82 between 3 and 5 miles. These statistics indicate that plant entry generates a localized shock affecting a limited but economically meaningful set of nearby incumbent firms, which motivates the use of narrow treatment rings and distance based comparisons in the main specifications.

4 Methodology

Our baseline design studies how healthcare premiums paid by incumbent firms change around nearby toxic plant openings. The framework in Section 2 implies two opposing forces. Plant entry can lower premiums by expanding the local insurance pool and improving risk diversification. At the same time, plant entry can raise expected future healthcare costs if the entrant reveals new information about local pollution risk. The baseline specification estimates the net effect of these two forces.

We follow the geographic ring approach used in recent work that identifies local effects using narrow distance bands around plausibly localized shocks. In our setting, we compute pairwise geographic distances between each toxic plant opening and all single-location incumbent firms. Treated firms are those located within a distance threshold X miles of the entering plant, where $X \in \{3, 2, 1.5, 1.25, 1\}$. Control firms are incumbent firms located between 3 and 5 miles of the same plant. This ring design is motivated by the evidence that environmental and other local effects attenuate sharply with distance, and by recent papers that use close geographic comparisons to isolate the local effects of plausibly exogenous shocks (Butts (2023), Currie et al. (2015), Diamond and McQuade (2019), Ganduri and Maturana (2024), LaPoint and Yale (2022), van Binsbergen et al. (2025)).

For each toxic plant opening, we construct a balanced panel of incumbent firms observed from three years before to three years after the opening year. When an incumbent firm is exposed to more than one toxic plant opening, we assign the firm to a single event following the assignment procedure in van Binsbergen et al. (2025): we pair the incumbent firm with the toxic plant in the earliest opening year for which it enters the treatment or control sample, and if multiple openings occur in that same year, we pair the firm with the nearest toxic plant opening. We then keep this pairing fixed throughout the event window. This procedure ensures that each incumbent firm contributes to only one local event and avoids overlap across nearby openings.

Our baseline estimating equation is

$$\log(Premium)_{ipt} = \beta \times Post_{pt} \times \mathbb{1}(Distance_{ip} < X) + \gamma_i + \delta_{pt} + \theta X_{it} + \varepsilon_{ipt} \quad (1)$$

where i indexes incumbent firms, p indexes toxic plant openings, and t indexes calendar year. The dependent variable, $\log(Premium)_{ipt}$, is the log of annual total health insurance premiums paid per participant by incumbent firm i in year t. $Post_{pt}$ equals one for years after plant p opens and zero otherwise. $\mathbb{1}(Distance_{ip} < X)$ equals one if incumbent firm i lies within X miles of the entering plant and zero if the firm lies in the 3 to 5 mile control ring.

The specification includes incumbent firm fixed effects, γ_i,which control for time-invariant firm characteristics, and plant $\times$ year fixed effects, δ_{pt}, which control for time-varying conditions common to firms linked to toxic plant opening p in year t. We also control for the number of plan participants, $log(Participants)_{it}$, as premiums are measured on a per participant basis and plan size may affect pricing. Standard errors are clustered at the toxic plant event level.[6]

The coefficient of interest, β, measures the within-firm relative change in healthcare premiums for firms within X miles of a toxic plant relative to firms in the 3 to 5 mile ring around the same plant, after the plant opens. Under the identifying assumption that, absent the plant opening, firms in the treatment and control rings would have followed similar trends, β can be interpreted as the causal effect of toxic plant entry on incumbent firms' healthcare premiums.

5 Results

5.1 Baseline Evidence: The Premium Puzzle

We begin with the central empirical puzzle in the paper by plotting coefficients from the cross-sectional regression of log healthcare premiums on standardized measures of local

[6] All results are robust to clustering standard errors at the incumbent firm level.

toxic exposure.[7] We measure local toxic exposure using three variables: (i) TRI Pounds, which captures toxic emissions, (ii) RSEI Cancer Hazard score, which captures toxicity-weighted cancer risk, and (iii) Number of Toxic Plants. Figure 1 shows that, in the cross-section, firms located in more polluted areas pay higher healthcare premiums. Across all three exposure measures, healthcare premiums are positively related to toxicity within 0 to 2 miles of the firm. Economically, a one standard deviation increase in toxic exposure within 0 to 2 miles is associated with a statistically significant 3% to 4% increase in healthcare premiums. By contrast, the corresponding coefficients for exposure measured within 2 to 4 miles are statistically insignificant and close to zero. These cross-sectional patterns are consistent with higher healthcare costs in more polluted local environments, but they do not isolate the causal effect of toxic plant exposure on healthcare premiums.

These cross-sectional patterns are descriptive. To estimate the causal effect of toxic plant openings on healthcare premiums, we exploit the timing of plant entry in a difference-in-differences design. Table 2 reports the baseline results. The estimates show that healthcare premiums fall for incumbent firms located near a newly entering toxic plant. The effect is negative across all treatment radii and becomes larger in magnitude as the treatment ring narrows. For firms within 3 miles of a newly entering toxic plant, healthcare premiums fall by 7.9%. Within 2 miles, healthcare premiums fall by 10.4%, and within 1.5 miles they fall by 14.4%. Translating the 2-mile estimate into dollar terms using the sample average healthcare premium of approximately $4,094 per participant and 966 plan participants (Table 1), the 10.4% decline corresponds to roughly $425 per participant per year, or approximately $411,000 in annual premium savings for the incumbent firm. The coefficients are economically meaningful and statistically significant. Further, the response is strongest for firms located closest to the entrant. This evidence contrasts with the positive cross-sectional pattern presented in Figure 1. Taken together, these findings suggest that the short-run effect of toxic plant entry differs from the broader

[7] We control for zipcode and year fixed effects and county × industry × size fixed effects. Industry is defined at the 4 digit NAICS level and size is defined using five size bins based on the number of employees at the firm.

cross-sectional relationship between pollution and healthcare premiums.

To rule out the possibility that the decline in average healthcare premiums is driven by changes in plan participation, we also examine the number of plan participants. Table 3 uses the log number of plan participants as the dependent variable and provides a direct test of whether nearby plant entry changes the size of the incumbent firm's insured workforce. The estimates are small and statistically insignificant across specifications, suggesting that toxic plant entry does not systematically change plan enrollment. This finding indicates that the premium declines in Table 2 are not driven by changes in participation and instead reflect changes in pricing.

We also test whether the premium decline reflects changes in plan design rather than pricing. Table A.I examines whether incumbents switch plan types after nearby toxic plant entry by using indicators for PPO and HMO plan offerings as outcome variables.[8] We find no evidence of systematic switching in plan type following entry, which supports the interpretation that the decline in premiums reflects pricing rather than a shift toward lower quality plans. We also reestimate the main specification after excluding incumbent firms that experienced toxic plant exits during the event window, and the results remain robust.[9]

Figure 2 reports event study estimates around toxic plant entry for firms within 0 to 3 miles of the entrant, using firms in the 3 to 5 mile ring as the control group. The pre-entry coefficients are statistically insignificant and close to zero, indicating no evidence of differential pre trends in premiums between the treatment and control rings. The effect is modest in the entry year and then becomes more negative over the following years, consistent with premiums adjusting gradually after local plant entry.

We next estimate how the premium response varies with distance from the entering

[8]PPO and HMO indicators are constructed using Form 5500 Schedule A, Part III (Welfare Benefit Contract Information), line 8 (Benefit and contract type), which reports the contract type for each plan. We define the indicator at the firm level and it equals one if any of the firm's plans is classified as PPO or HMO, and zero otherwise. PPO and HMO classifications are available only for filings after 2009.

