Drill Here, Drill Now, Pay Less: Labor Market and Air Quality Trade-offs of Oil and Gas Extraction*

Thuy Nguyen†

Job Market Paper

Click here to see the latest version

October 2, 2025

Abstract

The advent of hydraulic fracturing and horizontal drilling has transformed the oil and gas industry. While the fracking boom generates economic gains, the distribution of these benefits and the associated environmental harms remain underexplored. This work uses spatial econometric modeling to investigate the local and regional labor market and air quality trade-offs of oil and gas extraction in Pennsylvania, Ohio, and West Virginia.

The results reveal a double disparity that is as much generational as it is geographical. While younger, mobile workers, particularly non-locals, capture most of the economic benefits from oil and gas, local seniors living near production activity bear the bulk of the environmental health costs. Specifically, I find that oil and gas do not create jobs for local residents who work in their home county. Only a subset of locals, those who commute to jobs outside their home county, experience measurable gains, averaging eight additional jobs and $525,000 in earnings per 100,000 barrels of oil equivalent (BOE) produced. In comparison, the same level of production generates about 12 jobs and $1.2 million in earnings for non-local workers, amounting to about 40% more jobs and more than double the earnings compared to commuting locals.

In contrast, air quality impacts are the most severe when production occurs nearby. An additional 100,000 BOE produced within 1 to 2 km of a gridded cell increases its $PM_{2.5}$ concentration by 1.2 $\mu g/m^3$, on average. Leveraging fine-grained, age-specific population and vital statistics, I estimate that oil and gas extraction resulted in $62 billion in health damage—roughly one-quarter of total oil and gas revenue—in the tri-state region between 2001 and 2020. Local seniors suffered 80% of the health burden.

JEL Codes: J61, Q32, Q35, Q51, Q53, R12

Keywords: Hydrocarbon extraction, labor market, air pollution, spatial and distributional impacts

*I am deeply grateful for the years of thoughtful support and guidance from my advisor, Nick Muller, whose mentorship was indispensable to this work. I am also grateful to my committee members, Peter Adams, Akshaya Jha, and Erin Mansur, for their invaluable feedback, questions, and guidance. I thank Edson Severini, Jonathan Lhost, Jared Cohen, Neil Donahue, Valerie Karplus, Destenie Nock, Paulina Jaramillo, and Mitchell Small for their helpful comments. I also benefited from feedback provided by audience members at the AERE Summer Conference, Camp Resources XXXI, and the AERE@EEA Conference. Finally, I truly appreciate my colleagues in the program for their continued support and encouragement. I gratefully acknowledge the generous funding provided by Heinz Endowments. Any errors are entirely my own. This is a preliminary draft. Please do not circulate or cite.

†PhD Candidate, Engineering & Public Policy, Carnegie Mellon University. Email address: thuynguyen@cmu.edu

1 Introduction

Technological advancements in hydraulic fracturing, coupled with horizontal drilling, have fundamentally shifted the energy landscape by unlocking vast quantities of oil and gas from shale formations that were previously uneconomical to exploit. After extended periods of declining domestic production, the shale revolution made the United States (U.S.) the world's largest oil and gas producer. In March 2025, U.S. crude oil production averaged 13.5 million barrels per day, while natural gas gross withdrawals reached 131 billion cubic feet per day, each reaching historic highs [70, 71].

The impact of the shale boom on the economy and the environment has been profound. The surge in shale oil and gas production alleviated geopolitical concerns over the U.S. dependence on foreign oil imports. The resulting domestic natural gas supply glut fast-tracked the transition away from coal combustion for electricity, contributing to a substantial decline in greenhouse gas emission and air pollution in the U.S. At the household level, inexpensive gas has lowered electricity and heating bills.

Despite these benefits, the proliferation of shale oil and gas has also sparked intense policy debate. Proponents argue that shale development creates jobs and income for the local economy. Critics, on the other hand, emphasize the long-term costs of depleting non-renewable resources and the negative environmental externalities of fracking, advocating for restrictions or outright bans. Looking ahead, rising energy demand from data centers, artificial intelligence, electrification, and the integration of intermittent renewables ensure the continued importance of domestic oil and gas production. To the extent that policymakers try to balance these competing objectives, a complete understanding of the economic and environmental consequences of hydrocarbon extraction is critical to the development of sound oil and gas policies.

A primary challenge in assessing the consequences of oil and gas extraction is capturing the spatial heterogeneity of its impacts. Drilling often takes place in remote areas, with long-distance commuters making up a large portion of the workforce. As a result, non-local workers may capture a significant portion of the value generated, leaving less for the producing region. On the other hand, remote producing regions may also lack the infrastructure to support all aspects of resource extraction. Consequently, oil and gas activity may send local residents to other counties where upstream or support industries, such as equipment suppliers, contractors, and drilling services, are based. In these cases, earnings would flow back to workers' home communities in the producing counties.

What further distinguishes oil and gas from many industrial processes is that it releases emissions from numerous small, dispersed sources spread over a broad area, including drilling rigs, transport vehicles, pipeline compressor stations, and extraction infrastructure. Individual wells are themselves small in scale, but collectively, their sheer number means that emissions can accumulate to significant levels. In the context where large stationary sources and long-range transported emissions represent a greater share of pollution, identifying the marginal contribution of relatively small, dispersed sources becomes challenging. Moreover, air pollutants can travel long distances, undergo secondary formation, complicating efforts to isolate the contributions of oil and gas extraction from background sources.

This paper studies the effects of extraction on fine particulate matter 2.5 ($PM_{2.5}$), which arises from diesel engines, trucks, and heavy machinery used in drilling and production. These fine particulates can penetrate deep into lungs. Exposure to elevated $PM_{2.5}$ and its precursors is linked to increased mortality and a wide range of morbidity, including respiratory and cardiovascular diseases [53]. However, isolating their impact from oil and gas remains challenging due to the diffuse nature of the emissions, their complex interaction with broader atmospheric processes, and the limited availability of granular data.

Therefore, although the economic impacts of oil and gas production have been extensively studied, there is limited research on air quality consequences, and almost no work integrating labor and air quality outcomes in a spatially explicit framework. This paper addresses this gap by examining the spatial and demographic distribution of the labor and air quality impacts of oil and gas extraction. I provide new evidence of a pronounced geographic and generational disparity: much of the economic benefit from oil and gas accrues to young, non-local workers and, to a lesser extent, mobile local residents, while air pollution-related health damage falls disproportionately on older residents living near production activity.

In the first set of empirical analyses, using county-level annual labor market data from the U.S. Bureau of Economic Analysis and the U.S. Census Bureau, I construct measures tracing the geographical flows of employment and earnings across three groups of workers: (1) local residents working in their home county, (2) local residents commuting to work outside their home county, and (3) non-local workers [16, 62]. This decomposition allows me to weigh local economic gains against the value captured by non-local workers, while also assessing how benefits distribute within the local economy. I also examine job flows across monthly earnings brackets, industry segments, and age groups to identify the drivers of the observed labor market effects.

To capture the full scope of local economic impact, I also incorporate non-labor income, such as royalty payments to local mineral rights holders, and evaluate whether total benefits to local residents are commensurate with the overall economic value generated by oil and gas production. Finally, recognizing that migration is a key labor market adjustment mechanism, I use migration data from the Internal Revenue Service to determine whether extraction attracts new residents to producing areas or prompts outmigration. Together, these analyses provide a thorough assessment of the labor market consequences of oil and gas extraction.

In the second set of empirical analyses, I leverage the recently released Gridded Environmental Impacts

Frame, derived from the U.S. Census Bureau's confidential microdata, to assess the air quality consequences of oil and gas production [11]. The dataset provides high-resolution demographic information, including population counts by three age groups, along with air pollution concentration including $PM_{2.5}$, on a fine 0.01-degree grid. The severity of health impacts of $PM_{2.5}$ exposure depend not only on the magnitude of the increase in $PM_{2.5}$, but also on proximity to emission sources and the baseline health of the exposed population. Exploiting the fine-scaled pollution data, together with data on well location and production, I investigate the $PM_{2.5}$ impacts of oil and gas production in multiple, incrementally increasing distance bands, thereby mapping their spatial diffusion. The granular demographic data then permit precise, population-specific valuation of health damage, accounting for the baseline health of affected individuals over time and across space.

I focus on extraction in Pennsylvania, Ohio, and West Virginia over the period from 2001 to 2020. Prior to the shale boom, extraction activity in this region was relatively limited; however, these Appalachian Basin states have since experienced tremendous expansion of oil and gas production. The study period spans both the pre-shale baseline and the subsequent growth and maturation of shale development. Drilling and extraction decisions are not exogenous; firms may strategically target counties with lower wages or leasing costs. This potentially introduces bias due to reverse causality where increased drilling is a response to more favorable local economic conditions. In addition, time-varying omitted variables such as local policy changes could further bias the observed estimates if they are correlated with oil and gas activity and economic outcomes of interest. For causal estimates, I exploit exogenous geological variation in shale resources to instrument for oil and gas production, following the seminal work of Feyrer et al. (2017) [19]. The panel data structure also permits inclusion of geographic (county or grid cell) and year fixed effects to control for persistent spatial differences and common temporal shocks.

I find that the labor market returns of extraction are uneven, varying systematically by worker age, locations, and cross-county mobility. For local workers employed within their home county, oil and gas generates neither net job growth nor increases in total earnings. Instead, it shifts employment across the monthly earnings distribution, reducing jobs earning below \$1,250/month and increasing those earning above \$3,333/month. These shifts are modest and do not produce robust gains in total annual earnings. Therefore, they likely reflect a moderate local multiplier effect, where increased local spending driven by oil and gas raises earnings for existing workers without adding new local jobs. Consistent with this interpretation, I find no evidence of job creation for local workers within their home counties, neither in total nor across the three labor market segments and three age cohorts. This suggests that to benefit from extraction activity, workers may need to be geographically mobile, for example by commuting or having access to reliable transportation.

Indeed, I find that cross-county workers experience meaningful labor market gains. They include local residents who commute to work outside their home county and non-local workers entering the producing counties. For local commuting workers, a 1% increase in oil and gas extraction leads to a 0.015% increase in employment and a 0.022% increase in earnings, translating to roughly eight additional jobs and $525,000 in earnings for every 100,000 barrels of oil equivalent (BOE) produced, on average. Even so, non-local workers realize much larger economic returns, as a 1% increase in oil and gas extraction raises non-local employment and earnings by 0.025% and 0.057%, respectively, averaging about 12 additional jobs and $1.2 million in earnings for each 100,000 BOE. Importantly, non-local workers also gain on average 0.03% in earnings per job for every 1% increase in oil and gas production, roughly $41 per 100,000 BOE, while local commuting workers see no change in earnings per job. Further decomposition by three age brackets (29 and under, 30 to 55, and over 55) shows that job growth is the most pronounced among workers under 30 for both groups of mobile workers.

Together, the labor market results suggest that cross mobility plays an important role in determining whether workers realize labor market gains from oil and gas extraction. I find the main beneficiaries of extraction are young and mobile workers, who are most able and willing to take on jobs involving long commutes, temporary assignments, or site-specific work. Among these workers, non-local employees gain the most, benefiting from both extensive margin job creation and higher per-job earnings, while local commuting workers primarily gain through additional jobs. Therefore, oil and gas firms appear to mobilize their workforce over long distances to hydrocarbon-rich areas and supplement this workforce with local residents willing to commute.

I also find that royalty income provides an important economic return for local residents, accounting for around 30% of total personal income gains from extraction. Yet, local economic benefits remain smaller than earnings captured by non-local workers, even when evaluated using personal income, the broadest measure of economic returns for residents. Two key insights emerge from this observation. First, income from royalties does not fully offset the earnings outflows from the producing county to non-local workers. Second, the leakage of economic value would be even larger if one also accounted for royalties received by non-residents who hold mineral rights in the production area. Local personal income growth is also not proportional to increases in oil and gas revenue or total gross domestic product, confirming that only a fraction of the economic stimulus remains within the local economy. Lastly, I do not find robust evidence of in-migration or out-migration, suggesting that labor adjusts primarily through commuting and worker mobility rather than permanent population relocation.

Next, I examine the effects of oil and gas production on ambient particulate concentrations. The impacts on air quality are most pronounced near the site of production and diminish with distance. Notably, a 1%

increase in production within 1–2 km of a receptor cell raises its $PM_{2.5}$ levels by 0.036%. On average, this implies that an additional 100,000 BOE produced within 1–2 km of a gridded cell increases its $PM_{2.5}$ concentration by 1.2 $\mu g/m^3$. This effect is substantial, amounting to one-tenth of the U.S. Environmental Protection Agency (EPA)'s 2012 annual $PM_{2.5}$ standard of 12 $\mu g/m^3$ and about 13% of the current 2024 standard of 9 $\mu g/m^3$ [50]. The cumulative impact of production could therefore lead to nontrivial exceedances in affected areas, particularly because multiple wells often operate in close proximity. At the same time, these emissions come from numerous small, dispersed sources that are difficult to monitor and regulate.

Combining age-specific population data at the grid cell level with a standard concentration-response function [11, 36], I estimate that oil and gas production resulted in $62 billion (2020$) in health damage in the tri-state region between 2001 to 2020. Over the same period, oil and gas production yielded $265 billion (2020$) in revenue, implying that adverse health costs of extraction amounted to roughly 24% of the value it generated. Most premature deaths, about 70%, occurred in the final five years of the study period, and around 80% were among seniors over 65.

In total, the paper documents a clear divergence between who captures the economic gains and who suffers the environmental costs of extraction. Young, mobile, primarily non-local workers receive most of the economic gains, at the cost of older residents living near extraction sites, who bear the majority of adverse health effects. The unpriced health externalities offset about one-quarter of the economic value of oil and gas extraction.

This paper contributes to several strands of literature. First, it advances research on the local and regional economic consequences of natural resource extraction by documenting the uneven distribution of gains across workers of different ages, mobility, and geographic location. Most closely related work is Gittings and Roach (2020), which also leverages Census origin-destination data to examine the impact of oil and gas production on local versus non-local workers, but with a focus on workplace employment [20]. They find that new per-capita oil and gas production in a county raises workplace employment but reduces the share of local jobs held by local residents, instead drawing in workers from counties 25–200 miles away. They also find that average monthly earnings distributions shift upward for both groups. I build on their work in a number of ways. To begin, I extend their local/non-local framework by also incorporating local residents who work outside their home county. In doing so, my paper emphasizes that cross-county mobility shapes the distribution of oil and gas benefits, not only between local and non-local workers but also within the producing county. In addition to examining average earnings, my work also considers total annual earnings for each worker group. This distinction is important because monthly earnings can vary without producing meaningful changes in aggregate earnings, particularly when production affects employment duration, seasonal work, or workforce size. Finally, I break down job effects by age groups to document which segments of the workforce are most able to take advantage of natural resource booms, informing both policy and economic impact assessments.

Second, it contributes to the environmental health literature on oil and gas extraction by quantifying the $PM_{2.5}$ impacts of production and the resulting health damage. Prior empirical studies on the health effects of oil and gas largely focus on observed health outcomes, particularly among infants, with relatively fewer studies examining the intermediary mechanisms, such as increased ambient air pollution, through which extraction activity impacts health risks. Among those that investigate consequences of oil and gas on air quality, estimates typically rely on bottom-up source-level emission inventories and/or on air quality models that simulate pollutant dispersion using meteorological and chemical inputs. While informative, these mechanistic approaches rely on the assumption that observed extraction activity and the resulting modeled pollution dispersion fully determine the air quality burden, without accounting for indirect or endogenous changes in related polluting activity or behavioral responses.

This paper adds to the scant literature that estimates the causal effects of oil and gas on $PM_{2.5}$. The handful of studies to date measure pollution exposure based on proximity to oil and gas wells or pollution monitoring stations, both of which may be endogenously placed [22–24, 33, 52]. I address this by implementing a spatially exhaustive, grid-based framework across the full study region, which avoids potential siting bias. Furthermore, previous work either does not translate air quality impacts into health damage or relies on coarse, county-level population data. This paper combines pollution and age-specific population data, both at fine 0.01-degree grid-cell resolution, allowing for precise estimation of health impacts that accounts for baseline health across age groups and locations over time.

Finally, and perhaps most importantly, my paper bridges these two literatures by linking economic benefits and environmental costs in a spatially explicit manner, shedding light on the trade-offs and disparities from hydrocarbon extraction across both geographic location and age groups. To my knowledge, the only other study that considers both economic benefits and environmental costs is Mayfield et al. (2019) [40]. However, their analysis uses simulation models and descriptive regression analyses rather than a causal, micro-identified design. Specifically, they use life-cycle and air quality modeling, emission inventories, and descriptive regressions to estimate aggregate greenhouse gas emissions, air quality impacts, and employment effects throughout the Appalachian shale gas supply chain. This approach provides aggregate accounting and informs broad externality costs and policy design, but does not identify who gains or suffers or the nuances of labor market responses.

In contrast, I leverage exogenous geological variation, combined with detailed labor and pollution data, to causally identify the local and distributional impacts of oil and gas extraction, revealing which groups benefit, such as home or commuting workers, and how burdens and returns vary across populations and space.

Therefore, this work offers a complementary and more granular perspective on trade-offs that aggregate environmental-economic accounting may leave opaque.

The remainder of the paper proceeds as follows. Section 2 provides background on the shale revolution and the study region. Section 3 reviews the relevant literature. Section 4 describes the data construction and presents summary statistics. Section 5 outlines the empirical strategy, including the identification framework and the method for estimating excess premature deaths from extraction and their resulting economic damages. Section 6 presents the main results. Section 7 discusses the findings and concludes.

2 Background

2.1 The Shale Revolution

Technological breakthroughs in extracting shale resources transformed oil and gas production by replacing traditional, conventional extraction methods with advanced drilling and fracturing techniques that made it possible to extract large volumes of hydrocarbons from geological formations[1] once thought unreachable. Prior to the shale revolution, extraction targeted conventional reservoirs, surface-accessible formations where hydrocarbons had naturally migrated from deeper source rocks.[2] These reservoirs, characterized by high permeability and porosity, allowed fluid to flow (thus avoiding the need for well stimulation, i.e., techniques to improve oil and gas flow rate). These early conventional drilling methods required relatively simple technology but often produced wells that depleted quickly and delivered unpredictable production volumes. Exploration at the time evoked the image of a "big game trophy hunt," with operators drilling many dry holes in the hope of finding occasional high-yield ones [44]. Consequently, exploration efforts were geographically scattered, and production remained relatively modest.

The emergence of horizontal drilling and hydraulic fracturing revolutionized the industry by unlocking hydrocarbons trapped in low-permeability, low-porosity shale formations, source rocks once considered uneconomical to develop. Hydraulic fracturing (or fracking) pumps a high-pressure mix of water, sand, and chemicals into the rock to create tiny fractures. The sand keeps these cracks open, making it easier for oil and gas to flow into the well. Horizontal drilling turns the well sideways after reaching the target depth, allowing it to pass through oil and gas-rich rock for thousands of feet, which boosts extraction from a single site. Beyond dramatically increasing output, these techniques also changed the nature of oil and gas exploration itself. Modern unconventional development in shale now more closely resembles a systematic

[1] A formation is a distinct rock layer with consistent composition, texture, and other physical properties that set it apart from adjacent layers.

[2] A source rock is an organic-rich sedimentary layer where oil and gas form.

"manufacturing process" rather than high-risk prospecting [44]. Today, operators have much clearer knowledge of where these resources are and how extensive they are. Initial production rates are high, making revenues more predictable and reducing price volatility [26, 37, 44]. Economically speaking, fracking has both shifted the U.S. oil and gas supply curve outward and flattened it [26, 44].

2.2 Appalachian Basin

As a result of the shale revolution, the tri-state region of Pennsylvania (PA), Ohio (OH), and West Virginia (WV), located within the Appalachian Basin, has become a major center of oil and gas production in the contiguous U.S. Abundant in natural gas resources, the region sits atop the highly productive Marcellus and Utica shales and also intersects with older Devonian formations. Figure A1 in Appendix 7 maps out the major shale plays across the three states. Although the Appalachian Basin was home to the nation's first natural gas well in 1821 and first commercial oil discovery in 1859, hydrocarbon production across PA, OH, and WV remained relatively modest for decades. The shale revolution changed this, transforming the region from a marginal producer to a major contributor to the U.S. natural gas supply. In 2023, the tri-state region accounted for about one-third of U.S. dry natural gas production [43].

The fracking boom advanced at varying rates across the tri-state region, with considerable differences in its onset and adoption. Figure 1 presents annual oil and gas production for each state in Panel A and total oil and gas revenue for all three states combined in Panel B. PA saw a sharp increase in oil and gas output beginning around 2009, followed shortly thereafter by WV. It was not until 2013 that OH began experiencing a similar surge in its oil and gas output. Among the three states, PA, situated directly over the most productive portion of the Marcellus Shale, is by far the largest contributor, accounting for more than 60% of total oil and gas output in the region. Over the full sample period, total oil and gas output in the three states increased 25-fold. Between 2010 and 2020 alone, production grew by a factor of 11. While Panel A captures an overall steady increase in production across all three states, Panel B shows that total revenue fluctuates more markedly, reflecting the volatility of commodity prices. In particular, revenues in the region dropped sharply in 2015-2016 due to the global decline in oil prices. Nonetheless, cumulative oil and gas revenues are staggering, reaching $265 billion in 2020 dollars over the 2001–2020 period.

3 Related Literature

This paper builds on two strands of research: the economic consequences of natural resource extraction and their environmental externalities, particularly air pollution. While both speak to the local impacts of extraction, they are often studied in isolation, making it difficult to evaluate the net social welfare effects and distributional trade-offs. This section synthesizes key findings from both.

3.1 Hydrocarbon Extraction and Local Economies

An extensive literature documents how oil and gas booms impact local economies. Oil and gas activity tends to boost employment and wages, generates royalty payments, and raise public revenues in producing regions. Feyrer, Mansur, and Sacerdote (2017) estimate that every million dollars of new oil and gas production generates $80,000 in wage income and $132,000 in royalty and business income for the county. Focusing on the Marcellus shale region [19], Komarek (2016) finds that employment and wages per job increased by 7% and 11%, respectively, during the three years following the shale boom relative to the pre-boom period [35]. The study also identifies significant positive spillovers to non-traded sectors adjacent to extraction through the multiplier effect—that is, increased direct economic activity due to oil and gas stimulates additional downstream local spending and job creation, particularly in supporting industries such as construction and transportation.

Royalty rents from oil and gas firms to private mineral rights owners are another key channel through which oil and gas production generates local income. Brown, Fitzgerald, and Weber (2016) find that production in the Marcellus shale generated $235 in local royalty income per capita, totaling $2.15 billion in 2014 [7]. Across six major shale plays, total local royalty income reached $8.73 billion.[3] Royalty income also stimulates the local economy beyond the initial payments. Brown, Fitzgerald, and Weber (2019) estimate that each dollar of royalties creates 49 cents of additional local income [8].

Oil and gas extraction also generates revenues for local and state governments through severance taxes, property and sales taxes, lease payments, and in-kind contributions from firms [4, 45, 46]. While it also raises local spending on services such as road repairs, Newell and Raimi (2018a) find that 74% of local governments report net fiscal benefits, despite these added costs and revenue volatility [45].

However, these gains often prove temporary. Komarek (2016) finds that extractive-driven growth can reverse as quickly as it appears, with employment and wage gains eroding after four years. Spillover benefits to other sectors likewise do not persist [35]. Similarly, Allcott and Keniston (2018) document that while oil and gas booms raise population, employment, wages, and revenue productivity, busts lead to symmetric declines [2]. Jacobsen and Parker (2014), examining the oil boom-and-bust cycles of the 1970s and 1980s, find that local employment and income rose during the boom, but per capita income fell and unemployment compensation rose after the bust [31]. Feyrer, Mansur, and Sacerdote (2017) find that about two-thirds

[3] Local royalty income refers only to the royalty payments received by residents of the county where production takes place.

of the initial wage increase endures two years after production; however, royalty and business income are largely short-lived [19].

Over longer horizons, dependence on extraction may crowd out other forms of investment and growth, particularly in traded sectors (e.g., manufacturing), as rising wages and input costs erode competitiveness. This pattern reflects the classic symptoms of Dutch disease, in which resource booms inflate local prices and wages, making it harder for non-resource sectors to thrive. This has prompted researchers to explore the potential for a within-country "natural resource curse," asking whether resource-rich regions experience slower or more volatile economic development compared to their less resource-dependent counterparts. The evidence on the presence of a "resource curse" from oil and gas generally does not indicate a persistent negative effect. Allcott and Keniston (2018) find evidence against the "natural resource curse," showing that manufacturing is not crowded out but instead moves procyclically with local resource booms [2]. Weber (2014) shows that counties did not become more reliant on mining during the 2000s natural gas boom, as each gas-related mining job generated more than one job outside the sector [76]. Likewise, Komarek (2016) does not find any indications that the fracking boom displaced the traded manufacturing sector, providing no support for the Dutch disease hypothesis [35].

Within this literature, this work contributes to understanding the distributional labor market impacts of oil and gas extraction, particularly on the extent to which earnings and employment gains accrue to local versus non-local workers. Researchers have taken a variety of approaches to distinguish who benefits from extractive booms. These include (1) comparing place-of-work and place-of-residence economic measures; (2) estimating local economic effects across concentric distance bands to trace how wage and income propagate geographically; and (3) using origin–destination data to directly track commuting worker flows across regions. To ground this discussion, I summarize three studies that represent each methodological approach: Wrenn, Kelsey, and Jaenicke (2015) [80]; Feyrer, Mansur, and Sacerdote (2017) [20]; and Gittings and Roach (2020) [20].

Wrenn, Kelsey, and Jaenicke (2015) implement the first strategy by demonstrating that place-of-work employment data substantially overstate local job gains compared to place-of-residence measures in PA's Marcellus region [80]. They compare employment impacts estimated using Bureau of Economic analysis (BEA) and Bureau of Labor Statistics (BLS) data, which attribute jobs to the employer's location, with estimates based on PA tax records that capture where workers actually live. Using a difference-in-differences strategy, they find that employment gains from shale development are more than twice as large in the BEA and BLS data than in the tax-based data. Their results provide one of the early causal estimates suggesting that a significant share of shale-related employment accrues to out-of-county workers rather than local residents. Guettabi and James (2020) similarly contrast work-based measures from the BEA with residence-based employment and wage records from the Alaska Department of Labor. Using synthetic control framework, they find that the late-2000s oil boom significantly increased work-based employment but did not raise total residential employment in Alaska [25]. Instead, resident workers shifted from public to private sector jobs. Weinstein, Partridge, and Tsvetkova (2018) also do this but with residential earnings from BEA and workplace earnings from Economic Modeling Specialists International. However, because local residents working in their own county are counted in both place-of-residence and place-of-work data, the two metrics partially overlap [77]. This overlap limits the ability to isolate the portion of jobs and earnings that accrue to local residents compared with non-local workers commuting from outside the area.

Feyrer, Mansur, and Sacerdote (2017) apply the second approach by estimating how the economic impacts of oil and gas production diffuse geographically. They find that the total gains within a 100-mile radius are roughly three times larger than those confined to the producing county [19]. They combine wage and employment data from the BLS with tax return data from the Internal Revenue Service (IRS), which is based on where the worker files their taxes. However, this strategy may fail to capture job creation beyond the measured radius, particularly when workers commute from distant oil and gas hubs. For instance, oil and gas firms located in Texas, another major shale-producing state, may decide to drill in PA and deploy their workers there, or vice versa. These long-distance commutes fall outside the 0 to 200-mile radii considered in the study, and therefore important labor and income flows between distant shale regions may be overlooked.

Gittings and Roach (2020) use origin–destination employment data from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES), which enables them to implement all three approaches: comparing place-of-work and place-of-residence employment, examining spatial diffusion of economic effects by distance, and directly observing aggregated worker flows between home and job locations [20]. Focusing on the Marcellus and Utica shale regions, they find that new oil and gas production increases workplace employment, but not residential employment. As production increases, the share of workplace jobs held by local residents decline, while non-local employment rises, particularly among commuters from 25–50 and 125–200 miles away. They also find a rightward shift in the earnings distribution for both local and non-local workers working in the producer county. The share of locally held jobs paying under $1,250/month decreases, while non-locally held jobs paying $1,250–$3,333/month increase.

This study extends the work of Gittings and Roach (2020), leveraging the rich detail of LODES data on worker residence and employment locations. It introduces three important extensions. First, I expand the standard local/non-local framework by including local residents who work outside their home county, a group not captured in Gittings and Roach (2020). Specifically, I distinguish workers across three mutually exclusive groups: (i) local residents employed within their county of residence, (ii) local residents commuting to jobs outside their county, and (iii) non-residents commuting in. Resource booms could induce spatial labor reallocation and shifts in economic value flows not only across but also within counties and regions, as residents respond to wage differentials and job opportunities both in their home area and neighboring jurisdictions. Omitting these out-commuting local workers risks underestimating local labor supply adjustments. If oil and gas firms bring in external skilled labor, residents may find work in ancillary sectors (e.g., transportation, construction, hospitality) that expand regionally in response to the boom.

Second, I complement the employment analysis with a detailed decomposition of earnings gains to the three worker groups. Although LODES does capture job counts by monthly earnings bins, BEA data permit a more comprehensive estimation of total earnings accruing to each worker group. This is important because changes in average monthly earnings may not correspond to meaningful changes in total earnings. For instance, differences in earnings across worker groups or shifts in workforce composition can offset one another, resulting in relatively stable average earnings. Therefore, total earnings are necessary to capture the full economic impact. Together, by examining both shifts in monthly average and total earnings, my work provides insight into the distribution of earnings as well as the overall magnitude of gains from extraction.

Third, I extend the employment analysis by further disaggregating job flows by three age groups. The age-group decomposition reveals valuable information about the demographic composition of job creation as a result of oil and gas extraction. For instance, it shows whether employment gains among non-local commuting workers are concentrated in younger or older cohorts, and how this compares to the age distribution of local residents working locally or commuting out. Notably, it permits assessment of whether the age groups that benefit economically from oil and gas coincide with those that experience its environmental harms.

3.2 Negative Externalities of Oil and Gas Extraction

Oil and gas impose a vast range of negative externalities. These include environmental harms such as greenhouse gas emissions and air and water pollution; noise and light disturbances; and social disruptions such as increased crime, traffic, and strain on local amenities and services [1, 4, 18]. Air quality degradation is among the most consequential externalities, given their robust association with adverse health outcomes, including increased premature mortality, cardiovascular morbidity, and chronic respiratory dysfunction [68]. Children are particularly vulnerable to air pollution, especially during critical developmental stages such as in-utero, infancy, and early childhood, when exposure can impair lung function, hinder neurodevelopment, or aggravate existing conditions like asthma [74].

Oil and gas extraction involves a number of relatively small and dispersed sources of $PM_{2.5}$ and its precursors, nitrogen oxide (NOX), sulfur dioxide (SO2), and volatile organic compounds (VOCs), that can be significant when aggregated [53].[4] For example, drilling rigs and fracking pumps use diesel engines that emit primary $PM_{2.5}$ and precursors that form secondary $PM_{2.5}$. Diesel trucks are also used to transport fracking fluids, heavy equipment, and materials into and out of well sites. Despite the decline in total health damage from particulates from 2008-2014, Sergi et al. (2020) find that over one-fourth of U.S. counties saw an increase in health damage caused by $PM_{2.5}$ and the authors identify emission growth from oil and gas extraction as one of the primary causes of the increase [55].

A substantial body of research focuses on the direct health consequences of living in close proximity to oil and gas activity. Considering the developmental risks during pregnancy, a large subset of studies focuses on adverse birth outcomes, especially low birth weight and preterm birth, among infants born to mothers living near unconventional drilling operations. Multiple quasi-experimental studies find consistent evidence of negative effects [13, 27–29, 79]. Beyond birth outcomes, researchers have documented increased prevalence of respiratory symptoms, such as asthma exacerbation, nose and throat irritation, fatigue, and headaches, among populations near drilling sites [49, 60]. Other studies report elevated hospitalization rates, increased sexually transmitted infections, and greater healthcare utilization associated with drilling activity [5, 32, 34, 78].

Comparatively fewer studies examine the intermediary mechanisms, among them increased ambient pollution, through which oil and gas activity contributes to health risks. Where such estimates exist, they typically use emission inventories and air quality models [3, 9, 39, 40, 51, 53, 57, 58, 73]. Emission inventories aggregate source-specific data and engineering estimates, often based on emissions factors, to construct bottom-up accounts of pollution loads, while air quality models simulate pollutant dispersion using atmospheric chemistry and meteorological inputs. These deterministic air quality models offer detailed source attribution, but also make assumptions about emission factors, production processes, and atmospheric transport mechanisms.

To date, econometric attempts to causally identify their impacts on emission remain sparse [21, 81]. These regression-based models leverage quasi-experimental variation in oil and gas activity to estimate their net effect on observed pollutant concentrations. While econometric models may be less precise in tracing the physical pathways of pollution, their advantage lies in capturing the full range of behavioral and market responses, such as indirect effects on traffic, industrial activity, and other pollution-generating sectors that may accompany increased production. For instance, increased extraction may influence emissions through ancillary industrial operations, transportation demand, or changes in land use, all of which can be captured by regression frameworks if they co-vary with the treatment. This makes econometric approaches a

[4] See Appendix Table A.1 for detailed point sources of pollution from oil and gas extraction.

valuable complement to physical models, particularly when evaluating total social costs or designing policy interventions.