[9]Appendix Table A.II reports results excluding incumbents with exits within 3 miles and within 5 miles. While some treatment radii lose statistical significance because the exclusion reduces the sample size, the estimates remain negative across specifications.

plant. Figure 3 reports estimates from a specification that replaces the single treatment indicator with one mile distance bins around the entrant. The premium decline is concentrated in the bins closest to the entrant and becomes smaller as distance increases, with effects close to zero in the 3 to 4 mile ring. This pattern is consistent with a local response to plant entry rather than a broad area wide shock, and it indicates that both diversification and pollution signaling operate through nearby exposure and local insurance market conditions.

Taken together, these results highlight the premium puzzle: although more polluted areas have higher healthcare premiums in the cross section, toxic plant entry lowers healthcare premiums for nearby incumbent firms in the short run.

5.2 Evidence on Mechanisms

This section evaluates the mechanisms outlined in Section 2.2. The structure of the analysis follows the four predictions of the framework, examining diversification effects at plant entry, the role of insurance market structure, non-toxic entry, and information shocks about pollution risk.

5.2.1 Diversification and Pollution Signals at Entry

We begin by testing Prediction 1, which states that toxic plant entry should lower premiums through diversification, with larger effects when entrants bring more workers and weaker effects when entrants are more toxic.

Table 4 tests the diversification channel using heterogeneity in entrant employment. The estimates show that premium declines are concentrated among larger entrants. For entrants with above-median employment, the reduction in premiums is economically large across treatment radii. For entrants with below-median employment, the estimates are small and statistically insignificant. This pattern is consistent with the diversification channel. A larger entrant expands the local insurance pool by more, which improves risk

sharing and lowers the premium charged to nearby incumbent firms.[10]

Table 5 tests the pollution signaling channel. The estimates differ sharply by entrant toxicity. For low toxicity entrants, the premium effect is negative and large across all treatment rings. For high toxicity entrants, by contrast, the estimates are close to zero and statistically insignificant. This pattern fits our framework. Pool expansion tends to push premiums down, but a more toxic entrant also sends a stronger signal of future health risk. The two forces work in opposite directions, and for sufficiently toxic entrants the signal effect largely offsets the diversification effect.

Taken together, our results provide direct support for the model's first prediction. The premium decline is stronger when plant entry expands the pool more and weaker when the entrant brings a stronger adverse health signal.

5.2.2 Insurance Market Structure and Pooled Pricing

Prediction 2 states that the premium response should be concentrated in settings where pooled insurance pricing matters. Table 6 compares self insured and non self insured incumbent firms. The premium decline is concentrated among non self insured firms. For self insured firms, the estimates are positive but statistically insignificant. This contrast matters because self insured firms bear claims risk internally and are less exposed to pricing in the fully insured market. If the estimates instead reflected broad local changes such as labor market shifts, local demand, or provider supply, one would expect a similar response for both groups. The fact that the effect shows up only among non self insured firms points to the insurance pricing channel.

Table 7 provides a second validation test using local insurance market concentration. The premium decline is concentrated in treatment rings served by multiple insurers and

[10]The same logic implies a stronger response among smaller incumbent firms. Smaller firms are less likely to be underwritten based on their own experience and are more likely to be priced as part of a local pool, so they should benefit more from an increase in local pool size. To keep firms comparable around the size cutoff, we restrict the sample to incumbent firms with fewer than 150 employees and split them into two bins based on the subsample median. Appendix Table A.III reports the results. The estimates are more negative for smaller firms than for larger firms, consistent with insurer pricing becoming less dependent on local pooling as incumbent firm size increases.

is absent in rings served by a single insurance company. This pattern is again consistent with the pooling interpretation. In markets with multiple carriers, local pool expansion is more likely to affect pricing through competition and group risk assessment. In markets dominated by a single insurer, pricing is typically less responsive to marginal changes in pool composition. Similar links between insurer concentration and employer sponsored premiums have been documented by Lin and Zhou (2026) and Dafny (2010).

Insurance companies often price and pool risk within local administrative areas, such as counties. Table 8 tests this implication by comparing firms located in the same county as the entering toxic plant with firms located in a different county but still within the treatment radius. The premium decline is concentrated among firms in the same county, consistent with the geographic pooling channel.

Our results show that the premium decline is concentrated among non self insured firms, in the treatment rings served by multiple insurers, and within the same county as the entrant. In each case, the effect is strongest in settings where pooled insurance pricing is most likely to operate and where incumbent firms are most likely to be priced through that channel.

5.2.3 Non-Toxic Plant Entry and Pool Expansion

The baseline toxic plant entry design captures the net effect of pool expansion and pollution signaling. To isolate the diversification channel more directly, we study entry by non-toxic plants that expand the local employment pool but do not generate toxic emissions.

We identify non-toxic plant openings using the National Establishment Time-Series (NETS) database, which provides establishment-level information on founding year, employment, and location for U.S. firms. We define non-toxic entrants as establishments that (i) are founded after 1999, (ii) employ at least 20 workers in the founding year, (iii) never appear in the EPA TRI database under their own or affiliated D&B number, and (iv) are not located within five miles of a toxic plant opening during the estima-

tion window. This construction ensures that non-toxic entry expands the local workforce without introducing observable pollution signals. We then compute pairwise distances to incumbent firms and apply the same difference in differences design as in the toxic plant analysis, using the same treatment and control rings and the same ±3 year event window.

Table 9 reports difference in differences estimates for non-toxic plant entry. The estimates are strongly negative, with magnitudes larger than in the baseline toxic entry specification. Within 3 miles of a non-toxic entrant, incumbent firms' healthcare premiums fall by about 32%. This pattern matches the model's prediction: when the local pool expands without an offsetting pollution signal, the diversification effect should be stronger. The non-toxic entry results provide a cleaner test of the pool expansion channel and show that it can generate sizable premium declines.

We then sharpen the comparison using a matched triple difference design that pairs each toxic plant opening with a nearby non-toxic plant opening. Each toxic plant is matched to a non-toxic plant that opens in the same year and lies in the same employee size bucket. Employee size buckets are defined using the sample median, with small plants having at most 75 employees and large plants having more than 75 employees. Among eligible matches, we pair each toxic plant with the nearest non-toxic plant located 10 to 20 miles away. This restriction limits spatial overlap in the incumbent firm samples while keeping local conditions broadly comparable.

The matched triple difference specification is

$$\begin{aligned} \log(Premium)_{ij(p)t} = {} & \beta_1 \times \mathbb{1}(Distance_{ij(p)} < X) \times \mathbb{1}(Toxic)_{j(p)} \\ & + \beta_2 \times Post_{it} \times \mathbb{1}(Distance_{ij(p)} < X) \times \mathbb{1}(Toxic)_{j(p)} \\ & + \gamma_{i(p)} + \gamma_{Ring(p)t} + \gamma_{j(p)t} + \theta X_{it} + \varepsilon_{ij(p)t} \end{aligned} \tag{2}$$

where i indexes incumbent firms, $j(p)$ indexes entrant plants within matched pair p, and t indexes year. $\mathbb{1}(Toxic)_{j(p)}$ equals one for toxic entrants and zero for matched non-toxic entrants. $\gamma_{i(p)}$ are incumbent firm × matched pair fixed effects, $\gamma_{Ring(p)t}$ are

distance ring × year × matched pair fixed effects, and $\gamma_{j(p)t}$ are entrant plant × year × matched pair fixed effects. Distance rings are defined as 0-1, 1-2, 2-3, 3-4, and 4-5 miles from the entrant plant.

The coefficient of interest is β_2. It measures whether the post-entry premium response differs for toxic entrants relative to observationally similar non-toxic entrants. This specification isolates the pollution signaling channel by comparing toxic and non-toxic entry, thereby differencing out the pool expansion effect.