González et al. (2022) is one of the first studies to use a quasi-experimental design, leveraging daily variation in wind direction and fixed-location EPA air monitors across California (2006–2019) to identify the causal effects of oil and gas preproduction and production activities on pollution levels [21]. Focusing on $PM_{2.5}$ they find that each additional upwind preproduction well within 2 km of a monitor increases daily $PM_{2.5}$ concentrations by 2.35 $\mu g/m^3$, and by 0.97 $\mu g/m^3$ for wells 2–3 km away. For production, an additional 100 BOE within 1 km raises $PM_{2.5}$ by 1.93 $\mu g/m^3$. They also detect elevated concentrations of NO_2, CO, Ozone, and VOCs near both preproduction and production sites.

Zhang et al. (2023) adopt a complementary strategy in PA, using high-resolution satellite-based aerosol optical depth (AOD) data and a difference-in-differences framework comparing areas near drilled wells to those near permitted but undrilled wells [81]. Importantly, they incorporate wind-driven spillover weights based on Gaussian dispersion modeling into their empirical design, allowing them to distinguish between local and downwind effects of fracking activity. They find that average daily $PM_{2.5}$ concentration goes up by 0.062 $\mu g/m^3$ within 3 km of a fracked well and 0.017 $\mu g/m^3$ if they include all permitted wells, even those not yet drilled. They estimate that the observed increase in $PM_{2.5}$ concentration resulted in 20 additional deaths across 40 counties with at least one shale gas well in PA between 2010 and 2017.

I contribute to this literature in two primary ways. First, I implement a spatially exhaustive, grid-based framework for measuring exposure to oil and gas activity. Prior studies define treatment based on proximity to either extraction activity or pollution monitors. Zhang et al. (2023) examine $PM_{2.5}$ changes within circular areas around wells, and González et al. (2022) focus on well activity around monitors; both of which—wells and monitors—have been found to be endogenously sited with respect to demographics, economic conditions, strategic considerations, or regulatory stringency [22–24, 33, 52]. To address this, I assign oil and gas exposure to uniformly sized spatial cells that span the entire study region, regardless of well placement. This approach enables full spatial coverage, avoids gaps in unmonitored or undrilled areas, and allows for consistent estimation of pollution impacts across space. By decoupling exposure measurement from the siting of wells and monitors, this approach makes the results more generalizable and avoids conditioning on potentially endogenous placement of wells or monitors.

Second, I provide a detailed and spatially precise valuation of $PM_{2.5}$-related health damage from extraction by using fine-grained, age-specific population data. To the extent that wells are numerous and clustered, local differences in exposure become more pronounced, requiring high-resolution population data to accurately capture the magnitude and distribution of health impacts. To achieve this, I use the Gridded Environmental Impact Frame (EIF), a novel dataset derived from the Census Bureau's confidential Environmental

Impacts Frame microdata. This dataset combines geospatial information on pollution concentrations with high-resolution population counts, allowing me to account for the baseline health of the exposed population across both time and space.

4 Data Construction

To examine the economic and air quality impacts of oil and gas extraction, I compiled data on oil and gas activity, detailed economic outcomes, and air pollution to build comprehensive panel datasets at the county and grid cell levels from 2001 to 2020. The study region includes PA, OH, and WV.

Economic data. Existing research on the economic outcomes of extraction typically allocates impacts by place of work[5] and/or by place of residence.[6] While useful, these measures conflate distinct channels of labor market impact. Place-of-work statistics capture where jobs are located, but not whether earnings accrue to residents or commuters. Place-of-residence statistics capture where workers live but not where economic activity originates. More importantly, these measures overlap, and thus imperfectly capture the flows of income and employment generated by extraction across different groups of workers.

To unravel these flows of economic value, I draw on two complementary data sources. First, following Gittings and Roach (2020), I leverage the U.S. Census Bureau's LODES, which provides county-level annual employment data in three forms: by place of work, by place of residence, and by their origin–destination, recording both where workers live and where they work [16, 20]. In LODES, a worker's job location is defined by the physical or mailing address that employers report to the Quarterly Census of Employment and Wages (QCEW) or through Multiple Worksite Reports [10]. This administrative address may not always coincide with the site where the worker regularly performs their duties. This distinction between administrative address and actual worksite introduces some measurement error, though data from the Bureau of Labor Statistics indicates that such discrepancies are relatively low: approximately 5.6% of multiunit employers report inconsistently, representing about 4.5% of multiunit employment nationally [10]. Consequently, while LODES offers a highly granular view of employment patterns, there may be minor differences between the reported and actual work sites. The residence location for workers in LODES comes from federal administrative records.

Next, I obtain earnings data from the U.S. BEA's Local Area Personal Income (LAPI), which reports earnings by place of work, by place of residence, and includes a net residence adjustment to account for intercounty commuting [62]. Place-of-work earnings come from the QCEW [63]. To derive residence-based

[5] i.e., where extraction occurs.
[6] i.e., , where extraction workers live.

earnings, the BEA reallocates place-of-work earnings to the counties where income is received. To do so, it uses U.S. Census' Journey-to-Work (JTW) data to estimate county-to-county wage outflow ratios, representing the share of wages earned by workers who live outside their workplace county. Then, summing the earnings outflows from all other counties produces gross earnings inflows for each residence county. By reallocating earnings based on where income is actually received, LAPI mitigates the main discrepancy in LODES between administrative workplace addresses and actual worksites.

The key advantage of these data sources is their ability to capture both inter- and intra-county distributions of jobs and earnings. Furthermore, by jointly examining employment and earnings, I can identify whether impacts occur along the extensive margin (i.e., job creation) or earnings margin (i.e., higher earnings per job). Using these publicly available data, I construct a set of job and earnings measures that allow me to decompose the flows of economic value into the following categories:

1. **Home-county jobs & earnings**: Jobs & earnings held by locals working in their home county.
2. **Job & earnings inflows**: Jobs & earnings of locals working outside of their home county
3. **Job & earnings outflows**: Jobs & earnings of non-locals.

Note that these measures are defined to reflect the movement of economic values,[7] not the physical movement of labor. For example, job outflows in my framework represent jobs filled by non-local workers, that is, economic value that "flow out" of the local economy, even if the job itself is performed within the producing county. In addition, I use "job" and "employment" interchangeably throughout this study; however, it is important to note that LEHD data count jobs, not individuals, meaning a single worker may contribute multiple jobs to the total.[8]

These disaggregated measures provide detailed insight into the distribution of labor market gains from oil and gas extraction, which may be masked in more aggregated views. For example, if oil and gas creates new local jobs for residents in rural extraction counties where employment opportunities are scarce, I might observe an increase in home-county employment alongside a decline in job inflows. This would occur if previously out-commuting residents now find work locally. In such a case, the place-of-residence employment measure (which sums home-county jobs and job inflows) may register a muted or even null employment effect,[9] despite the real welfare gains from reduced commute. Job and earnings effects may also diverge. As Weber (2012) conceptualized, a booming sector may attract out-commuting local workers back home

[7]This differs from LODES, which define inflows and outflows according to the commuting patterns of workers.

[8]While it is possible to restrict the analysis to primary jobs (i.e., the highest-paying job held by a worker), this study focuses on total job counts to reflect the full employment impact of oil and gas activity. This represents another deviation from Gittings and Roach (2020), who examines effects on primary jobs.

[9]An extreme case would be if all new local jobs were filled by former out-commuters, offsetting job inflows entirely.

without necessarily offering them higher wages than they previously earned elsewhere [75]. In such cases, the primary benefit comes from shorter commutes rather than increased earnings. Thus, the deconstructed job and earnings measures provide a more complete picture of the effects of the labor market.

From LAPI, I obtain additional measures of local economic activity. One important channel through which oil and gas extraction benefits local economies is royalty payments derived from mineral rights. To capture this, I use LAPI's combined measure of dividend, interest, and rental income.[10] I also use personal income as a more comprehensive indicator that includes all sources of income. Both measures are residence-based. Next, I examine the effects of oil and gas on personal income relative to oil and gas revenue and gross domestic product (GDP). Oil and gas revenue presents the total production value to oil and gas firms.[11] GDP represents the value of goods and services produced in a county, and thus captures the total value of oil and gas resource harvested from the county, including the direct and indirect impacts of resource extraction on other sectors and counties. Therefore, the variation in personal income/oil and gas revenue and personal income/real GDP is indicative of how much of the economic value of oil and gas extracted from a county accrues to residents within the county. For instance, a negative coefficient on personal income/real GDP would indicate that the extraction of oil and gas has a smaller stimulus effect on local income relative to the total value it generates.

Air pollution data. To examine variation in local air quality, I use $PM_{2.5}$ estimates from the recently released Gridded Environmental Impacts Frame (EIF) from the U.S. Census Bureau, which integrates several high-resolution geospatial datasets on air pollution and population by several demographic groups [11]. The data are provided on a fixed $0.01^\circ \times 0.01^\circ$ grid, corresponding to approximately 1.2 km^2 resolution in North America. In particular, I leverage the annual mean $PM_{2.5}$ concentrations originally provided by the Atmospheric Composition Analysis Group. They use the GEOS-Chem chemical transport model to simulate the geophysical relationship between aerosol optical depth (AOD) and $PM_{2.5}$ and relate satellite-based AOD retrieval to surface $PM_{2.5}$ concentrations [15].

In addition, I use the granular age-specific population counts both as a control in my empirical model and to value health damage resulting from oil and gas extraction. The Gridded EIF provides age-specific population counts in three broad groups: under 18, 19–65, and over 65 [36]. Epidemiological evidence suggests that exposure to particulate matter disproportionately affects infants and adults aged 30 and above. To estimate the 30–65 population at the grid-cell level, I calculate the share of individuals aged 30-65 within the 19-65 group at the county level and apply this proportion to each cell's 19-65 population. I describe this

[10]According to BEA, rental income "consists of the net income from the rental of tenant-occupied housing by persons, the imputed net income from the housing services of owner-occupied housing, and the royalty income of persons from patents, copyrights, and rights to natural resources." [61]

[11]To estimate oil and gas revenue, I use West Texas Intermediate crude oil price and Henry Hub natural gas price data from FRED [65, 66].

process in greater details in Section 5.3.

Oil and gas data. I obtained monthly well-level oil and gas production data from Enverus (formerly DrillingInfo) [17]. I convert natural gas volumes from cubic feet to barrels of oil equivalent (BOE) and construct a unified production measure by summing oil and gas output, then aggregate this measure to the annual level. For my economic analysis, I aggregate production to the county-year level. To assess air quality impacts, I construct a spatial measure of oil and gas extraction by aggregating production within incremental concentric distance bands (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10-15 km, and 15-20 km) from the centroid of each grid cell. This approach allows me to identify how oil and gas activity at varying proximities affects local ambient pollution.

4.1 Summary statistics

After merging, the final dataset comprises 4,200 county-year and four million cell-year observations, covering 210 counties and approximately 237,000 wells in Pennsylvania, Ohio, and West Virginia between 2001 and 2020. Table 1 presents average characteristics at the county and grid-cell levels over the sample period. All reported values are annual averages, either at the county or cell level.

The average county reports $5.86 billion in personal income and generates $6.92 billion in GDP, with substantial variation across the sample. The decomposed labor market measures reveal pronounced and interesting differences between locally employed workers and those commuting across counties. On average, home-county earnings (i.e., earnings of locals who work in their home county) are about twice as large as earnings inflows (i.e., earnings of locals who work elsewhere). Combined, these two components constitute place-of-residence earnings, which account for about 64% of total personal income.[12] Dividends, interest, and rental income contribute an additional 17% on average, implying that transfer payments[13] account for the remaining 19%.

The scale of cross-county work activity is substantial. The total number of inbound and outbound workers is approximately 1.7 times the number of local workers employed within the county. Despite this, their combined annual earnings are comparable to local earnings, implying that commuters, on average, either work fewer hours or receive lower wages. In addition, job and earnings inflows and outflows are roughly equal, suggesting that commuting workers primarily travel to other counties within the broader tri-state region.

[12] Personal income comprises three main components: place-of-residence earnings, property income (dividends, interest, and rental income), and transfer payments.

[13] Personal current transfer receipts are "receipts of persons from government and business for which no current services are performed. Current transfer receipts from government include Social Security benefits, medical benefits, veterans' benefits, and unemployment insurance benefits. Current transfer receipts from business include liability payments for personal injury and corporate gifts to nonprofit institutions." — U.S. Bureau of Economic Analysis [62]

Oil and gas production is highly skewed, with a mean of 3.21 million BOE per county-year and a standard deviation more than five times the mean, indicating significant heterogeneity in extraction intensity. At the grid-cell level, oil and gas intensity varies widely across distance bands, with more distant bands covering larger areas and capturing more production. Average production is approximately 8,000 BOE within 1 km of a cell centroid, rising to over 1.3 million BOE within 15–20 km. Average annual $PM_{2.5}$ concentrations are 10.24 $\mu g/m^3$. Each cell has an average population of about 8,000 people.

To further document heterogeneity across worker groups, figure 2 displays the annual average number of jobs for the three worker groups in the tri-state region, segmented by labor market types in Panel A, age groups in Panel B, and monthly earnings bins in Panel C. The breakdown relies on the industry, age, and earnings groupings available in the LODES origin-destination data. LODES classifies employment into three broad labor market segments: Goods Producing,[14] Trade, Transportation, and Utilities,[15] and All Other Services.[16]

The industry, age, and earnings profiles for the mobile inflow and outflow workers are essentially the identical. The majority of jobs fall under the All Other Services segment, comprising 66% of home-county workers' jobs compared to 58% for commuting workers. The difference arises from a larger proportion of commuters working in Trade, Transportation, and Utilities, at 24%, versus 17% for home-county workers. The share of workers in Goods Producing industries is roughly similar across all groups, at about 17-18%. Panel B shows that the age distribution is comparable across the three worker groups, with home-county workers having a slightly smaller share of workers aged 29 or younger and a larger share of workers aged 55 or older. This is consistent with the notion that older workers may be less willing or able to commute. Workers aged 30 to 54 make up over half of the labor force. Earnings among mobile workers are skewed toward the higher earnings bracket, with 41% in the over $3,333/month bin versus 35% for home-county workers.

[14] The Goods Producing category includes the following NAICS sectors: 11 (Agriculture, Forestry, Fishing, and Hunting); 21 (Mining, Quarrying, and Oil and Gas Extraction); 23 (Construction); and 31–33 (Manufacturing).

[15] The Trade, Transportation, and Utilities category includes the following NAICS sectors: 22 (Utilities); 42 (Wholesale Trade); 44–45 (Retail Trade); and 48–49 (Transportation and Warehousing).

[16] The All Other Services category includes the following NAICS sectors: 51 (Information); 52 (Finance and Insurance); 53 (Real Estate and Rental and Leasing); 54 (Professional, Scientific, and Technical Services); 55 (Management of Companies and Enterprises); 56 (Administrative and Support and Waste Management and Remediation Services); 61 (Educational Services); 62 (Health Care and Social Assistance); 71 (Arts, Entertainment, and Recreation); 72 (Accommodation and Food Services); 81 (Other Services, except Public Administration); and 92 (Public Administration).

5 Empirical Strategy

5.1 Economic Model Specification

To study the effects of oil and gas production on economic outcomes, I estimate the following baseline regression:

$$\text{arcsinh}(Y_{i,t}) = \alpha \text{arcsinh}(\text{Oil \& Gas Production}_{i,t}) + \text{X}_{it} + \lambda_i + \lambda_t + \mu_{i,t} \tag{1}$$

where $Y_{i,t}$ denotes an economic outcome[17] in county i and year t, Oil & Gas Production$_{i,t}$ is oil and gas production measured in barrels of oil equivalent, X$_{i,t}$ is a vector of observed county-specific attributes,[18] λ_i is the county fixed effect controlling for persistent county-specific characteristics, and λ_t are year fixed effects controlling for shocks common to all counties in year t, and $\mu_{i,t}$ is the idiosyncratic error term.

I use the inverse hyperbolic sine (IHS) transformation to approximate a log transformation for both oil and gas and my outcome variables, allowing for zeros in oil and gas production and addressing skewness in both production and economic outcomes. The literature has highlighted potential problems with interpreting treatment effects when using the IHS transformation on outcomes with zeros [12, 41, 47]. However, these issues arise because the outcome itself can be zero, creating a mixture of extensive- and intensive-margin effects that are implicitly weighted by the scale of measurement. However, in this study, all outcome variables are strictly positive.

The IHS transformation functions essentially as a smooth approximation to the natural logarithm, preserving the interpretability of coefficients as approximate elasticities while also addressing skewness and accommodating zeros in explanatory variables, such as oil and gas production. As a result, the coefficient α on the independent variable of interest, Oil & Gas Production$_{i,t}$, approximate elasticities,[19] representing the percentage change in an outcome given a percentage increase in oil and gas production, holding other things constant.

While the underlying geological formations and hydrocarbon resources are plausibly exogenous to local economic conditions, drilling and extraction decisions may not be. To mitigate endogeneity concerns, I follow Feyrer et al. (2017) in exploiting exogenous geological variation in the distribution of shale resources

[17] I examine the impacts of oil and gas extraction on the following county-level economic outcomes: jobs and earnings by place-of-residence and place-of-work, home-county jobs/earnings, job/earnings inflows, job/earnings outflows, personal income, personal income/oil and gas revenue, in- and out-migration.

[18] County controls include population, average temperature, and average precipitation. All controls are IHS-transformed.

[19] Bellemare and Wichman (2019) show that for large values of the outcome and explanatory variables, the IHS transformation approximates the natural logarithm, allowing coefficients to be interpreted as elasticities as those in a log–log model [6]. As a rule of thumb, they suggest that this approximation is reasonable when the value of outcome and independent variables both exceed 10. My economic variables far exceed this arbitrary threshold. As expected, the elasticity estimates are essentially unchanged when derived following Bellemare and Wichman (2019). Thus, my results report the α coefficients from equation (1) without the adjustment, and separately present predicted effects in levels that incorporate the adjustment.

to instrument for oil and gas production [19]. Specifically, the intensity of oil and gas activity is modeled using county fixed effects to account for persistent differences in production levels across counties, and the interaction between shale plays and year dummies,[20] which capture the timing of the rollout of oil and gas production within a given shale play. I estimate the following equation:

$$\text{arcsinh}(\text{Oil \& Gas Production}_{i,t}) = \lambda_i + \lambda_{st} + \epsilon_{i,t} \tag{2}$$

where λ_{st} denotes a vector of shale play-year dummies, and ϵ_{it} is the error term.

This specification weighs the exposure of each county to shale production by the fraction of the county's area that overlaps with the shale formation [18]. The predictions from equation (2) is a valid instrumental variable (IV) because they capture the overall timing of production growth across the entire shale play, rather than relying on oil and gas activity within any individual county, thereby minimizing potential correlation with county-specific unobservables. Standard errors are clustered by county and year to account for serial correlation and contemporaneous shocks within counties over time.

5.2 Air Quality Model Specification

I apply a similar empirical strategy to examine the impact of oil and gas activity on $\text{PM}_{2.5}$, adapting for key differences in the structure of the data and the nature of the outcome. Specifically, the analysis uses pollution measurements at the grid-cell level (rather than counties) and accounts for the fact that pollution can disperse spatially beyond the point of emission. In contrast to economic outcomes, which are typically measured and experienced within administrative boundaries, air pollution is governed by atmospheric processes and disperses continuously across space, requiring a design that is not constrained by jurisdictional lines and captures spatial variation in exposure.

To this end, I define production-based exposure as the total volume of production occurring within concentric distance bands, specifically, 0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10-15 km, and 15-20 km, from the centroid of each pollution grid cell. To estimate how production within different distance bands impacts pollution concentrations in the receptor (focal) cell, I estimate the following baseline specification:

$$\text{arcsinh}(\text{PM}_{j,t}) = \alpha^d \text{arcsinh}(\text{Oil \& Gas Production}^d_{j,c}) + \text{X}_{jt} + \lambda_j + \lambda_t + \nu^d_{i,t} \tag{3}$$

where $\text{PM}_{j,t}$ denotes $\text{PM}_{2.5}$ concentration, measured in $\mu g/m^3$, of cell j in year t, Oil & Gas $\text{Production}^d_{j,c}$ represents oil and gas production within a d-km (e.g., 1-2 km) distance band of cell j in year t, X_{jt} is a vector

[20] Shale plays enter the model as categorical variables indicating which shale formation(s), if any, are present within the county.

of cell-level control variables.[21] λ_j is the cell fixed effect to account for cell time-invariant characteristics.

The IV specification then becomes:

$$\text{arcsinh}(\text{Oil \& Gas Production}^d_{j,t}) = \lambda_j + \lambda^d_{s,t} \times S^d_j + \eta^d_{j,t} \tag{4}$$

$\lambda^d_{s,t}$ denotes a vector of shale play-by-year dummies, defined by all cells within a d-km distance band, S^d_j denotes the proportion of grid cells within a d-km distance band j that lie on at least one shale play.

The IV specification for the air quality analysis in equation (4) parallels that of equation (2), absorbing baseline differences in production levels within cell's distance bands and leveraging exogenous temporal variation in production intensity linked to geological formations, while also accounting for spatial heterogeneity in shale play presence across varying distance band sizes. They differ mainly in the spatial unit of analysis (i.e., county versus cell) and in the construction of oil and gas production measure. Lastly, I cluster standard errors by cell and year.

5.3 Premature Mortality Attributable to Oil and Gas Production

Extensive epidemiological research has demonstrated that exposure to $PM_{2.5}$ elevates the risk of premature mortality. Indeed, increased mortality risk accounts for the majority of health damage linked to air pollution (EPA 2011). However, pollution levels alone do not fully capture the public health burden, as the true impact jointly depends on exposure levels, the size of the exposed population, and the baseline vulnerability of different demographic groups.

To estimate the mortality consequences of changes in $PM_{2.5}$ concentrations due to oil and gas, I apply a concentration–response function (CRF), which relates fine particulate exposure to all-cause mortality risk among adults. Then, by combining estimated pollution changes with high-resolution population counts and county-level mortality rates for two age groups: 30-65 and over 65, I account for the fact that identical pollution increments can result in very different health burdens depending on who and how many are exposed. For each grid cell j, age group a, and year t, the change in the mortality rate is calculated as:

$$\Delta m_{j,a,t} = m_{0,j,a,t} \left(1 - \frac{1}{\exp\left(\beta \cdot \Delta \text{PM}_{2.5,j,t}\right)}\right) \tag{5}$$

where $m_{0,j,a,t}$ is the baseline annual mortality rate for age group a in state s in year t. β is the natural

[21] I control for the IHS-transformed population count and IHS-transformed oil and gas production in surrounding distance bands to account for spatial spillovers and correlations across the source zone. For each focal distance band, I control production within up to four immediately adjacent bands. For example, when estimating the effect of production at 2–3 km, I control for production in the 0–1 km, 1–2 km, 3–5 km, and 5–10 km bands. Results are largely unchanged if I instead include only the two closest surrounding bands. For distance bands at the edges of the distribution, such as 0–1 km or 15–20 km, I include the adjacent bands that exist within the study region.

log of the relative risk[22] corresponding to a 10 $\mu g/m^3$ increase in ambient $PM_{2.5}$ pollution. $\Delta PM_{2.5,j,t}$ is the estimated average change in $PM_{2.5}$ attributable to total oil and gas production in cell c and year t.

I use the dose–response relationship from the 2009 American Cancer Society (ACS) cohort study [36].[23] The study estimates a relative risk estimate of all-cause mortality of 1.06 (95% CI: 1.04–1.08) per 10 $\mu g/m^3$ increase in $PM_{2.5}$ for adults aged 30 and older [36]. A more recent analysis of the American Cancer Society cohort by Turner et al. (2016) report identical risk estimates [59]. The baseline mortality rate information $m_{0,j,a,t}$ come from the Centers for Disease Control and Prevention's (CDC's) WONDER database [42]. I use county-level mortality rates for the 30-65 and 65+ age groups whenever available. For suppressed or low-quality data, particularly in sparsely populated counties, I substitute state-level age-specific mortality rates. Each 0.01-degree grid cell is then assigned the mortality rate of its corresponding county.

The expected change in the number of deaths in each cell, age group, and year is the product of the change in mortality rate and the corresponding population:

$$\Delta \text{Mortality}_{j,a,t} = \Delta m_{j,a,t} \times \text{Population}_{j,a,t} \tag{6}$$

My population data from the Gridded EIF, which provides population counts at the 0.01-grid cell for three age groups: under 19, 19-65, and over 65. The population at risk of premature mortality from chronic $PM_{2.5}$ include individuals 30 and above [36]. I estimate population of the 30-65 at the cell level it in two steps. First, I calculate the proportion of the 19-65 population that falls within the 30-65 range using county-level age-specific population estimates from the CDC. Then, I apply this county-level proportion to the corresponding 19-65 population counts in each grid cell from the Gridded EIF. This approach allows me to estimate cell-level population counts for the 30-65 age group while maintaining the fine spatial granularity of the original data.

Next, I aggregate these mortality changes across all grid cells, age groups, and years to obtain the total mortality burden:

$$\Delta \text{Mortality}_{\text{Total}} = \sum_j \sum_d \sum_t \Delta \text{Mortality}_{j,a,t} \tag{7}$$

To translate the health impact into monetary social costs, I use the Value of a Statistical Life (VSL), which quantifies the monetary value associated with small changes in mortality risk across a population. Rather than assigning a value to any individual life, the VSL reflects the aggregate willingness to pay for marginal

[22]Relative risk is defined as the ratio of outcome probabilities between exposed and unexposed populations.

[23]This functional form assumes the CRF is log-linear in relative risk (RR). Specifically, RR $= \exp(\beta \cdot \Delta PM_{2.5})$. The attributable fraction is therefore $1 - \frac{1}{\text{RR}}$, reflecting the proportion of baseline mortality attributable to the pollutant increment.

reductions in the probability of death. For example, if 100,000 people each pay a small amount to reduce their individual risk by 1 in 100,000, the total willingness to pay equals the VSL [72]. The valuation applies a VSL of \$7.4 million (\$2006), consistent with EPA guidelines [72]. The monetized damage from elevated premature mortality due to oil and gas is given by:

$$\text{Total Damage} = \Delta\text{Mortality}_{\text{Total}} \times \text{VSL} \tag{8}$$

6 Results

This section presents the results in three parts. Section 6.1 examines the labor market consequences of oil and gas extraction. I begin in Section 6.1.1 by examining broad measures of jobs and earnings by place of work and residence. In Section 6.1.2, I decompose these measures to trace how the economic value of oil and gas flows through the labor market. Section 6.1.3 explores heterogeneous employment responses across industry segments and age groups. I then turn to royalty and personal income in Section 6.1.4, assessing whether local income growth keeps pace with value creation from oil and gas. Finally, I study migration responses as a key channel of labor market adjustment.

Section 6.2 analyzes the impact of air quality. Section 6.2.1 estimates $PM_{2.5}$ effects of production across increasing distance bands, and Section 6.2.2 uses these estimates to value the health damage due to elevated $PM_{2.5}$ exposure. Section 6.2.3 maps the damage over space and time, demonstrating how the geographic shift of oil and gas operations has altered the distribution of impacts. Section 6.2.4 presents the results under alternative dose–response relationships.

Section 6.3 investigates dynamics, asking how the relationship between oil and gas extraction and labor market and air quality outcomes has evolved over time, with particular attention to comparing the pre- and post-shale eras.

6.1 Labor Market Impacts

6.1.1 Place-of-work versus Place-of residence

The first set of results reported in Table 2 compares the impact of oil and gas extraction on economic measures based on place of residence with those based on place of work. The first two columns of Table 2 focus on job effects, while the last two present earnings effects. The exercise aims to replicate what has commonly been done in the literature, which often relies on one type of measure or, in some cases, compares the two. The coefficients approximate elasticities, representing the percentage change in an outcome given a percentage increase in oil and gas production, holding other things constant. To provide a more intuitive understanding of the results, I also report the predicted level effects per 100,000 BOE alongside the coefficient estimates.[24]

Consistent with the literature, I find larger effects on work-based measures than residence-based measures, indicating leakages of economic value out of producing counties [20, 80]. I find that the employment measured by place of work increases by 0.010%, which corresponds to an average of about 10 additional jobs for every 100,000 BOE. This result aligns with the findings of Gittings and Roach (2020) [20]. In contrast, the effect on resident employment is two-thirds as large and is not statistically significant. Although resident employment did not increase, their earnings grow by 0.013%, which translates to $973,000 for the average county for each 100,000 BOE increase in oil and gas production. However, work-based earnings effects are about twice as large. Specifically, earnings by place of work increase by 0.024%, or about $1.92 million for an additional 100,000 BOE.

The exercise is informative but comes with caveats. First, place-of-work and place-of-residence measures overlap since both of them include home-county workers. The overlap clouds the interpretation of how oil and gas production affects the local labor market, as it combines the effects on different and potentially heterogeneous groups of workers. Furthermore, while the results suggest that non-local workers capture a significant share of the economic benefits of oil and gas extraction, Table 2 does not allow for a direct estimation of the outflows of economic value from the producing area. Additionally, the place-of-residence measure, which combines labor market outcomes for home-county and inflow workers, may obscure important within-county variations in the distribution of economic benefits.

6.1.2 Decomposing Flows of Job and Earnings

To address the limitations, I decompose the place-of-work and place-of-residence measures to study the labor market effects in greater detail in Table 3. Table 3 examines both job and earnings outcomes across three worker groups, with the first three columns focusing on jobs and the latter three on earnings. Home-county measures reflect the job and earnings of residents who work within their home county. Inflow variables capture the jobs and earnings of residents who commute outside their home county for work, while outflow variables represent the economic outcomes of out-of-county workers coming into the producing area.

[24] Formally, elasticity is defined as $\varepsilon_{yx} = \frac{\partial y}{\partial x} \cdot \frac{x}{y}$. Rearranging gives the marginal effect $\frac{\partial y}{\partial x} = \varepsilon_{yx} \cdot \frac{y}{x}$. Assuming local constancy, the marginal effect at the means is approximated by $\frac{\partial y}{\partial x} \approx \hat{\varepsilon}_{yx} \cdot \frac{\bar{y}}{\bar{x}}$. For a finite change Δx, the implied change in y is $\Delta y \approx \Delta x \cdot \hat{\varepsilon}_{yx} \cdot \frac{\bar{y}}{\bar{x}}$. All predicted level effects are scaled to changes of 100,000 BOE. Note that I apply Bellemare and Wichman (2019)'s approach to derive elasticities from arcsinh-arsinh regressions, where $\hat{\varepsilon}_{yx} = \hat{\alpha} \frac{\sqrt{\bar{y}^2+1}}{\sqrt{\bar{y}}} \cdot \frac{\bar{x}}{\sqrt{\bar{x}^2+1}}$ [6]. In addition, $\hat{\alpha}$ represents the estimated coefficients from equations (1) and (3) [6]. Finally, to account for the fact that oil and gas production varies substantially and is often concentrated in rural regions, instead of using the simple mean of y, I use a weighted average of y where the weights correspond to oil and gas production. This weighting ensures that observations with higher oil and gas production exert proportionally greater influence on the average outcome.

This framework offers a thorough understanding of how oil and gas production affects both inter-county (comparing local and non-local workers) and intra-county (comparing home-county and inflow workers) economic flows.

Columns 1-3 estimate the effect of oil and gas production on employment. I find a null effect of oil and gas production on home-county jobs (column 1), suggesting that oil and gas production does not create jobs for locals within the producing county. This implies that oil and gas production generates little to no induced labor demand within the local area. Even so, there is some evidence that oil and gas increase employment for out-commuting local workers. Specifically, a 1% increase in oil and gas extraction leads to a 0.015% increase in job inflows or about eight jobs for each 100,000 BOE (column 2). However, the employment elasticity for non-local outflow workers is over 1.65 times larger, with a 1% increase in oil and gas production raising outflow jobs by 0.025% (column 3). For the average county, this increase represents about 12 additional jobs going to non-locals. Overall, the results indicate that oil and gas production exclusively creates jobs for mobile commuting workers, with the majority of employment gains accruing to non-local workers.

Next, I turn to earnings to examine whether extraction primarily drives job creation or also leads to gains in earnings through higher wages or increased hours worked. Columns 4 through 6 report the estimated effects on total earnings. Consistent with employment results, home-county earnings are not significant, but earnings inflows are. A 1% increase in oil and gas raises earnings inflows by 0.022%, which is about $525,000 in aggregate, on average. However, earning outflows to non-locals are more than twice as large. Specifically, a 1% increase in extraction raises earning outflows by 0.057%. On average, this translates to about $1.2 million for every 100,000 BOE.

Aggregate earnings are the sum product of job count, hours worked, and the prevailing wage rates that correspond to those jobs. To identify the channels of the earnings gains reported in Table 3, I examine how oil and gas impacts the distribution of jobs across three earnings brackets in Table 4, namely those earning $1,250/month or less, $1,251 to $3,330/month, and over $3,330/month, and earnings per job in Table 5. This allows me to say something about whether observed gains arise from extensive-margin changes (i.e., job creation) or earnings margins effect (i.e., higher wages).