Table 10 compares toxic and non-toxic entrants in the matched triple difference design. The coefficient on the triple interaction is positive, which implies that the premium decline is smaller for toxic entrants than for matched non-toxic entrants. Put differently, after holding entry timing, local conditions, and entrant size fixed through the matching design, toxic entry leads to a smaller reduction in premiums than non-toxic entry. This is precisely the pattern implied by our model: both types of entry expand the local pool, but only toxic entrants bring an adverse pollution signal that offsets part of the diversification gain.

Overall, the matched comparison highlights the tradeoff between two competing channels. Pool expansion puts downward pressure on premiums, while pollution signaling works in the opposite direction. The observed response around toxic plant entry reflects the net effect of these channels.

While non-toxic plant entry isolates the diversification channel by expanding the local workforce without introducing pollution signals, our next test isolates the pollution signaling channel by studying information shocks that change perceived health risk without affecting pool size.

5.2.4 Carcinogen Reclassification and Pollution Signals

We next study information shocks that change perceived pollution risk without affecting local employment or the size of the insurance pool. For this analysis, we use carcinogen reclassification events from the National Toxicology Program's (NTP) *Report on Car-*

cinogens (RoC), which periodically updates the scientific classification of chemicals as "known" or "reasonably anticipated" human carcinogens. Following Gormley and Matsa (2011) and van Binsbergen et al. (2025), we treat the year of each RoC release as an information shock for incumbent firms located near TRI facilities that report using the reclassified chemical.

We obtain the list of chemicals newly listed or upgraded to carcinogenic across RoC editions and match them to TRI facilities using Chemical Abstracts Service (CAS) registry numbers. Each reclassification event is then linked to facilities that reported using the chemical prior to the reclassification. We exclude facilities that begin using the chemical only after the report date to avoid post-treatment selection. This yields approximately 5,200 plant–reclassification events between 1999 and 2022.

The estimation sample follows the same single-location incumbent definition and pairwise distance construction as in the toxic plant entry analysis. We study incumbent firms in a balanced window from one year before to one year after the reclassification event. Treated firms are those located within X miles of a facility using the reclassified chemical, and control firms are those located between 3 and 5 miles of the same facility. Because reclassification shifts perceived risk but does not change the local pool, this design isolates the pollution signaling channel independently of pool expansion.

Table 11 studies cases in which a chemical used by a nearby toxic plant is newly listed or upgraded as carcinogenic. Unlike plant entry, these events do not change local pool size. Instead, they shift information about health risk while keeping the underlying local workforce fixed. The estimates are positive across treatment radii. Firms located within 2 miles of a reclassified toxic plant experience a premium increase of roughly 14% compared with firms in the 3 to 5 mile control ring.

This evidence speaks directly to the pollution signaling mechanism. When the perceived carcinogenic risk rises, premiums rise as well, even though the pool does not change. The reclassification results also help reconcile the premium puzzle. Cross sectionally, more polluted places have higher premiums, and when perceived risk rises through

reclassification, premiums rise. Around toxic plant entry, by contrast, premiums fall because the diversification effect from pool expansion can dominate in the short run. Taken together, the results suggest that pollution signaling works in the expected direction, but its effect can be offset by diversification when entry expands the local pool.

6 Conclusion

Does toxic plant entry raise or lower the health insurance costs of nearby employers? For this simple question we find a counterintuitive answer. Cross-sectionally, firms in more polluted areas pay higher premiums, and standard theory predicts that entry should push costs even higher as environmental exposure worsens worker health. We find the opposite. Premiums for incumbent firms within 2 miles of a new toxic plant fall by approximately 10%, a pattern that reflects local insurance market structure rather than an absence of health risk. Because firm-specific underwriting is costly, insurers pool small employers geographically and price premiums based on local pool composition. Entry expands the pool, diversifies idiosyncratic risk, and lowers premiums for pooled incumbents before cumulative health costs materialize in claims data. The heterogeneity supports this mechanism: declines are larger when entrants bring more workers, attenuated when entrants are more toxic, and absent among self-insured firms that bypass pooled pricing. Non-toxic plant entry generates declines roughly twice as large, while carcinogen reclassifications raise premiums by 14%, providing separate evidence on the two forces predicted by the framework. The broader implication is that a firm's health insurance costs depend not only on its own workforce and plan design, but also on local industrial activity and insurers' perceptions of environmental health risk.

References

Aghamolla, C., J. Jain, and R. T. Thakor (2025, June). When private equity comes to town: The local economic consequences of rising healthcare costs. Working paper, University of Minnesota.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*(3), 488–500.

Almazan, A., A. De Motta, and S. Titman (2007). Firm location and the creation and utilization of human capital. *Review of Economic Studies 74*(4), 1305–1327.

Brot-Goldberg, Z., Z. Cooper, S. V. Craig, L. R. Klarnet, I. Lurie, and C. L. Miller (2024). Who pays for rising health care prices? Evidence from hospital mergers. Working Paper 32301, National Bureau of Economic Research.

Bundorf, M. K., B. Herring, and M. V. Pauly (2010). Health risk, income, and employment-based health insurance. In *Forum for Health Economics and Policy*, Volume 13, pp. 1–35. De Gruyter.

Butts, K. (2023). JUE insight: Difference-in-differences with geocoded microdata. *Journal of Urban Economics 133*, 103493.

Cooper, Z., S. V. Craig, M. Gaynor, and J. Van Reenen (2019). The price ain't right? Hospital prices and health spending on the privately insured. *Quarterly Journal of Economics 134*(1), 51–107.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review 105*(2), 678–709.

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

Currie, J., J. G. Zivin, J. Mullins, and M. Neidell (2014). What do we know about short- and long-term effects of early-life exposure to pollution? *Annual Review of Resource Economics 6*(1), 217–247.

Dafny, L., M. Duggan, and S. Ramanarayanan (2012). Paying a premium on your premium? Consolidation in the US health insurance industry. *American Economic Journal: Economic Policy 4*(1), 153–177.

Dafny, L. S. (2010). Are health insurance markets competitive? *American Economic Review 100*(4), 1399–1431.

Diamond, R. and T. McQuade (2019). Who wants affordable housing in their backyard? an equilibrium analysis of low-income property development. *Journal of Political Economy 127*(3), 1063–1117.

Dougal, C., C. A. Parsons, and S. Titman (2022). The geography of value creation. *Review of Financial Studies 35*(9), 4201–4248.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and É. L. Pinder (2003). The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine 138*(4), 273–287.

Ganduri, R. and G. Maturana (2024). Do property rehabs affect neighboring property prices? *Journal of Urban Economics 143*, 103694.

Gao, J., S. Ge, L. Schmidt, and C. J. T. Trillo (2025). *How do health insurance costs affect low- and high-income workers?*

Giroud, X. and H. M. Mueller (2017). Firm leverage, consumer demand, and employment losses during the Great Recession. *Quarterly Journal of Economics 132*(1), 271–316.

Giroud, X. and H. M. Mueller (2019). Firms' internal networks and local economic shocks. *American Economic Review 109*(10), 3617–3649.

Gormley, T. A. and D. A. Matsa (2011). Growing out of trouble? Corporate responses to liability risk. *Review of Financial Studies 24*(8), 2781–2821.

Graff Zivin, J. and M. Neidell (2013). Environment, health, and human capital. *Journal of Economic Literature 51*(3), 689–730.

Greenstone, M., R. Hornbeck, and E. Moretti (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Quarterly Journal of Economics 125*(2), 536–598.

LaPoint, C. and S. Yale (2022). Property tax sales, private capital, and gentrification in the us. *Private Capital, and Gentrification in the US (September 15, 2022)*.