Table 4 reveals interesting shifts in the distribution of jobs by earnings that are not apparent in Table 3. While Table 3 shows no significant effect on total employment or aggregate earnings for home-county workers, Table 4 documents a decline in low-earning jobs and a corresponding increase in jobs paying above $3,333/month, signaling an upward shift in earnings. The negative point estimate in column 2 also indicates a decline in jobs earning $1,251-$3,333/month, albeit without conventional statistical significance. Consistent with this finding, Table 5 shows that home-county workers experience an overall increase in earnings per job of 0.0137%, which corresponds to an average increase of $36.5.

Several mechanisms could be at play. One interpretation is that oil and gas activity modestly increased local labor demand for home-county workers, enough to raise wages or hours worked for existing employees but insufficient to generate net job creation. Such effects likely stem from the local economic multiplier effect, where oil and gas-related spending increases demand for local housing, services, and other sectors. An increase in labor market tightness in this setting would also increase the bargaining power of workers, pushing wages higher without expanding employment. Another possibility is that oil and gas enables workers to relinquish lower-wage jobs for higher-paying ones. However, as shown later in Table 8, there is no significant employment effect for home-county workers across three labor market segments, suggesting that any job transitions, if occurring, do not involve movement between these segments.

Local out-commuting workers experience different effects. Unlike home-county workers, who seem to transition from lower-paying to higher-paying jobs without net job or earnings growth, inflow workers primarily see an increase in jobs earning over $3,333/month. However, despite the increase in higher paying jobs, their annual earnings per job do not increase significantly (Table 5), implying that these workers may be working fewer months per year, on average. It remains unclear whether this represents voluntary choices or involuntary labor market constraints. It could reflect increased leisure time or greater flexibility in work schedules. Alternatively, it might capture seasonal or cyclical reductions in employment tied to fluctuations in oil and gas production, which, while generally more stable than drilling, can still experience downturns due to market conditions. Nonetheless, the decline in months worked may indicate a welfare gain for local inflow workers, as they maintain stable annual earnings per job while working fewer months. Since earnings per job do not increase significantly, the total earnings growth for inflow workers shown in column 6 of Table 3 results mainly from an increase in the number of jobs (i.e., the extensive margin).

On the other hand, non-local workers continue to experience the largest labor market returns. They record the greatest gains in higher-paying jobs (column 9 of Table 4) and in earnings per job (column 3 of Table 5). Similar to inflow workers, outflow workers also see more jobs with earnings above $3,333/month. On average, non-local workers gain an additional $41 in earnings per job for every 100,000 BOE of oil and gas produced.

The detailed decomposition reveals two important insights. First, only workers who are both willing and able to commute outside their home county see net job creation from oil and gas production. This signals cross-county labor mobility as a central factor in shaping the distribution of labor market gains in resource-rich regions, as workers must overcome commute-related constraints (e.g., transportation costs, travel time, and availability of transit options) to access these employment opportunities.

Second, labor market benefits manifest differently in magnitude and margin across the three worker groups. For home-county workers, employment shifts from low-paying jobs to higher paying jobs, likely driven by increased demand for local services from oil and gas. This suggests some movement into better-paying positions without net increases in total jobs or earnings. Inflow workers enjoy more higher paying jobs that raise total earnings, although earnings per job remain unchanged, possibly due to shorter working periods of these jobs. Non-local commuters experience the largest gains, as new higher paying jobs raise both employment and earnings per job, resulting in significant growth in total earnings.

6.1.3 Heterogeneous Effects by Industry and Age

To gain insights into the potential mechanisms through which oil and gas influence the labor market, this section further disaggregates employment by three broad labor market segments and three age groups. Although the sectoral breakdown is relatively coarse and thus cannot precisely isolate direct oil and gas employment or its spillovers, it still offers useful insight into which segments of the labor market expand or contract in response to oil and gas production and whether workers shift between labor markets. Meanwhile, age-based breakdown examines whether younger or older workers disproportionately benefit, potentially capturing differences in mobility, skills, or opportunity costs. Together, these breakdowns help frame understanding of how resource-driven demand shocks translate into labor market outcomes across broad sectors and demographics.

The three labor market segments differ in their exposure to oil and gas-driven demand shocks and their capacity to absorb mobile labor. The Goods Producing category, which includes Mining, Quarrying, and Oil and Gas Extraction (NAICS 21), is the closest to capturing direct effects of oil and gas. However, because it also includes Construction, Manufacturing, and Agriculture, the estimated effect may misrepresent the true direct job impact of oil and gas extraction. Oil and gas operations also depend heavily on the Trade, Transportation, and Utilities segment for trucking, pipeline transport, and power supply at well sites. Finally, the All Other Services segment has limited direct involvement in oil and gas extraction and therefore likely captures indirect spillovers. The age bins include three cohorts: young adults (29 or under), middle-aged workers (30–54), and older workers (over 55).

Starting with home-county workers, Table 6 shows no significant effects across labor market or age groups, consistent with the null results for total jobs reported in Table 3. At the same time, Table 4 indicates a shift from low-paying to better paying jobs among home-county workers. Together, these findings suggest that the transition occurs within the same labor market segments and age groups, reinforcing the interpretation that workers are receiving higher wages as incumbents rather than through new job creation. Consequently, there are no net increases in employment.

Table 8 presents the employment impacts across labor market segments for commuting workers, including both inflow and outflow groups. Unlike home-county workers, mobile workers experience job growth primarily in the Goods Producing sector, followed by Trade, Transportation, and Utilities. The results suggest that firms may bring in skilled non-local workers to fill gaps in the local labor market, supplemented to a lesser extent by mobile local workers.[25] At the same time, the job growth among out-commuting local workers may also result from producing counties lacking the industrial base needed to support oil and gas activity locally. Increased oil and gas output raises demand for critical inputs to the extraction process, such as steel, cement, and specialized equipment. In the absence of local inputs, oil and gas operators must source them elsewhere.[26] This demand may prompt input producers to expand operations and scale up production, thereby creating additional employment opportunities in areas better equipped with industrial infrastructure, such as machine shops, fabrication yards, or equipment suppliers.

Similarly, oil and gas production increases the need for transportation and logistics services, including trucking and freight. These activities may be based in commercial or logistical hubs with better highway access and warehouse capacity, rather than in the producing counties themselves. As a result, even local residents may find jobs tied to oil and gas activity outside their home county, contributing to the observed pattern of larger job outflows than inflows in these segments.

Interestingly, in the All Other Services segment, job gains are statistically significant, albeit weak, only for non-local workers. One possible explanation is that in addition to extraction workers and truckers, oil and gas firms may also bring in non-local specialists, such as engineers, environmental consultants, geologists, regulatory experts, and royalty managers, to provide technical and administrative support essential to oil and gas operations.

Moreover, job outflows consistently exceed inflows across all labor market segments. In Goods Producing, outflows are nearly double inflows in both elasticity and predicted effect; in Trade, Transportation, and Utilities, outflow elasticity is about 60% larger. Taken together, these results confirm that oil and gas mostly benefits non-local workers and generates limited induced jobs from increased local spending. Spillover effects mainly occur through increased job opportunities outside of the producing county for mobile local workers in sectors closely linked to oil and gas.

Table 7 presents the job inflow and outflow effects across three age brackets. For inflow workers, job gains go to workers under 30, with smaller but detectable effects for those 55 and older. In contrast, job gains for non-local workers span all three age bins, suggesting that oil and gas demands a mix of mobile young labor for physically demanding fieldwork as well as older, more experienced workers for technical, supervisory, or

[25] Oil and gas industry regularly employs highly skilled workers. According to Bureau of Labor Statistics, the top five occupations in oil and gas extraction industry (NAICS 211100) in 2020 were engineers (10.62%), extraction workers (10.33%), computer occupations (7.53%), financial specialists (7.45%), and material moving workers (7.44%) [64].

[26] For example, Minnesota and Wisconsin supply much of the sand used in fracking despite not sitting on top of shale formations [48].

regulatory roles that support oil and gas extraction and operations. Among non-local workers, the youngest group shows the most elastic employment response, while individuals aged 30 to 54 see the largest predicted effect in levels, reflecting their larger share of the workforce. In both the inflow and outflow groups, young adults respond the most strongly to oil and gas-driven job growth. Since young adults are more likely to be single, rent rather than own their homes, and have fewer family obligations, they may be more willing to commute or relocate for work and better able to respond to the spatially dispersed and transient labor demands of the resource sector.

6.1.4 Local Capture and Leakage of Economic Benefits

Until now, my analysis has focused on labor market outcomes, without considering other sources of income such as royalty payments. To document the full range of economic impacts from oil and gas extraction, I examine additional measures of value that capture different channels of economic returns. One key channel for local benefits from oil and gas is royalty payments from mineral rights. I use LAPI's combined measure of dividends, interest, and rental income to examine asset-based earnings, and personal income as the most comprehensive measure of local economic benefits, since it includes all potential sources of income. Both measures are residence-based.

Table 9 reports the results for dividends, interest, and rental income in column 1 and overall personal income in column 2. The results confirm that royalties represent an important source of local income for extraction counties in the tri-state region. A 1% increase in oil and gas raises dividends, interest, and rental income by 0.016%, or roughly $316,000 per 100,000 BOE for the average county. The comprehensive personal income measure increases by 0.0088% or $1.05 million for every 100,000 BOE, on average.

Assuming the observed effect in column 1 is entirely attributable to royalty income, I estimate that royalties account for roughly 30% of total local economic benefits, as measured by personal income, from extraction, while earnings inflows represent another 50%. Notably, the predicted personal income effect ($1.05 million) for residents is smaller than the predicted earnings outflows effect ($1.19 million). This implies that 1) income gains from royalty do not entirely offset earnings captured by non-locals and 2) value outflows would be even more substantial if I were to consider the leakage of royalties by nonresidents who possess mineral rights within the production area.

In 2014, local mineral ownership of oil and gas in the Marcellus shale was 55% in 2014 [7]. To estimate the potential scale of value leaving the county, I perform a simple back-of-the-envelope calculation assuming the same local ownership as in 2014 and uniform royalty rates for residents and nonresidents. Under this assumption, I estimate that the combined leakages of earnings and royalties amount to roughly $1.45 million for every 100,000 BOE[27] for the average county.

To gauge the extent to which total economic benefits remain in extraction counties versus flowing outward, accounting for both labor and non-labor income, I compare personal income growth[28] to two broader measures of economic value: oil and gas revenue and total county GDP. Oil and gas revenue captures the direct market value of extracted resources, while GDP represents the total value of goods and services produced within the county. These measures encapsulate not only local gains but also leakages of economic value to non-local firms, workers, or landowners. In contrast, personal income is strictly residence-based and serves as a comprehensive measure of local economic stimulus, capturing labor earnings, dividends, and royalty payments received by county residents.

Table 10 presents the effects of oil and gas production on the ratios of personal income to oil and gas revenue and total GDP. A negative coefficient on the personal income-to-revenue or -GDP ratio would imply that the rise in personal income observed in Table 9 does not keep pace with the value of oil and gas exaction to firms or the county GDP. Indeed, column 1 shows that as oil and gas increases, the share of revenue that translates into personal income for local residents declines sharply, indicating that much of the direct resource wealth flows to nonlocal firms, workers, or landowners. Even when moving to broader GDP measures, which incorporate a wider range of economic activities including sectors less prone to economic leakage such as healthcare, the negative relationship persists, although with diminished magnitude (columns 2). This persistence across the two measures suggests that resource extraction reduces the share of economic benefits accruing to local residents. In other words, these findings demonstrate that resource booms do not proportionally translate into local economic gains.

6.1.5 Tracking Flows of People

While job and income measures offer valuable insight into labor market effects, they tend to capture short-term employment changes rather than long-term shifts. If workers are not merely commuting but permanently relocating, it indicates persistent economic changes that extend beyond transient labor adjustments. For instance, if extraction attracts workers to settle in the producing area, then Table 3 overlooks an important aspect of value creation. Unlike transient job inflows, in-migration signifies a more durable expansion of the local workforce. For the extraction county, these are not simply jobs being allocated across space. The new arrivals do more than fill jobs. They establish households, spend locally, and contribute to a broader tax

[27] Calculation: Predicted local royalty income effects, proxied by dividends, interest, and rent, average $316,000 per 100,000 BOE. The non-local portion is $(316\text{K}/55\% \times 45\%)\times \approx \$259\text{K}$. Adding earnings outflows of $1.19M yields total value outflows of about $1.45 million per 100,000 BOE.

[28] According to BEA, personal income as measured by BEA includes rental income, which includes royalty payments, along with earnings from wages, self-employment, and other sources. Rental income "consists of the net income from the rental of tenant-occupied housing by persons, the imputed net income from the housing services of owner-occupied housing, and the royalty income of persons from patents, copyrights, and rights to natural resources" [61].

base. Conversely, if extraction leads to out-migration, whether due to housing pressures, environmental concerns, or economic displacement, it could indicate that local job creation is not sufficient to retain workers or that other negative spillovers outweigh employment gains. Without also examining migration, employment measures alone can understate or mischaracterize the broader economic impact of oil and gas.

Table 11 shows how oil and gas affects in-migration in column 1 and out-migration in column 2, using IRS population migration data based on annual address changes from individual tax returns [56]. The results show no statistically significant effects, indicating that oil and gas extraction does not induce meaningful changes in population flows. Instead, the labor market effects of oil and gas appear to be transitory, driven by the mobility of workers rather than long-term settlement. The absence of population growth despite an influx of non-local workers in extraction counties is consistent with the limited evidence of induced employment spillovers. Table 11 indicates that existing residents also do not relocate in response to increased oil and gas activity, suggesting that the inflow of jobs generated by oil and gas may also be largely short-term or that residents face barriers that prevent them from moving.

6.2 Air Quality Impacts

6.2.1 Estimating Spatial Pollution Gradients

Having thoroughly examined the economic consequences of oil and gas extraction, I now shift focus to its impact on local air quality and how pollution disperses across space. To do this, I examine $PM_{2.5}$ concentrations at the grid-cell level and estimate how oil and gas production within increasing distance bands surrounding a given receptor cell affects its observed pollution level. This approach captures both the localized impacts from nearby oil and gas activity and the broader dispersion and accumulation of pollutants as they travel over longer distances.

Figure 3 presents the estimated air quality. Panel A displays the estimated pollution elasticities alongside 95% confidence intervals and Panel B presents the corresponding predicted changes in $PM_{2.5}$ concentrations per 100,000 BOE along with the range of effects implied by the regression results. Note that panel (B) only displays the effects for statistically significant estimates at the 5% level. Table 12 reports the estimates underlying Figure 3.

Figure 3 reveals a non-monotonic spatial pattern in pollution response. Pollution concentrations respond most strongly to nearby production, but the effects are not statistically significant for production occurring immediately adjacent, within 1 km. This pattern may reflect the atmospheric dispersion and chemical formation processes that govern $PM_{2.5}$ concentrations. At very short distances, pollutant plumes may rise and disperse before settling at ground level, resulting in lower measured concentrations near the emission source

[30, 54]. Turbulent air near the source can also lift emissions upward, delaying their accumulation at ground level [30, 54]. Moreover, $PM_{2.5}$ consists not only of directly emitted particles (primary $PM_{2.5}$), but also of secondary particles formed through atmospheric reactions involving precursor gases such as NOx, SO_2, and VOCs. These chemical transformations require both time and distance to occur, which may explain why pollution effects are more detectable at short-to-moderate distances rather than immediately adjacent to production. This spatial lag reflects well-established atmospheric processes and underscores the importance of examining the full dispersion pattern of pollution.

Figure 3 shows a spatial gradient in the estimated effects. Pollution impacts diminish as production originates farther from the receptor cell. A 1% increase in production within 1–2 km raises $PM_{2.5}$ concentrations by 0.036%, amounting to approximately 1.19 $\mu g/m^3$ per 100,000 BOE. At 2–3 km, the estimated PM2.5 elasticity slightly increases to 0.039%, but the predicted effect decreases to 0.77 $\mu g/m^3$ per 100,000 BOE. This occurs because overall production at 2–3 km is higher, so a 1% increase in production corresponds to a larger absolute change in volume. The magnitude and significance of these effects decrease with distance. For production occurring beyond 3 km, the predicted effects per 100,000 BOE ranges from 0 to 0.16 $\mu g/m^3$ and is statistically insignificant at conventional thresholds.

Taken together, my economic and air quality findings thus far point to a clear disparity: employment and earnings gains largely benefit non-local workers, while the health costs are overwhelmingly borne by local residents.