Lin, H. and Y. Zhou (2026). Distributional effect of insurer concentration on premiums: Evidence from employer-sponsored health insurance. *The RAND Journal of Economics*. Forthcoming.

Linehan, K. (2010, December). Self-insurance and the potential effects of health reform on the small-group market. Issue Brief No. 840, National Health Policy Forum. Accessed 2026-02-25.

Ouimet, P. and G. Tate (2025). Who does it hurt if everyone gets the same? Implications of standardizing nonwage benefits for the labor market. *Working Paper*.

Pauly, M. V. (1970). The welfare economics of community rating. *Journal of Risk and Insurance 37*(3), 407–418.

Rothschild, M. and J. E. Stiglitz (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics 90*(4), 629–649.

Tong, J. (2024). Health care costs and corporate investment. *Review of Financial Studies 37*(4), 1078–1127.

U.S. House of Representatives (2025, December). H. rept. 119-408: Self-Insurance Protection Act. Committee Report. Accessed 2026-02-25.

van Binsbergen, J. H., J. F. Cocco, M. Grotteria, and S. L. Naaraayanan (2025). The impact of carcinogenic risk exposure on housing values: Estimates from chemical reclassifications. *Review of Financial Studies*. Forthcoming.

Figure 1: Healthcare Premiums around Toxic Plants

This figure plots the baseline association between local toxic exposure and healthcare premiums of nearby firms. The coefficients correspond to regressions of log(average healthcare premium) on standardized measures of aggregate toxic exposure within 0 to 2 miles and 2 to 4 miles of the firm location. Exposure measures include log(TRI Pounds), log(RSEI Cancer Hazard), and log(Number of Toxic Plants), each standardized to have mean zero and unit standard deviation. The regressions include zipcode fixed effects, year fixed effects, and county × industry × size fixed effects, where industry is defined at the 4 digit NAICS level and size is defined using five size bins based on the number of employees at the firm. The figure reports point estimates with 95% confidence intervals, and standard errors are clustered at the zipcode level.

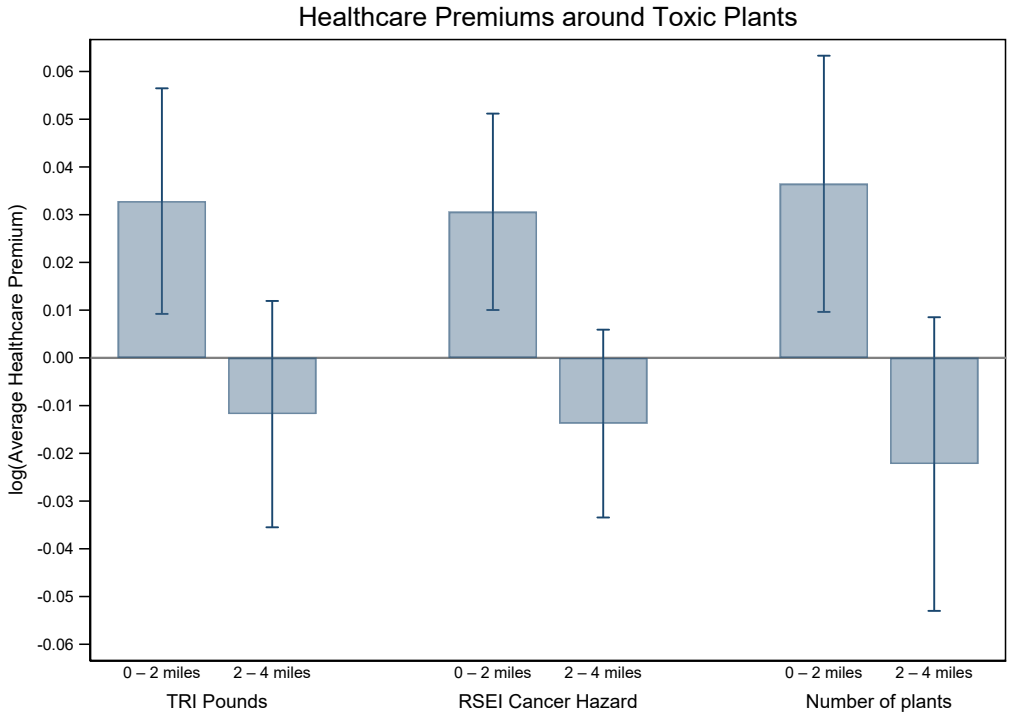


Figure 2: Dynamics around Toxic Plant Entry

This figure plots the dynamic difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' average healthcare premiums. The coefficients are obtained from the following event study specification estimated within a 0 to 3 mile treatment ring:

$$log(Premium)_{ipt} = \sum_{\tau \neq -1} \beta_\tau \times \mathbb{1}(EventTime_{pt} = \tau) \times \mathbb{1}(Distance_{ip} < 3) + \gamma_i + \delta_{pt} + \varepsilon_{ipt}$$

for incumbent firm i, toxic plant p, in year t. The specification includes fixed effects for the incumbent firm (γ_i) and toxic plant $\times$ year (δ_{pt}). The plotted coefficients (β_τ) measure the change in log healthcare premiums for incumbents located within 3 miles of the toxic plant ("treated") relative to incumbents located between 3 to 5 miles ("control"), normalized to event time $\tau = -1$. Event time (τ) is defined relative to the year of toxic plant opening. The figure reports point estimates and 95% confidence intervals with standard errors clustered at the toxic plant level.

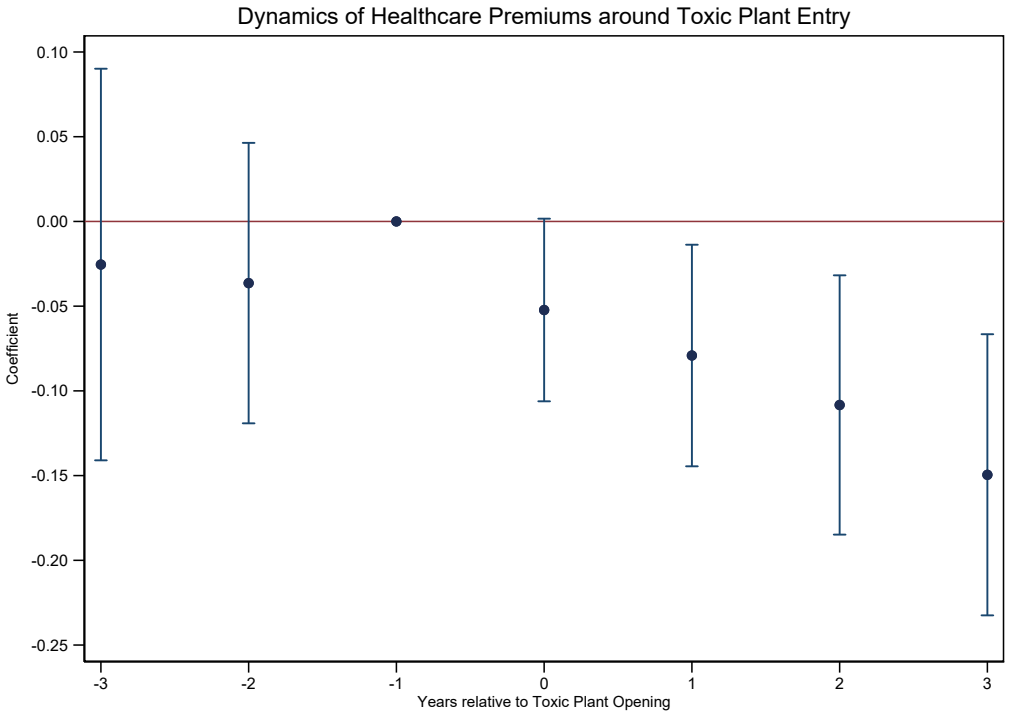


Figure 3: Estimates as a function of distance

This figure plots the distance gradient difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' average healthcare premiums. The coefficients are obtained from the following specification using one-mile distance bins around the entering toxic plant:

$$log(Premium)_{ipt} = \sum_{d \in \{0\text{–}1,\, 1\text{–}2,\, 2\text{–}3,\, 3\text{–}4\}} \beta_d \times \mathbb{1}(Post_{it} = 1) \times \mathbb{1}(Distance_{ip} \in d) + \gamma_i + \delta_{pt} + \varepsilon_{ipt}$$

for incumbent firm i, toxic plant p, in year t. The specification includes fixed effects for the incumbent firm (γ_i) and toxic plant $\times$ year (δ_{pt}). The plotted coefficients (β_d) measure the change in log healthcare premiums for incumbents located in distance bin d miles from the toxic plant relative to incumbents located between 4 to 5 miles control group, comparing the post-entry period to the pre-entry period. Post ($Post_{pt}$) equals one for incumbent firm i in year t in the post-entry period, that is, after the toxic plant opening year. The figure reports point estimates and 95% confidence intervals with standard errors clustered at the toxic plant level.

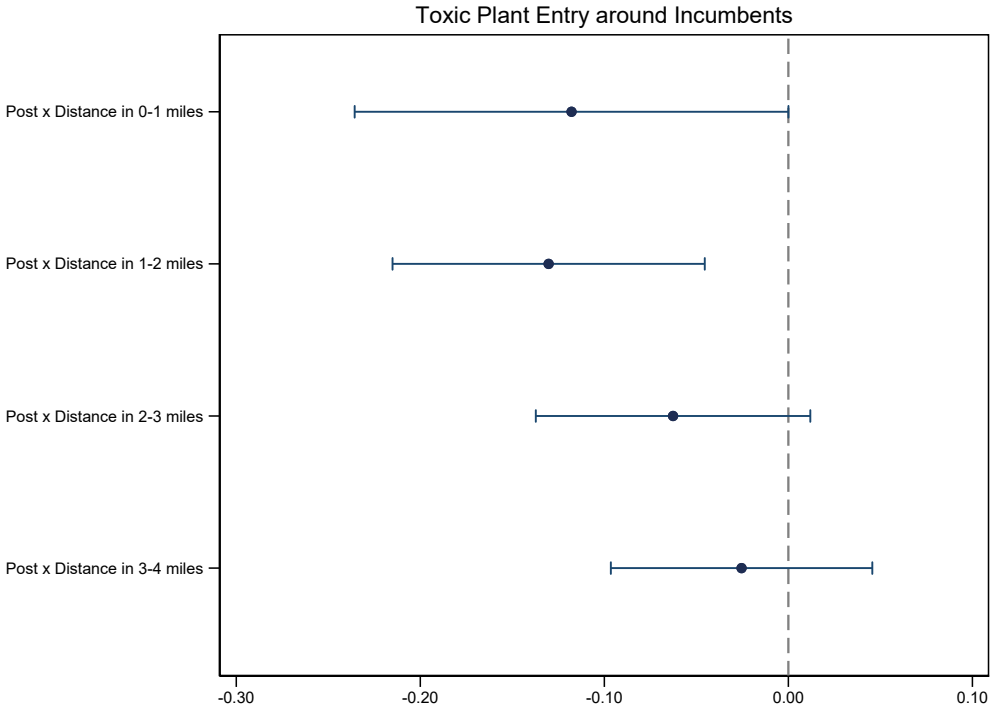


Table 1: Panel A: Summary Statistics for Incumbent Firms

This table reports summary statistics for incumbent firms located near toxic plants. The sample includes incumbent firm-year observations from three years before to three years after toxic plant entry. Summary statistics are reported separately for incumbent firms located within 0 to 1 miles, 0 to 2 miles, 0 to 3 miles, and 3 to 5 miles of a toxic plant. *Average healthcare premium* is the annual total health insurance premium paid per participant by the incumbent firm. *Number of plan participants* is the total number of participants enrolled in the firm's health insurance plan in a given year. *Self insured* equals 1 if the incumbent firm bears medical claims risk directly and 0 otherwise. *Number of insurance companies* is the count of distinct insurance carriers serving firms within the specified distance radius. *Sales* measures total plant revenue expressed in millions of U.S. dollars. The table reports the mean, standard deviation, 25th percentile, median, 75th percentile, and number of observations for each variable.

	Mean	**StDev**	**P25**	**P50**	**P75**	**N**
			0 to 1 miles			
Average healthcare premium ($)	4079.55	3974.1	863.95	2911.93	6288.52	4228
Number of plan participants	830.74	2299.93	125	270	661.5	4228
Self insured (%)	20.44	40.34	0	0	0	4228
Number of insurance companies	3.31	3.23	1	2	4	4228
Sales ($ million)	61.04	289.04	8.4	21.1	44.26	3029
			0 to 2 miles			
Average healthcare premium ($)	4093.8	4021.59	896.42	2860.47	6241.74	10160
Number of plan participants	966.3	2684.15	126	282	719.5	10160
Self insured (%)	20.28	40.21	0	0	0	10160
Number of insurance companies	6.3	7.43	2	3	8	10160
Sales ($ million)	65.33	374.68	6.9	19.4	41.29	7191
			0 to 3 miles			
Average healthcare premium ($)	4191.4	4130.55	919.88	2946.03	6321.5	17796
Number of plan participants	1009.46	2766.96	131	297	766	17796
Self insured (%)	19.66	39.75	0	0	0	17796
Number of insurance companies	12.84	20.03	2	5	14	17796
Sales ($ million)	80.45	543.4	6.5	19.17	42	12433
			3 to 5 miles			
Average healthcare premium ($)	4274.91	4098.12	1026.46	2999.94	6474.48	17558
Number of plan participants	1173.31	3456.07	121	286	747	17558
Self insured (%)	18.70	38.99	0	0	0	17558
Number of insurance companies	12.96	16.11	3	7	16	17558
Sales ($ million)	81.29	524.93	5.39	17.3	41.57	11974

Panel B: Summary Statistics for Toxic Plant Entrants

This table reports summary statistics for toxic plant entrants in their year of entry. *TRI pounds* measures the total quantity of toxic releases reported to the EPA Toxic Release Inventory in the entry year. *RSEI cancer hazard* is the EPA Risk-Screening Environmental Indicator cancer hazard score in the entry year, which captures the toxicity weighted exposure risk from reported emissions. *Number of employees* measures total employment at the toxic plant. *Sales* measures total plant revenue expressed in millions of U.S. dollars. *Number of incumbents within 1 mile*, *within 2 miles*, and *within 3 miles* count the number of distinct incumbent firms located within the respective distance rings of the toxic plant. *Number of incumbents in 3 to 5 miles* counts the number of distinct incumbent firms located between 3 and 5 miles of the toxic plant. The table reports the mean, standard deviation, 25th percentile, median, 75th percentile, and number of observations for each variable.