6.2.2 Monetizing Health Impacts of Air Pollution

While changes in pollution concentrations provide a physical measure of local environmental impact, they do not capture broader social costs arise from increased exposure to fine particulates, most notably the elevated risk of premature mortality. In this section, I translate the estimated pollution effects into economic damage using a dose-response function, informed by the epidemiological literature, that links changes in exposure to $PM_{2.5}$ to mortality risk. I then monetize the resulting health impacts using the value of a statistical life (VSL), yielding spatially explicit estimates of the health costs imposed by oil and gas extraction.

Figure 4 presents the estimated annual health damage attributable to oil and gas production in the tri-state region between 2001 and 2020. The valuation is based on the estimated $PM_{2.5}$ impacts from production within the 1–2 km and 2–3 km distance bands, which are both statistically and economically significant. Panel A reports excess premature mortality (left y-axis) and the corresponding monetized damage (right y-axis), disaggregated by two at-risk age groups: 30–65 and over 65. Panels B and C plot the underlying data that support the valuation in Panel A. Specifically, Panel B shows total population over time and Panel

C shows average baseline mortality rates across counties in the study area.

Cumulative damage over this period amounts to roughly \$62 billion, representing 24% of the total \$265 billion in oil and gas revenue (see Figure 1). Health damage sharply increased beginning around 2010, with an almost exponential rise over the following five years. In contrast, population size remained relatively stable (Panel B), while baseline mortality rates generally declined (Panel C), aside from a spike in 2020 due to the COVID-19 pandemic. These patterns suggest that most of the increase in health damage is attributable to the surge in production driven by the shale boom. The exception is 2020, when the damage continued to grow despite little changes in production, driven by elevated mortality due to the pandemic.

The bulk of health damage, about 70%, occurred between 2015 and 2020, with more than half concentrated in the final three years of the sample. Nearly 80% of these damages fell on individuals over age 65, underscoring the disproportionate burden borne by the most vulnerable residents, who face higher baseline mortality risks. This pattern stands in stark contrast to labor market evidence showing that local workers aged 30 and above experienced little to no employment benefits from oil and gas. Accordingly, the allocation of health costs and economic gains is unequal not only between local and non-local populations but also within local communities themselves. Older residents, in particular, bear a disproportionate share of the health impacts while receiving few, if any, of the economic benefits.

6.2.3 The Geography of Harm

While much of the increase in health damages stems from higher oil and gas output, the migration of drilling activity toward peri-urban and urban areas, particularly the Pittsburgh metropolitan area, further compounds the effect. Figure 5 maps the spatial distribution of oil and gas wells by their first year of production for 2005, 2010, 2015, and 2020, illustrating how production has shifted over time and space. Figure 6 then maps the estimated health damages from production for the same years across the region.

Between 2005 and 2010, there was some expansion into adjacent areas, but activity remained concentrated near established sites, and the impacts in areas newly exposed to production were moderate overall. After 2010, damages spread widely into regions that had previously seen little or no production-related activity, and the scale of impacts in these new areas grew substantially. Specifically, extraction activity in PA expanded to both ends of the state: toward the Pittsburgh region in the southwest, a historically industrial area, and northeast into largely rural counties. In OH, new wells expanded into largely rural southeastern counties, but also emerged around the Columbus and Cleveland metropolitan areas. In WV, new wells cluster in the Northern Panhandle, a relatively densely populated area, and around Monongalia County, one of the more populous counties in the state. Of the total \$62 billion in estimated health damage, PA accounts for over half ($40 billion), followed by OH ($14 billion) and WV ($8 billion) (see Figure 7). The high damages in PA result from its extensive gas production combined with the encroachment of drilling activity into the densely populated Pittsburgh area.

6.2.4 Alternative Risk and Damage Estimates

Estimates of health damage from particulate pollution rest on assumptions about the relationship between particulate exposure and mortality. Therefore, it is important to assess how the results change under different credible parameter values. In the main analysis, I adopt the β coefficient from Krewski et al. (2009) as the default parameter linking changes in $PM_{2.5}$ concentrations to excess mortality. This study has been widely used in regulatory assessments and provides a conservative risk estimate. To assess sensitivity, I re-estimate health damage using the higher β coefficient reported in the Harvard Six Cities (H6C) study by Lepeule et al. (2012), which finds a substantially larger mortality response per unit increase in $PM_{2.5}$ exposure and identify the at-risk population as individuals aged 25 and older [38]. Table 13 summarizes the risk estimates (i.e., β coefficient) from the two epidemiological studies evaluating mortality risks associated with long-term exposure to fine particulates.

Figure 8 then presents the annual health damage estimates under each relative risk assumption, with cumulative damages over the full analysis period (2001–2020) annotated. Estimated damages vary considerably depending on the choice of epidemiological parameter: total damage under H6C reach roughly $131 billion, more than double the $62 billion estimated using the ACS study. If the H6C estimate more accurately represents the true mortality risk from $PM_{2.5}$ in the region, the health cost of extraction would reach an astonishing 50% of the total oil and gas revenue. It is important to note that my estimates only include individuals over 30, so they likely represent a lower bound. Damage could be even larger under the H6C parameter if the 25–30 age group identified as at-risk in the H6C study, as well as infants under one year old who also face elevated $PM_{2.5}$-related mortality risk, were included.

6.3 Dynamic Effects of Extraction

In this section, I examine how the effects of oil and gas extraction change over time to understand the persistence of economic and air quality impacts observed. I do this by interacting production with the number of years since 2008, when hydraulic fracturing began in the study region.[29] This approach allows me to track how the relationship between extraction and my outcome variables evolves over time , providing

[29] Bartik et al. (2019) identifies 2008 as the first year of hydraulic fracturing in the Marcellus shale play and 2012 in the Utica shale play [4].

suggestive evidence of whether these effects persist and how fracking may have changed the economic and air quality consequences of extraction.

Figure 9 presents the dynamic economic effects of oil and gas extraction, focusing specifically on earnings and job flow variables. Panel A shows the main effects of production on economic outcomes in the reference year (i.e., 2007, one year before the onset of fracking), while Panel B displays the estimated coefficients of the interaction term with years since the reference year. Together, the panels illustrate both the baseline relationship in the reference year and how the effects evolve over time relative to the baseline.

Consistent with the main findings in Table 3, Panel A of Figure 9 shows that, in the reference year, non-local workers see the largest total job and earnings effects, followed by local out-commuting workers, while home-county workers do not experience meaningful returns from extraction. Across all groups, earnings effects are consistently larger and precisely estimated.

Panel B then shows changes in these effects relative to the baseline levels depicted in Panel A. Starting with earnings, home-county workers see minor deviations in earnings relative to the baseline, experiencing small positive incremental effects that remain relatively flat over time and are no longer statistically different from zero after five years. For cross-county workers, two notable patterns emerge. First, earnings effects were relatively lower in the years preceding the reference year but steadily increased, suggesting a growing reliance on commuting workers leading up to the onset of fracking, especially among non-local workers. Second, the additional earnings effects jump noticeably immediately after the reference year, reaching a peak around five years later. This jump aligns with the exponential rise in oil and gas output, particularly in PA from 2008 to 2012, as shown in Figure 1. The timing suggests that the rapid expansion of extraction generated strong short-term demand for labor, particularly from cross-county workers.

Job effects follow a similar trajectory to earnings but are smaller in both magnitude and statistical significance. They do not rise immediately after the reference year, instead increasing about three to four years later and peaking roughly five years after 2007, around the same time as earnings effects. The exception is home-county workers, who experience a sharp increase two years after the reference year but quickly return to baseline.

Figure 10 presents the dynamic effects of oil and gas extraction on $PM_{2.5}$ across six distance bands, including 0–1 km, 1-2 km, 2–3 km, 3–5 km, 5–10 km, and 10–15 km. Unlike Figure 9, to reduce computational burden, years are grouped in four-year intervals. The baseline period is 2004–2007, and the first and last groupings are shorter: 2001–2003 and 2020, respectively. Panel B shows that, relative to the baseline effects in Panel A for 2004–2007, the $PM_{2.5}$ impacts of extraction in subsequent periods are persistently smaller across distance bands. These results offer suggestive evidence that fracking may have moderately reduced the incremental air quality impacts of extraction. In the early years, drilling was more geographically dispersed as firms searched for productive shale pockets, spreading emissions more widely. As fracking unlocked access to shale resources, production shifted toward concentrated production in resource-rich “sweet spots,” localizing emissions and reducing incremental air quality impacts. Concurrently, improvements in drilling efficiency, the adoption of multi-well pads, and enhanced emission controls may have also reduced $PM_{2.5}$ intensity per unit of output.

It is important to stress that these results should be interpreted as suggestive evidence only. Establishing causal effects of fracking would require a quasi-experimental design, such as a difference-in-differences approach with an appropriate control group. Furthermore, the magnitude of the deviations shown in Panel B is modest relative to the corresponding baseline effects shown in Panel A. Thus, any observed “gains” (i.e., reductions in adverse air quality impacts) appear fairly limited.

7 Discussion and Conclusion

This paper provides new evidence on the economic and environmental consequences of oil and gas extraction, revealing stark asymmetries in who gains and who bears the costs. Three overarching findings arise from this work. First, the labor market returns of local extraction are regional. I find that cross-county labor mobility drives both inter- and intra-county distribution of labor gains from oil and gas production. Specifically, locals working in the producing county see no statistically meaningful gains in employment or earnings, indicating minimal induced labor demand. Instead, the economic gains primarily reflect job creation for mobile workers, especially non-locals but also, to a lesser extent, residents who commute out of their home counties, whose mobility allows them to meet the geographically dispersed labor demands of the resource sector. Indeed, the strongest job effects are observed among younger and prime-age mobile workers, who are more likely to relocate or commute in response to oil and gas-driven labor demand.

Earnings effects mirror the employment results: only mobile local and non-local workers experience gains, but non-locals see earnings increase roughly twice as large. Moreover, the gains for non-local workers arise from both extensive and earnings margins, whereas local commuting workers experience gains primarily along the extensive margin. Yet even as the value of oil and gas production grows, the share of local personal income—including royalty payments—relative to extraction revenue and GDP declines. This leakage indicates that resource booms do not translate into proportional local economic development. Rather, much of the generated wealth flows to external firms, workers, and landowners.

Second, air pollution and health costs remain local. Although employment benefits bypass many local residents, the resulting health damage from elevated $PM_{2.5}$ disproportionately affects nearby communities. The most severe effects occur from production within 1–3 km. Over the study period, these health damages total \$62 billion, amounting to one-quarter of oil and gas revenue. The health burden falls disproportionately on older adults over 65, who face higher baseline mortality risks and constitute 80% of the total health damage.

Finally, these findings together show that while the oil and gas boom generates significant economic activity, the people who benefit most are often different from those who bear the environmental costs. Younger and prime-age non-local workers capture most of the labor market benefits, while senior residents living near extraction sites absorb the bulk of health damage from elevated $PM_{2.5}$. This asymmetry demonstrates that oil and gas production stimulates economic activity without translating proportionally into local welfare gains, leaving producing regions to shoulder a disproportionate share of environmental and public health costs, even as economic value flows out to non-local workers and absentee landowners.

These findings carry several important policy implications. The findings are relevant for local policymakers considering resource extraction as a means to spur local economic activity. The limited returns to the local economy, coupled with the evident leakage of economic value and air quality degradation, should cause local leaders to hesitate before implementing policies that facilitate hydrocarbon extraction and shale development. This is particularly important because existing federal oversight of oil and gas is limited [48]. Thus, the responsibility for shaping and implementing oil and gas policies lies with the state and local governments. In the past, states have attempted to exert control over local actions with the objective of maximizing oil and gas production [48]. This work provides evidence of important distributional effects at the county level. The trade-offs faced by producing counties may not translate directly to the state level, meaning that local and state economic interests may not always align. Moreover, cross-border economic activity and pollution transport imply that oil and gas-related policies may be relevant even for localities and states that neither produce oil and gas nor allow its extraction.

Tables

Table 1: Summary Statistics

Variable	Mean	Std. Dev.
County-level measures		
Outcome variables		
Personal Income ($ billion)	5.86	11.5
Real Gross Domestic Product ($ billion)	6.92	15.5
Local earnings ($ billion)	2.53	5.86
Earnings inflows ($ billion)	1.22	2.22
Earnings outflows ($ billion)	1.18	3.09
Local jobs (thousand)	29.5	65.9
Job inflows (thousand)	24.2	49
Job outflows (thousand)	24.7	34.9
Dividends, interest, and rental income ($ billion)	0.98	2.18
Independent and control variables		
Oil and Gas (million BOE)	3.21	16.5
Population (thousand)	124	214
Temperature (°F)	51.7	2.47
Precipitation (in.)	3.75	0.654
Cell-level measures		
Outcome variable		
$PM_{2.5}$ ($\mu g/m^3$)	10.24	2.7
Independent and control variables		
Oil and Gas by Source Zones (thousand BOE)		
0-1 km	7.54	107.4
1-2 km	23.2	224
2-3 km	39	327
3-5 km	125	844
5-10 km	580	3,274
10-15 km	962	4,828
15-20 km	1,330	6,092
Population (thousand)	8.06	20.8

Notes: The table shows annual averages; dollar values are inflation-adjusted to 2020 dollars.

Table 2: Impact of Oil and Gas on Job and Earnings by Place of Residence and Place of Work

	Job by place of		Earnings by place of	
	Residence	Work	Residence	Work
	(1)	(2)	(3)	(4)
Oil and Gas Production	0.0066	0.0099**	0.0133**	0.0243***
	(0.0048)	(0.0046)	(0.0048)	(0.0063)
Pred. Lvl Effect / 100,000 BOE	7.3	10.2	973K	1.92M
Observations	3,990	3,990	4,200	4,200
R^2	0.99523	0.99749	0.99747	0.99615
Within R^2	0.06450	0.20468	0.21023	0.14638

Notes: The dependent variables are the inverse hyperbolic sine of county-level total jobs and earnings, measured by place of residence and place of work. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. The earnings data cover 2001–2020 for 210 counties, resulting in 4,200 county-year observations. The job data cover 2002–2020, yielding 3,990 county-year observations. Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Impact of Oil and Gas on Decomposed Job and Earnings to Three Worker Groups

	Job outcomes			Earnings outcomes		
	Home-county	Inflows	Outflows	Home-county	Inflows	Outflows
	(1)	(2)	(3)	(4)	(5)	(6)
Oil and Gas Production	-0.0026	0.0153*	0.0253***	0.0109	0.0220**	0.0568***
	(0.0045)	(0.0077)	(0.0084)	(0.0067)	(0.0079)	(0.0146)
Pred. Lvl Effect / 100,000 BOE	-1.5	8.4	11.7	537K	525K	1.19M
Observations	3,990	3,990	3,990	4,200	4,200	4,200
R^2	0.99678	0.98791	0.99274	0.99523	0.99468	0.98620
Within R^2	0.10687	-0.02103	0.03708	0.18571	-0.04772	-0.10928

Notes: The dependent variables are the inverse hyperbolic sine of county-level jobs and earnings for three worker groups: home-county workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work). The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports predicted effects, representing the predicted level change in the dependent variable for a 100,000 BOE increase in oil and gas, computed using the estimated elasticity at the mean of the dependent variable and oil and gas. In fixed-effects IV regressions, the "within R^2" is calculated as $R^2_{\text{within}} = 1 - \frac{SSR_{\text{within}}}{TSS_{\text{within}}}$, where SSR_{within} is the sum of squared residuals after removing fixed effects, and TSS_{within} is the total sum of squares of deviations from group means. In 2SLS, SSR_{within} can exceed TSS_{within}, producing unbounded negative values that do not reflect model fit. Therefore, the "within R^2" is reported for transparency, but it should not be interpreted as a conventional measure of model fit. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Impact of Oil and Gas on the Number of Employment Across Three Monthly Earnings Brackets

	Home-county Jobs			Job Inflows			Job Outflows		
	$\leq$ \$1250	\$1251 to \$3333	$>$ \$3333	$\leq$ \$1250	\$1251 to \$3333	$>$ \$3333	$\leq$ \$1250	\$1251 to \$3333	$>$ \$3333
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Oil and Gas Production	-0.0204***	-0.0058	0.0283***	0.0034	0.0039	0.0391***	0.0081	0.0122	0.0515***
	(0.0068)	(0.0052)	(0.0092)	(0.0067)	(0.0079)	(0.0128)	(0.0076)	(0.0075)	(0.0156)
Pred. Lvl Effect / 100,000 BOE	-2.9	-1.1	6.5	0.4	0.7	10.0	0.9	1.7	11.1
Observations	3,990	3,990	3,990	3,990	3,990	3,990	3,990	3,990	3,990
R^2	0.99510	0.99535	0.99136	0.98728	0.98566	0.98331	0.99335	0.99176	0.98244
Within R^2	-0.01453	0.00078	-0.02838	0.04547	-0.00751	-0.19857	0.14002	0.02831	-0.10233