	Mean	**StDev**	**P25**	**P50**	**P75**	**N**
TRI pounds	35182.28	116404.4	9.1	1130	16168.5	7963
RSEI cancer hazard (million)	30671.86	140746.9	0.01	0.49	42.10	7963
Number of employees	99.32	285.99	10	35	100	3694
Sales ($ million)	43.94	310.53	1.7	6.7	22.54	2098
Number of incumbents within 1 mile	0.25	0.81	0	0	0	7995
Number of incumbents within 2 miles	0.73	1.95	0	0	1	7995
Number of incumbents within 3 miles	1.61	5.33	0	1	2	7995
Number of incumbents in 3 to 5 miles	1.82	4.97	0	0	2	7995

Table 2: Toxic Plant Entry (DiD)

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variable is $Post \times \mathbb{1}(Distance < X)$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	-0.079***	-0.104***	-0.144***	-0.143**	-0.156**
	(0.028)	(0.037)	(0.051)	(0.057)	(0.065)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	26,012	19,178	16,307	15,367	14,534
Adjusted R^2	0.808	0.810	0.812	0.810	0.813

Table 3: Effect of Toxic Plant Entry on Plan Participation

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' plan participants. The dependent variable is log(Number of plan participants), defined as the log of the number of participants enrolled in the incumbent firm's health insurance plan. The independent variable is $Post \times \mathbb{1}(Distance < X)$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Number of plan participants)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	0.011	0.022	0.017	-0.006	-0.017
	(0.022)	(0.029)	(0.038)	(0.045)	(0.046)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	26,012	19,178	16,307	15,367	14,534
Adjusted R^2	0.928	0.930	0.930	0.930	0.932

Table 4: Toxic Plant Entry: Number of employees in the toxic plant

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing entry by small versus large toxic plants. Plant size is measured by the number of employees at the toxic plant in the entry year and split at the sample median. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times Less\ Employees$ and $Post \times \mathbb{1}(Distance < X) \times More\ Employees$, where *Post* equals one for years after the toxic plant opening and zero otherwise. *Less Employees* and *More Employees* indicate whether the number of employees at the toxic plant in the entry year is below or above the sample median, respectively. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles) × Less employees	0.020	0.038	-0.075	-0.091	-0.103
	(0.076)	(0.097)	(0.105)	(0.127)	(0.139)
Post × 𝟙(Distance < X miles) × More employees	-0.158***	-0.174**	-0.234**	-0.240*	-0.241
	(0.069)	(0.086)	(0.104)	(0.128)	(0.155)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	9,146	6,963	5,872	5,510	5,159
Adjusted R^2	0.781	0.780	0.786	0.780	0.784

Table 5: Toxic Plant Entry: Cancer causing emissions by the toxic plant

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing entry by low versus high toxicity plants. Toxicity is measured using the RSEI Cancer Hazard score, which captures the toxic plant's emissions of cancer causing chemicals in the entry year. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times Low\ Toxicity$ and $Post \times \mathbb{1}(Distance < X) \times High\ Toxicity$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. $Low\ Toxicity$ and $High\ Toxicity$ indicate whether toxicity in the entry year is below or above the sample median, respectively. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles) × Low toxicity	-0.120***	-0.156***	-0.245***	-0.266***	-0.282***
	(0.040)	(0.053)	(0.076)	(0.087)	(0.098)
Post × 𝟙(Distance < X miles) × High toxicity	-0.020	-0.027	-0.006	0.014	0.014
	(0.041)	(0.053)	(0.064)	(0.073)	(0.086)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	25,542	18,752	15,982	15,072	14,252
Adjusted R^2	0.807	0.808	0.810	0.808	0.811

Table 6: Toxic Plant Entry: Self-Insured Incumbents

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing self insured versus non self insured incumbent firms. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times Self\ Insured$ and $Post \times \mathbb{1}(Distance < X) \times Not\ Self\ Insured$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. $Self\ Insured$ indicates that the incumbent firm is self insured and $Not\ Self\ Insured$ indicates that the incumbent firm is not self insured. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles) × Self Insured	0.155	0.346	0.101	0.133	0.221
	(0.128)	(0.264)	(0.115)	(0.136)	(0.028)
Post × 𝟙(Distance < X miles) × not Self Insured	-0.069**	-0.114***	-0.186***	-0.183***	-0.144**
	(0.028)	(0.037)	(0.056)	(0.063)	(0.070)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	20,925	15,340	15,493	14,586	11,642
Adjusted R^2	0.846	0.847	0.818	0.816	0.852

Table 7: Toxic Plant Entry: Single Insurance Company in the treatment ring

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing treatment rings served by a single insurance company versus multiple insurance companies. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times Single\ Insurance\ Company$ and $Post \times \mathbb{1}(Distance < X) \times Multiple\ Insurance\ Companies$, where *Post* equals one for years after the toxic plant opening and zero otherwise. *Single Insurance Company* indicates that the treatment ring is served by one insurance company and *Multiple Insurance Companies* indicates that the treatment ring is served by more than one insurance company. X takes values 3, 2, or 1 mile (columns (1) to (3)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)		
Treatment (Distance in miles)	**3**	**2**	**1**
	(1)	(2)	(3)
Post × 𝟙(Distance < X miles) × Single Insurance Company	0.139*	-0.032	-0.033
	(0.084)	(0.091)	(0.123)
Post × 𝟙(Distance < X miles) × Multiple Insurance Companies	-0.082***	-0.116***	-0.177**
	(0.028)	(0.039)	(0.078)
Incumbent Firm FE	YES	YES	YES
Plant × Year FE	YES	YES	YES
Observations	21,914	16,068	12,200
Adjusted R^2	0.841	0.845	0.850

Table 8: Toxic Plant Entry: Incumbents in same county as the toxic plant

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing incumbent firms located in the same county with those located in a different county as the entering toxic plant. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times Same\ County$ and $Post \times \mathbb{1}(Distance < X) \times Different\ County$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. $Same\ County$ indicates that the incumbent firm is located in the same county as the entering toxic plant and $Different\ County$ indicates that the incumbent firm is located in a different county. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles) × Same county	-0.070*	-0.110**	-0.153**	-0.137*	-0.114
	(0.039)	(0.051)	(0.062)	(0.076)	(0.084)
Post × 𝟙(Distance < X miles) × Different county	0.049	0.029	-0.211	-0.154	0.187*
	(0.081)	(0.122)	(0.233)	(0.256)	(0.097)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	15,164	10,891	9,122	8,550	8,073
Adjusted R^2	0.828	0.823	0.826	0.822	0.826

Table 9: Non-Toxic Plant Entry (DiD)

This table reports difference-in-differences estimates of the effect of non-toxic plant entry on incumbent firms' healthcare premiums. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variable is $Post \times \mathbb{1}(Distance < X)$, where $Post$ equals one for years after the non-toxic plant opening and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the non-toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	-0.317***	-0.240*	-0.400**	-0.355*	-0.433**
	(0.121)	(0.137)	(0.187)	(0.181)	(0.190)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	4,476	3,264	2,519	2,214	1,932
Adjusted R^2	0.691	0.671	0.641	0.633	0.638

Table 10: Toxic Plant Entry versus Non-Toxic Plant entry (DiDiD)

This table reports triple difference estimates comparing the effects of toxic plant entry versus non-toxic plant entry on incumbent firms' healthcare premiums. Each toxic plant is matched to a nearby non-toxic plant that opened in the same year and is in the same employee size bucket. Employee size buckets are defined based on the sample median, with small (employees ≤ 75) and large (employees > 75). Among eligible matches, each toxic plant is paired with the nearest non-toxic plant located 10 to 20 miles away, to avoid spatial overlap in incumbent firms while maintaining similar local conditions. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The sample includes properties within a 5-mile radius of both toxic and non-toxic matched plants. The independent variable of interest is the triple interaction $Post_{it} \times \mathbb{1}(Distance_{ij(p)} < X) \times \mathbb{1}(Toxic)_{j(p)}$, and coefficients are obtained from the following specification:

$$\begin{aligned}\log(Premium)_{ij(p)t} = {} & \beta_1 \times \mathbb{1}(Distance_{ij(p)} < X) \times \mathbb{1}(Toxic)_{j(p)} \\ & + \beta_2 \times Post_{it} \times \mathbb{1}(Distance_{ij(p)} < X) \times \mathbb{1}(Toxic)_{j(p)} \\ & + \gamma_{i(p)} + \gamma_{Ring(p)t} + \gamma_{j(p)t} + \theta X_{it} + \varepsilon_{ij(p)t}\end{aligned}$$

Here, i indexes incumbent firms, $j(p)$ indexes entrant plants j within matched pair p, and t indexes year. $Post_{it}$ equals one for years after the entrant plant opening and zero otherwise. $\mathbb{1}(Toxic)_{j(p)}$ equals one for toxic entrants and zero for matched non-toxic entrants. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance_{ij(p)} < X)$ equals one for incumbent firms within X miles of the entrant plant and equals zero for incumbent firms located between 3 and 5 miles of the same entrant plant. $\gamma_{i(p)}$ are incumbent firm $\times$ matched-pair fixed effects. $\gamma_{j(p)t}$ are entrant plant $\times$ year $\times$ matched pair fixed effects. $\gamma_{Ring(p)t}$ are distance ring $\times$ year $\times$ matched-pair fixed effects, where distance rings are defined as 0 to 1, 1 to 2, 2 to 3, 3 to 4, and 4 to 5 miles from the entrant plant. All specifications control for the number of plan participants. Standard errors clustered at the plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post $\times$ $\mathbb{1}$(Distance $<$ X miles) $\times$ $\mathbb{1}$(Toxic)	0.384***	0.409	0.362***	0.335***	0.334***
	(0.127)	(0.254)	(0.118)	(0.117)	(0.117)
Incumbent Firm $\times$ Matched-Pair FE	YES	YES	YES	YES	YES
Year $\times$ Distance $\times$ Matched-Pair FE	YES	YES	YES	YES	YES
Year $\times$ Plant $\times$ Matched-Pair FE	YES	YES	YES	YES	YES
Observations	14,154	10,920	9,501	9,041	8,953
Adjusted R^2	0.740	0.732	0.742	0.738	0.739

Table 11: Chemical Reclassification Event (DiD)

This table reports difference-in-differences estimates of the effect of chemical reclassification of toxic plant on nearby firms' healthcare premiums. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the firm. The independent variable is $Post\ Reclassification \times \mathbb{1}(Distance < X)$, where $Post\ Reclassification$ equals one for years after the reclassification event year and zero otherwise. The reclassification event year is the year in which a chemical is newly listed or upgraded as carcinogenic in the National Toxicology Program Report on Carcinogens, for a chemical used by the toxic plant. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for nearby firms within X miles of the relevant toxic plant (treatment ring) and equals zero for firms located between 3 and 5 miles of the same plant (the control ring). All specifications include firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post Reclassification × 𝟙(Distance < X miles)	0.067	0.143***	0.138**	0.187***	0.215***
	(0.044)	(0.055)	(0.057)	(0.065)	(0.070)
Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	7,346	5,184	4,093	3,647	3,157
Adjusted R^2	0.811	0.803	0.800	0.802	0.810

Internet Appendix

A Theoretical Framework

This appendix formalizes the mechanisms underlying our empirical findings. We develop a model where risk-averse insurers price health insurance using mean-variance preferences over pool-level outcomes. When firm-specific risk assessment is costly, insurers pool small firms geographically. We show that larger pools reduce idiosyncratic health risk through diversification, lowering premiums for all firms, while observable pollution signals predict higher future health costs, creating an offsetting force.

A.1 Setup

Environment. Consider a local insurance market with N incumbent firms, each employing L_i workers. Total employment is $L = \sum_{i=1}^{N} L_i$.

Worker Health. Each worker j at firm i has health type $\theta_{ij,t}$ representing expected medical costs, drawn i.i.d. from distribution $F(\theta)$ with mean μ and variance $\sigma^2 > 0$. Health evolves according to:

$$\theta_{ij,t+1} = \theta_{ij,t} + \delta \cdot \text{Exposure}_{ij,t} + \epsilon_{ij,t+1} \tag{A.1}$$

where $\delta > 0$ captures how toxic exposure worsens health, and $\epsilon_{ij,t+1}$ is an idiosyncratic shock with $[\epsilon_{ij,t+1}] = 0$ and $(\epsilon_{ij,t+1}) = \sigma_\epsilon^2$.

Information Friction. Insurers observe firm sizes L_i and environmental exposure E_i (from EPA emissions data), but cannot observe individual worker health θ_{ij} or firm-specific average health $\bar{\theta}_i = \frac{1}{L_i}\sum_{j=1}^{L_i} \theta_{ij}$ due to HIPAA privacy protections and costly underwriting.

Costly Firm-Specific Pricing. To assess and price each firm individually requires firm-specific underwriting with fixed cost $F > 0$. For small firms ($L_i < 50$ or 100 employees), this cost exceeds the benefit. Following Pauly (1970), insurers therefore pool small firms within local geographic areas.

A.2 Insurer Pricing: Mean-Variance Framework

The insurer is risk-averse and prices insurance using mean-variance preferences over the pool's realized claims. For a pool with L workers, the insurer's value function is:

$$V = [\bar{\theta}] - \lambda \cdot (\bar{\theta}) \tag{A.2}$$

where $\lambda > 0$ reflects risk aversion. The insurer sets premiums to cover expected costs plus a risk premium for variance:

$$P = \alpha \cdot [\bar{\theta}] + \lambda \cdot (\bar{\theta}) + \beta \cdot \bar{E} + \gamma \tag{A.3}$$

where:

- $\alpha > 0$ converts expected health costs to dollars
- $\lambda \cdot (\bar{\theta})$ is the risk premium for bearing idiosyncratic uncertainty

- $\beta \cdot \bar{E}$ reflects how observable emissions $\bar{E}$ predict future health deterioration
- γ captures fixed administrative costs

By the law of large numbers, with i.i.d. health shocks across workers:

$$(\bar{\theta}) = \frac{\sigma^2}{L} \tag{A.4}$$

Larger pools have lower variance, reducing the risk premium insurers require.

A.3 Plant Entry Effects

At time $t = 0$, a new plant enters with L_{new} workers and emissions level E_{new}. Post-entry, the pool has $L' = L + L_{\text{new}}$ workers.

A.3.1 Diversification Effect

The post-entry pool variance is:

$$(\bar{\theta}') = \frac{\sigma^2}{L + L_{\text{new}}} < \frac{\sigma^2}{L} = (\bar{\theta}) \tag{A.5}$$

reduction in variance lowers the risk premium:

$$\Delta P_{\text{diversification}} = \lambda \sigma^2 \left(\frac{1}{L + L_{\text{new}}} - \frac{1}{L} \right) < 0 \tag{A.6}$$

Economic Intuition: More workers diversify away idiosyncratic health risk. The insurer faces less uncertainty about realized claims, so charges a lower risk premium. This benefit is shared by all firms in the pool.

A.3.2 Signal Effect

Insurers observe the entrant's emissions E_{new} (from EPA TRI data) and use them to forecast future health costs. The post-entry average emissions signal becomes:

$$\bar{E}' = \frac{L \cdot \bar{E} + L_{\text{new}} \cdot E_{\text{new}}}{L + L_{\text{new}}} \tag{A.7}$$

The signal-driven premium change is:

$$\Delta P_{\text{signal}} = \beta \cdot \frac{L_{\text{new}}}{L + L_{\text{new}}} \cdot (E_{\text{new}} - \bar{E}) \tag{A.8}$$

This can be positive (if $E_{\text{new}} > \bar{E}$) or negative (if $E_{\text{new}} < \bar{E}$).

Economic Intuition: Observable emissions serve as a signal of future health deterioration under information asymmetry (Akerlof, 1970; Rothschild and Stiglitz, 1976). High-emission entrants increase the pool's expected future costs even before medical claims materialize.