Notes: The dependent variables are the inverse hyperbolic sine of county-level job counts for three worker groups: home-county workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work), measured across three monthly earnings brackets: \$1,250/month or less, \$1,251/month to \$3,333/month, and over \$3,333/month. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. The earnings data cover 2001–2020 for 210 counties, resulting in 4,200 county-year observations. The job data cover 2002–2020, yielding 3,990 county-year observations. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Impact of Oil and Gas on Earnings per Job

	Earnings per Job		
	Home-county	Job inflows	Job outflows
	(1)	(2)	(3)
Oil and Gas Production	0.0137*	0.0059	0.0310**
	(0.0077)	(0.0101)	(0.0130)
Pred. Lvl Effect / 100,000 BOE	36.5	6.8	40.5
Observations	3,990	3,990	3,990
R^2	0.71835	0.85134	0.78032
Within R^2	0.00788	0.00615	-0.06727

Notes: The dependent variables are the inverse hyperbolic sine of county-level earnings per job, calculated by dividing total earnings by total jobs, for three worker groups: home-county workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work). The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Impact of Oil and Gas Production on Home-County Employment by Labor Market Segment and Age Group

	Home-county jobs					
	Labor Market Segment			Age Groups		
	Goods Prod.	Trade, Transp., Utils.	Other Serv.	29 or younger	30 to 54	55 or older
	(1)	(2)	(3)	(4)	(5)	(6)
Oil and Gas Production	-0.0002	-0.0056	-0.0025	-0.0018	-0.0056	-0.0020
	(0.0063)	(0.0051)	(0.0047)	(0.0103)	(0.0070)	(0.0046)
Pred. Lvl Effect / 100,000 BOE	0.0	-1.7	-0.4	-0.2	-0.5	-0.7
Observations	3,990	3,990	3,990	3,990	3,990	3,990
R^2	0.98701	0.99410	0.99661	0.99558	0.99632	0.99594
Within R^2	0.00754	0.04673	0.11398	0.09221	0.11753	0.03602

Notes: The dependent variables are the inverse hyperbolic sine of county-level job counts for home-county workers (locals working in their home county). Columns 1–3 present results by three labor market segments: Goods Producing, Trade, Transportation, and Utilities, and All Other Services. Columns 4–6 present results by three age groups: 29 or younger, 30 to 54, and 55 or older. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Impact of Oil and Gas on Inflow and Outflow Employment by Age Group

	Job Inflows			Job Outflows		
	29 or younger (1)	30 to 54 (2)	55 or older (3)	29 or younger (4)	30 to 54 (5)	55 or older (6)
Oil and Gas Production	0.0185**	0.0107	0.0174*	0.0316***	0.0223**	0.0246**
	(0.0079)	(0.0074)	(0.0090)	(0.0093)	(0.0086)	(0.0089)
Pred. Lvl Effect / 100,000 BOE	2.2	3.2	2.3	3.4	5.5	2.6
Observations	3,990	3,990	3,990	3,990	3,990	3,990
R^2	0.98762	0.98770	0.98609	0.99167	0.99202	0.99126
Within R^2	-0.03316	-0.00019	-0.03640	0.04262	0.06356	-0.00240

Notes: The dependent variables are the inverse hyperbolic sine of county-level job counts for inflow (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work) across three labor market segments: Goods Producing, Trade, Transportation, and Utilities, and All Other Services. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Impact of Oil and Gas on Inflow and Outflow Employment by Labor Market Segment

	Job Inflows			Job Outflows		
	Goods Prod. (1)	Trade, Transp., Utils. (2)	Other Serv. (3)	Goods Prod. (4)	Trade, Transp., Utils. (5)	Other Serv. (6)
Oil and Gas Production	0.0262**	0.0173**	0.0112	0.0488**	0.0271**	0.0133*
	(0.0094)	(0.0072)	(0.0089)	(0.0184)	(0.0103)	(0.0072)
Pred. Lvl Effect / 100,000 BOE	2.6	2.3	3.6	5.1	2.9	3.3
Observations	3,990	3,990	3,990	3,990	3,990	3,990
R^2	0.97881	0.98751	0.98727	0.96430	0.99146	0.99307
Within R^2	-0.06429	-0.01745	-0.01484	-0.04352	0.02155	0.07501

Notes: The dependent variables are the inverse hyperbolic sine of county-level job counts for inflow (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work) across three labor market segments: Goods Producing, Trade, Transportation, and Utilities, and All Other Services. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Impact of Oil and Gas on Dividends, Interest, & Rent, and Income

	Dividends, Interest, & Rent	Personal Income
	(1)	(2)
Oil and Gas Production	0.0160***	0.0088***
	(0.0047)	(0.0028)
Pred. Lvl Effect / 100,000 BOE	316K	1.05M
Observations	4,200	4,200
R^2	0.99744	0.99885
Within R^2	0.22562	0.33198

Notes: The dependent variables are the inverse hyperbolic sine of county-level dividends, interest, & rent and personal income. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10: Impact of Oil and Gas on Personal Income Ratios

	$\frac{\text{Personal Income}}{\text{Oil \& Gas Revenue}}$ (3)	$\frac{\text{Personal Income}}{\text{GDP}}$ (4)
Oil and Gas Production	-0.8571***	-0.0449***
	(0.0281)	(0.0093)
Observations	2,705	4,200
R^2	0.99479	0.81602
Within R^2	0.96745	-0.18851

Notes: The dependent variables are the inverse hyperbolic sine of personal income to oil and gas revenue and GDP. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 11: Impact of Oil and Gas on In-migration and Out-migration

	In-migration (1)	Out-migration (2)
Oil and Gas Production	0.0158	0.0118
	(0.0351)	(0.0361)
Pred. Lvl Effect / 100,000 BOE	0.8	0.6
Observations	4,200	4,200
R^2	0.93596	0.94054
Within R^2	0.03196	0.05048

Notes: The dependent variables are the inverse hyperbolic sine of county-level in- and out-migration counts. The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports the predicted level effects, showing the expected change in the dependent variable for a 100,000 BOE increase in oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 12: Impact of Oil and Gas on Particulate Matter 2.5

	Receptor-cell $PM_{2.5}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Oil and Gas Production Within 0-1 km	0.0405						
	(0.0265)						
Oil and Gas Production Within 1-2 km		0.0360**					
		(0.0132)					
Oil and Gas Production Within 2-3 km			0.0393*				
			(0.0223)				
Oil and Gas Production Within 3-5 km				0.0266			
				(0.0174)			
Oil and Gas Production Within 5-10 km					0.0158		
					(0.0122)		
Oil and Gas Production Within 10-15 km						0.0144	
						(0.0103)	
Oil and Gas Production Within 15-20 km							0.0060
							(0.0047)
Pred. Lvl Effect / 100,000 BOE	4.12	1.19	0.77	0.16	0.02	0.01	0.00
Observations	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280
R^2	0.94567	0.95233	0.94828	0.96549	0.97591	0.97707	0.98025
Within R^2	-1.8153	-1.4700	-1.6797	-0.78794	-0.24833	-0.18816	-0.02327

Notes: The dependent variable is the inverse hyperbolic sine of annual average $PM_{2.5}$ in the receptor cell. The main explanatory variable is the inverse hyperbolic sine of oil and gas production within concentric distance bands (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10–15 km, and 15–20 km) from each receptor cell. Coefficients can be interpreted as the percent change in $PM_{2.5}$ given a 1% change in oil and gas production within a distance band of the cell. All regressions control for the inverse hyperbolic sine of cell population and production in the four adjacent bands, and include cell and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by cell and year, are shown in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 13: Risk Estimates for Long-Term $PM_{2.5}$ Exposure–Related Mortality

Study	At-Risk Age Group	Beta Coefficient (SE)
Krewski et al. (2009) – ACS	30 & Older	β = 0.0058 (0.000963)
Lepeule et al. (2012) – H6C	25 & Older	β = 0.0013 (0.00335)

Notes: Table reports β coefficients linking long-term $PM_{2.5}$ exposure to all-cause mortality risk from two major cohort studies [36, 38]. Beta coefficients represent the log-linear change in mortality risk per 1 $\mu g/m^3$ increase in annual average $PM_{2.5}$. Standard errors are shown in parentheses. Cohort acronyms: ACS = American Cancer Society Cancer Prevention Study II; H6C = Harvard Six Cities Study. The ACS cohort identifies the at-risk population as adults aged 30 and older, while the H6C cohort identifies the at-risk population as adults aged 25 and older. Adapted from Luke et al. (2025) [14].

Figures

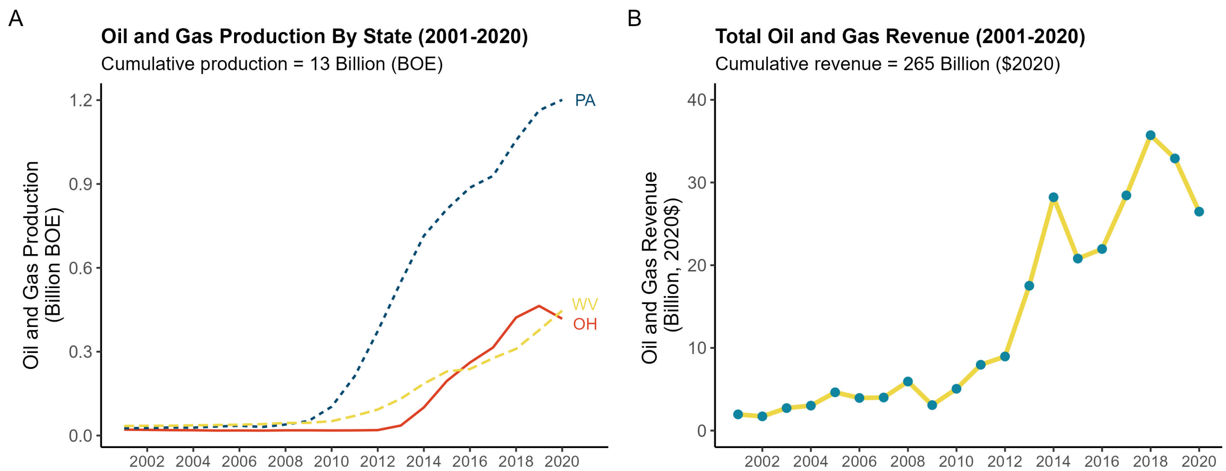

Figure 1: Oil and Gas Production and Revenue in Pennsylvania, Ohio, and West Virginia (2001 - 2020)

Notes: Panel A shows annual oil and gas production in billion barrels of oil equivalent (BOE) separately for Pennsylvania, Ohio, and West Virginia. Panel B presents total oil and gas revenue for all three states combined in 2020 dollars. Production data are sourced from Enverus [17], and revenue estimates are calculated using West Texas Intermediate crude oil and Henry Hub natural gas prices from FRED [65, 66].

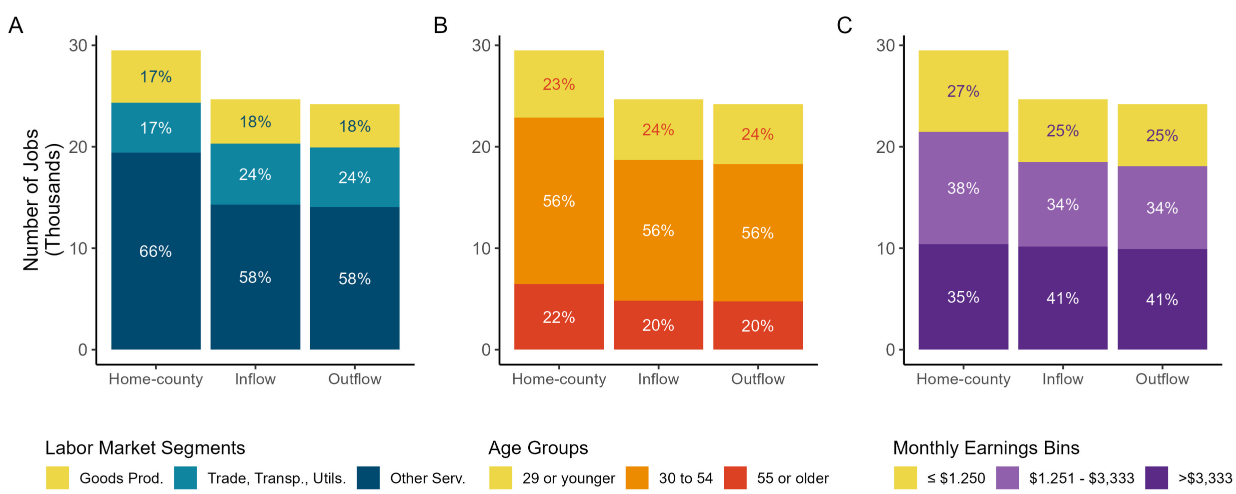

Figure 2: Annual Average Employment Distribution by Labor Market Segment, Age Cohort, and Monthly Earnings Bin

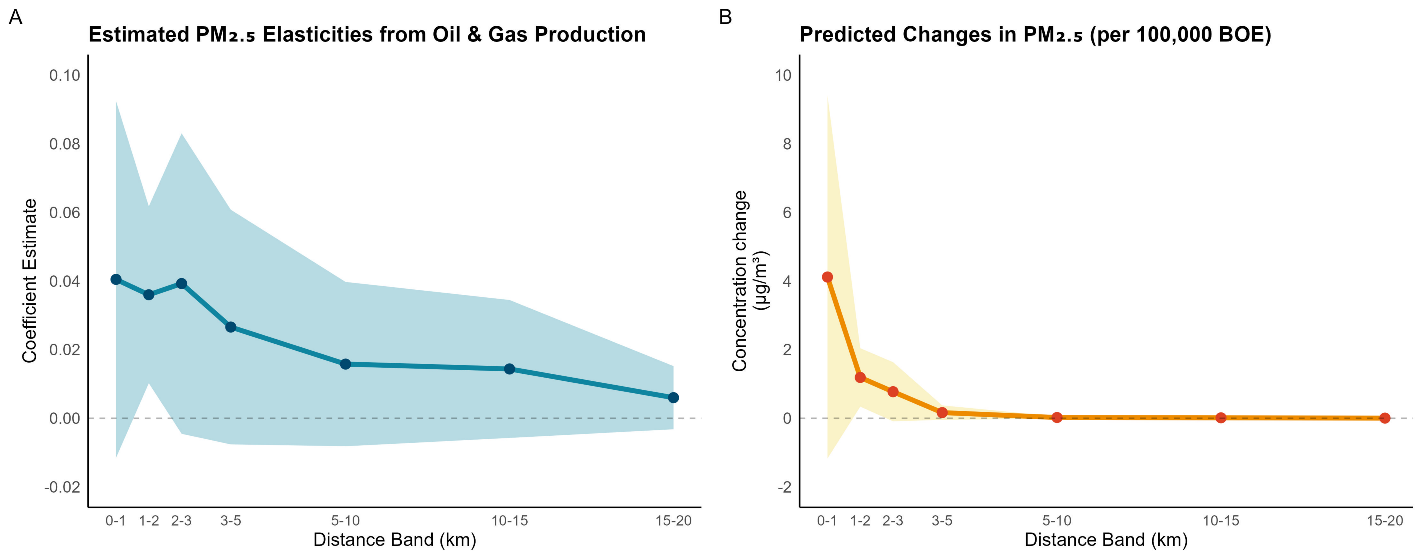

Figure 3: Impacts of Oil and Gas on Particulate Matter 2.5

Notes: This figure shows results from regressions where the dependent variable is the inverse hyperbolic sine of cell-level $PM_{2.5}$. Panel A presents estimated coefficients for oil and gas production measured within the following distance bands: 0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10-15 km, and 15-20 km. Coefficients can be interpreted as the percent change in PM $PM_{2.5}$ given a 1% change in production within each band. Panel B shows predicted level effects, representing the expected change in $PM_{2.5}$ (in micrograms per cubic meter) for a 100,000 BOE increase in oil and gas production.