A.3.3 Total Effect

The total premium change combines both channels:

$$\Delta P_{\text{total}} = \underbrace{\lambda\sigma^2\left(\frac{1}{L+L_{\text{new}}} - \frac{1}{L}\right)}_{(-)\text{ diversification}} + \underbrace{\beta \cdot \frac{L_{\text{new}}}{L+L_{\text{new}}} \cdot (E_{\text{new}} - \bar{E})}_{(+/-)\text{ signal}} \tag{A.9}$$

Proposition 1 (Main Predictions)**.** *Equation (A.9) generates the following predictions:*

(1) Baseline Effect: *For typical entrants where $E_{new} \approx \bar{E}$, the diversification effect would dominate:*

$$\frac{\partial P}{\partial L_{new}} < 0 \tag{A.10}$$

(2) Employment Size: *Larger entrants would create stronger diversification. The marginal benefit of adding workers is positive.*

(3) Toxicity Heterogeneity: *High-emission entrants would generate smaller (or positive) effects:*

$$\frac{\partial(\Delta P)}{\partial E_{new}} = \beta \cdot \frac{L_{new}}{L+L_{new}} > 0 \tag{A.11}$$

(4) Non-Toxic Entrants: *When $E_{new} = 0$, the signal effect would reinforce diversification:*

$$\Delta P_{non\text{-}toxic} = \Delta P_{diversification} + \beta \cdot \frac{L_{new}}{L+L_{new}} \cdot (-\bar{E}) \tag{A.12}$$

A.4 Self-Insured Validation

Fully-Insured Firms purchase insurance and pay the pooled premium P from equation (A.3). Their costs depend on pool-level variance $(\bar{\theta}) = \sigma^2/L$.
Self-Insured Firms pay medical claims directly:

$$P_i^{\text{self}} = \alpha \cdot [\bar{\theta}_i] + \lambda_i \cdot (\bar{\theta}_i) + \text{Admin}_i \tag{A.13}$$

where $(\bar{\theta}_i) = \sigma^2/L_i$ depends only on firm i's own size.

Proposition 2 (Self-Insured Placebo)**.** *Plant entry should have zero effect on self-insured firms' health costs:*

$$\frac{\partial P^{self}}{\partial L_{new}} = 0, \quad \frac{\partial P^{fully}}{\partial L_{new}} \neq 0 \tag{A.14}$$

If effects operated through labor market competition, provider supply expansion, or general economic growth, both firm types would be affected. Only insurance pooling would generate differential effects for fully-insured versus self-insured firms.

A.5 Information Shocks

When chemical carcinogenicity is publicly reclassified (for example, through National Toxicology Program reports), insurers would receive new information about health risks without any change in pool size.

The signal strength parameter would increase from β to β', where $\beta' > \beta$. This would generate:

$$\Delta P_{\text{info shock}} = (\beta' - \beta) \cdot \bar{E} > 0 \tag{A.15}$$

Proposition 3 (Risk Revelation)**.** *Exogenous information shocks that reveal higher health risks should generate immediate premium increases without diversification effects.*

Table A.I: Effect of Toxic Plant Entry on Insurance Plan Offerings

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' plan offerings. The dependent variable in Panel A is *PPO Plan*, an indicator equal to 1 if the incumbent firm offers a PPO plan in year t and 0 otherwise. The dependent variable in Panel B is *HMO Plan*, an indicator equal to 1 if the incumbent firm offers an HMO plan in year t and 0 otherwise. The independent variable of interest is $Post \times \mathbb{1}(Distance < X)$, where *Post* equals one for years after the toxic plant opening and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for log(Number of plan participants). Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

Panel A

	PPO Plan				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	0.006	0.007	-0.001	0.014	0.006
	(0.018)	(0.025)	(0.028)	(0.032)	(0.041)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	10,857	7,972	6,717	6,253	5,890
Adjusted R^2	0.683	0.684	0.678	0.680	0.669

Panel B

	HMO Plan				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	0.019	0.027	0.005	-0.003	-0.0005
	(0.019)	(0.021)	(0.024)	(0.027)	(0.034)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	10,557	7,737	6,494	6,028	5,693
Adjusted R^2	0.788	0.790	0.792	0.790	0.788

Table A.II: Toxic Plant Entry: Excluding Nearby Toxic Plant Exits

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums. In Panel A, the sample excludes incumbent firms that experienced a toxic plant exit within 0 to 3 miles during the event window from three years before to three years after toxic plant entry. In Panel B, the sample excludes incumbent firms that experienced a toxic plant exit within 5 miles during the event window. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variable is $Post \times \mathbb{1}(Distance < X)$, where $Post$ equals one for years after the toxic plant opening and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

Panel A

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	-0.139**	-0.154**	-0.193*	-0.158	-0.150
	(0.063)	(0.077)	(0.099)	(0.118)	(0.139)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	9,466	7,695	6,799	6,495	6,226
Adjusted R^2	0.807	0.813	0.815	0.813	0.812

Panel B

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles)	-0.176*	-0.190	-0.299**	-0.283*	-0.304
	(0.098)	(0.119)	(0.150)	(0.170)	(0.196)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	4,443	3,456	2,898	2,701	2,537
Adjusted R^2	0.780	0.779	0.779	0.775	0.771

Table A.III: Toxic Plant Entry: Incumbent firm size

This table reports difference-in-differences estimates of the effect of toxic plant entry on incumbent firms' healthcare premiums, comparing smaller versus larger incumbent firms. The sample is restricted to incumbent firms with fewer than 150 employees. Incumbent firm size is measured by the number of employees at the firm and grouped into two bins using the subsample median. The dependent variable is log(Healthcare Premium), defined as the log of annual total health insurance premiums paid per participant by the incumbent firm. The independent variables are $Post \times \mathbb{1}(Distance < X) \times \mathbb{1}(\leq 100\ Employees)$ and $Post \times \mathbb{1}(Distance < X) \times \mathbb{1}(100\ to\ 150\ Employees)$, where *Post* equals one for years after the toxic plant opening and zero otherwise. $\mathbb{1}(\leq 100\ Employees)$ equals one for incumbent firms with at most 100 employees and zero otherwise. $\mathbb{1}(100\ to\ 150\ Employees)$ equals one for incumbent firms with between 100 and 150 employees and zero otherwise. X takes values 3, 2, 1.5, 1.25, or 1 mile (columns (1) to (5)). The indicator $\mathbb{1}(Distance < X)$ equals one for incumbent firms within X miles of the entering plant (treatment ring) and equals zero for incumbent firms located between 3 and 5 miles of the same plant (the control ring). All specifications include incumbent firm fixed effects and toxic plant × year fixed effects, and control for the number of plan participants. Standard errors clustered at the toxic plant level are reported in parentheses. *** denotes significance at 1%, ** at 5%, and * at 10%.

	log(Healthcare Premium)				
Treatment (Distance in miles)	**3**	**2**	**1.5**	**1.25**	**1**
	(1)	(2)	(3)	(4)	(5)
Post × 𝟙(Distance < X miles) × 𝟙(≤ 100 employees)	-0.101**	-0.138**	-0.197**	-0.186	-0.194
	(0.029)	(0.056)	(0.097)	(0.121)	(0.147)
Post × 𝟙(Distance < X miles) × 𝟙(100-150 employees)	0.023	0.012	-0.134	-0.181	-0.090
	(0.062)	(0.101)	(0.138)	(0.161)	(0.144)
Incumbent Firm FE	YES	YES	YES	YES	YES
Plant × Year FE	YES	YES	YES	YES	YES
Observations	10,246	7,182	6,049	5,659	5,332
Adjusted R^2	0.841	0.837	0.836	0.832	0.839