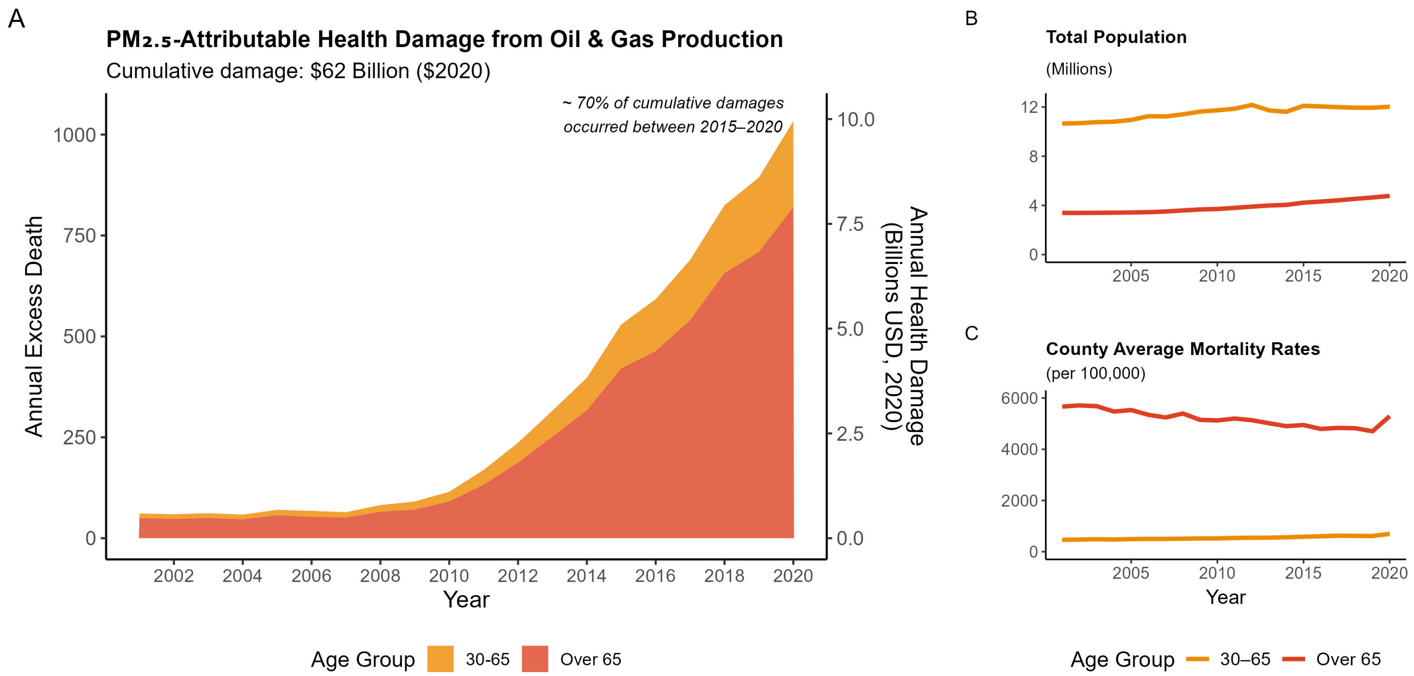

Figure 4: Health Damage of Oil and Gas Production

Notes: This figure presents the monetized health damage derived from the statistically significant coefficients in Figure 3 for the 1–2 km and 2–3 km distance bands. Panel A shows estimated annual premature mortality and resulting monetized annual health damage attributable to elevated $PM_{2.5}$ from oil and gas production in Pennsylvania, Ohio, and West Virginia from 2001 to 2020, for two at-risk age groups: 30–65 and over 65. Panel B plots total population trends, and Panel C displays county average mortality rates over the same period in the tri-state region. Together, Panels B and C provide demographic and baseline health context for interpreting the mortality-related damages in Panel A.

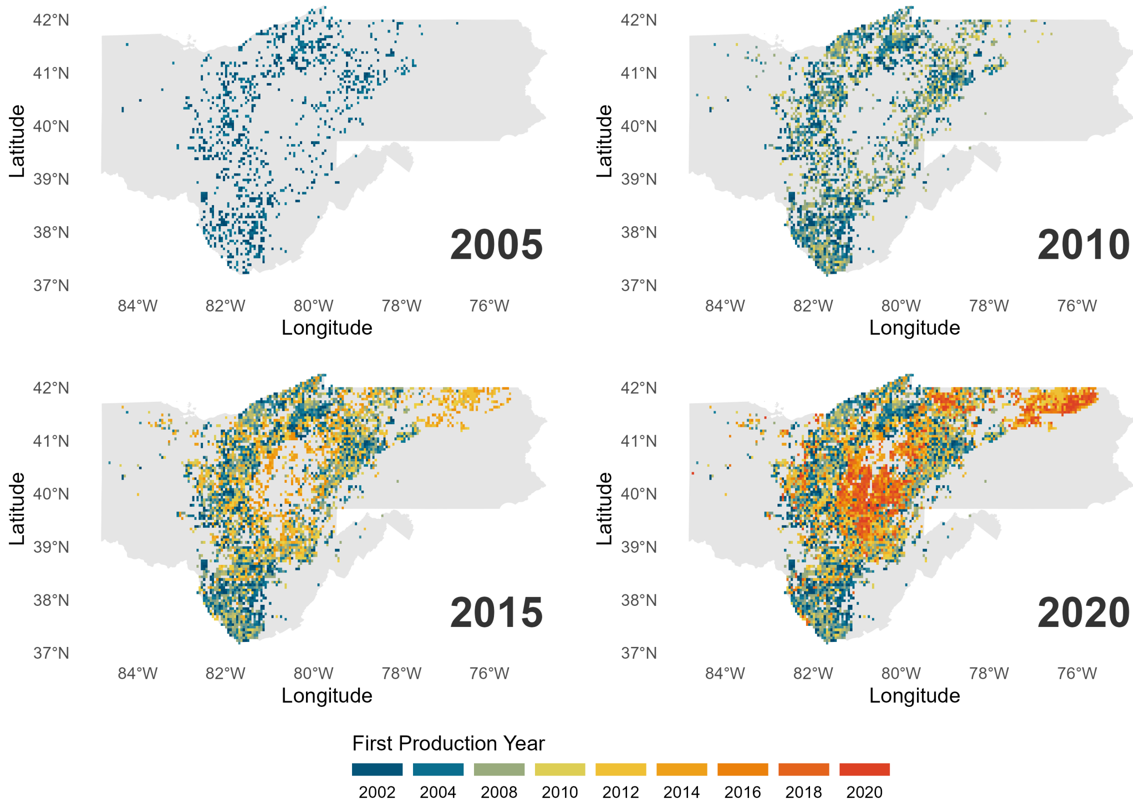

Figure 5: Geographic Distribution of Oil and Gas Wells by Year of First Production Over Time

This figure maps the spatial distribution of & and gas wells based on their first year of production, grouped into four years: 2005, 2010, 2015, and 2020. Wells are color-coded by their first production year to illustrate how new development has expanded or shifted over time.

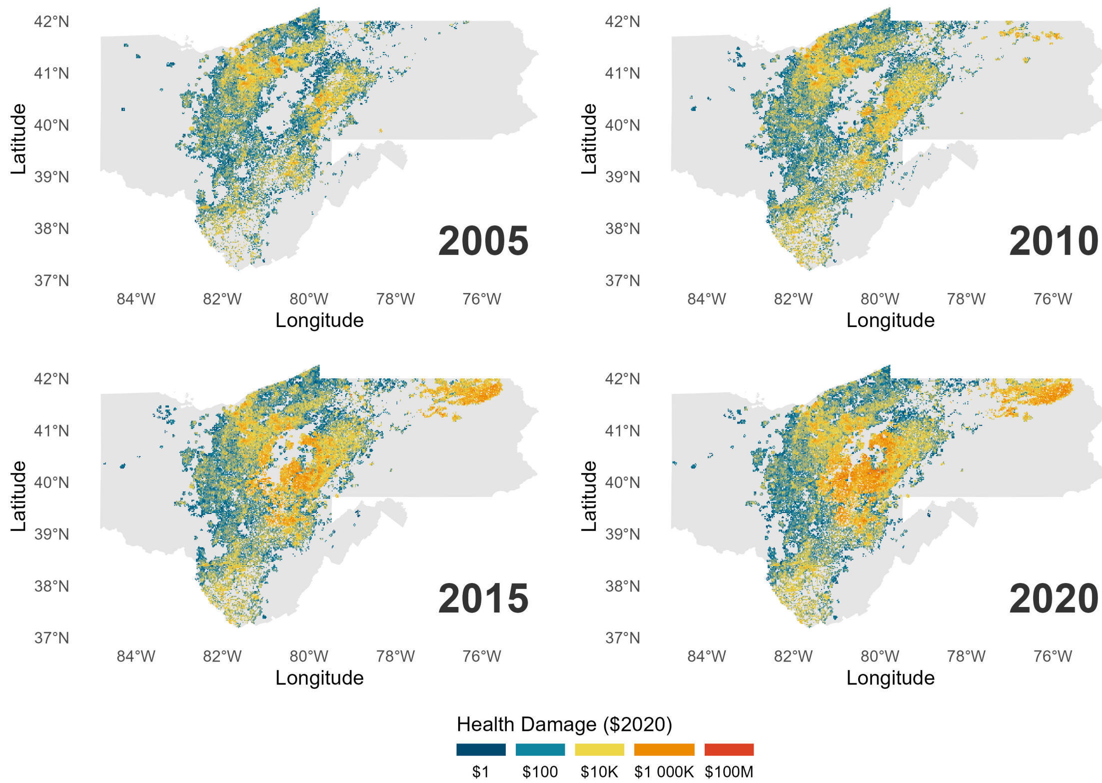

Figure 6: Geographic Distribution of Health Damages from Oil and Gas Over Time

Notes: This figure maps granular health damages from oil and gas production in the tri-state region, showing how impacts vary geographically across four years: 2005, 2010, 2015, and 2020. Colors or shading indicate the magnitude of damage in each area.

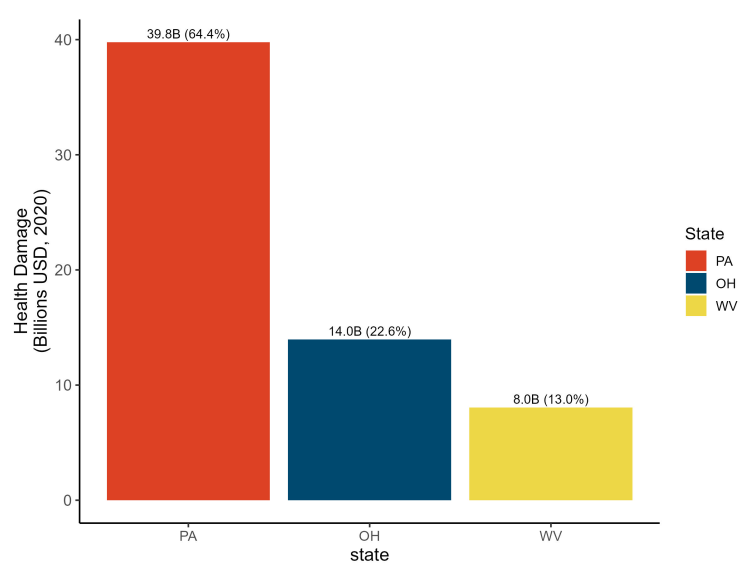

Figure 7: Cumulative Health Damage By State (2001-2020)

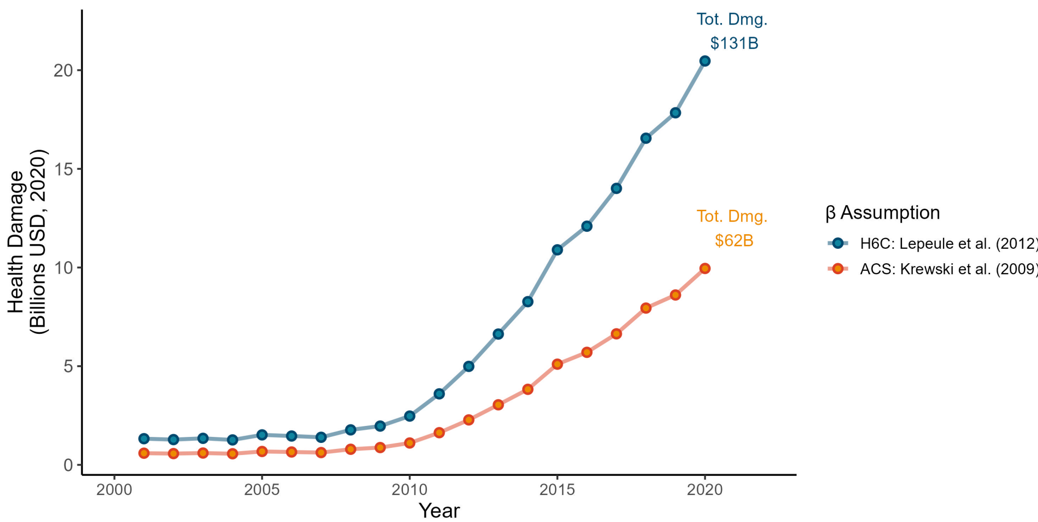

Figure 8: Estimated Health Damages from Long-Term $PM_{2.5}$ Exposure Using Alternative Risk Assumptions

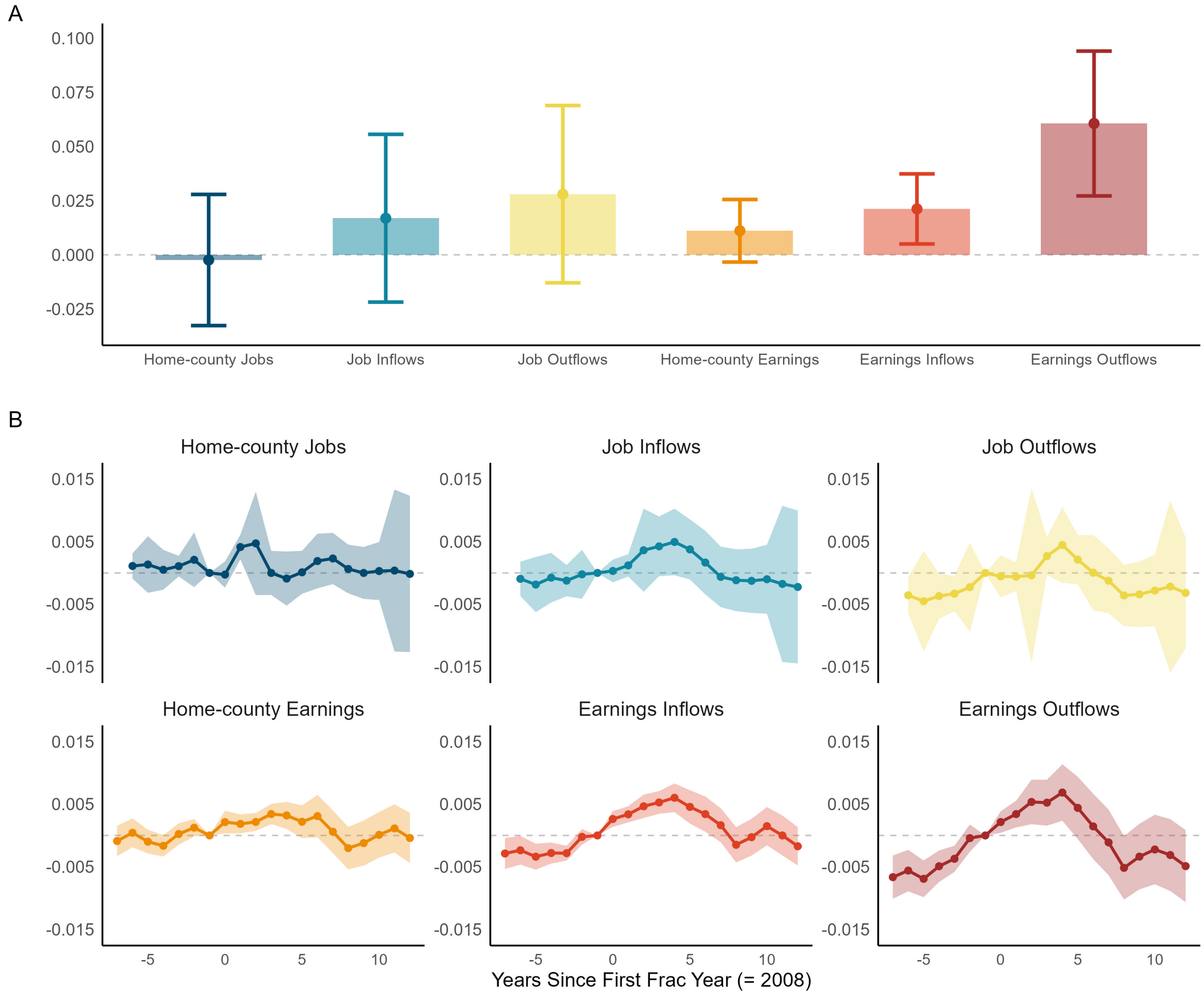

Figure 9: Dynamic Economic Effects of Oil and Gas Extraction

Notes: This figure presents the dynamic economic effects of oil and gas extraction. Shaded areas represent 95% confidence intervals. Panel A shows baseline earnings and job outcomes in 2007, one year before the onset of fracking in the region. Panel B displays interaction effects relative to this reference year.

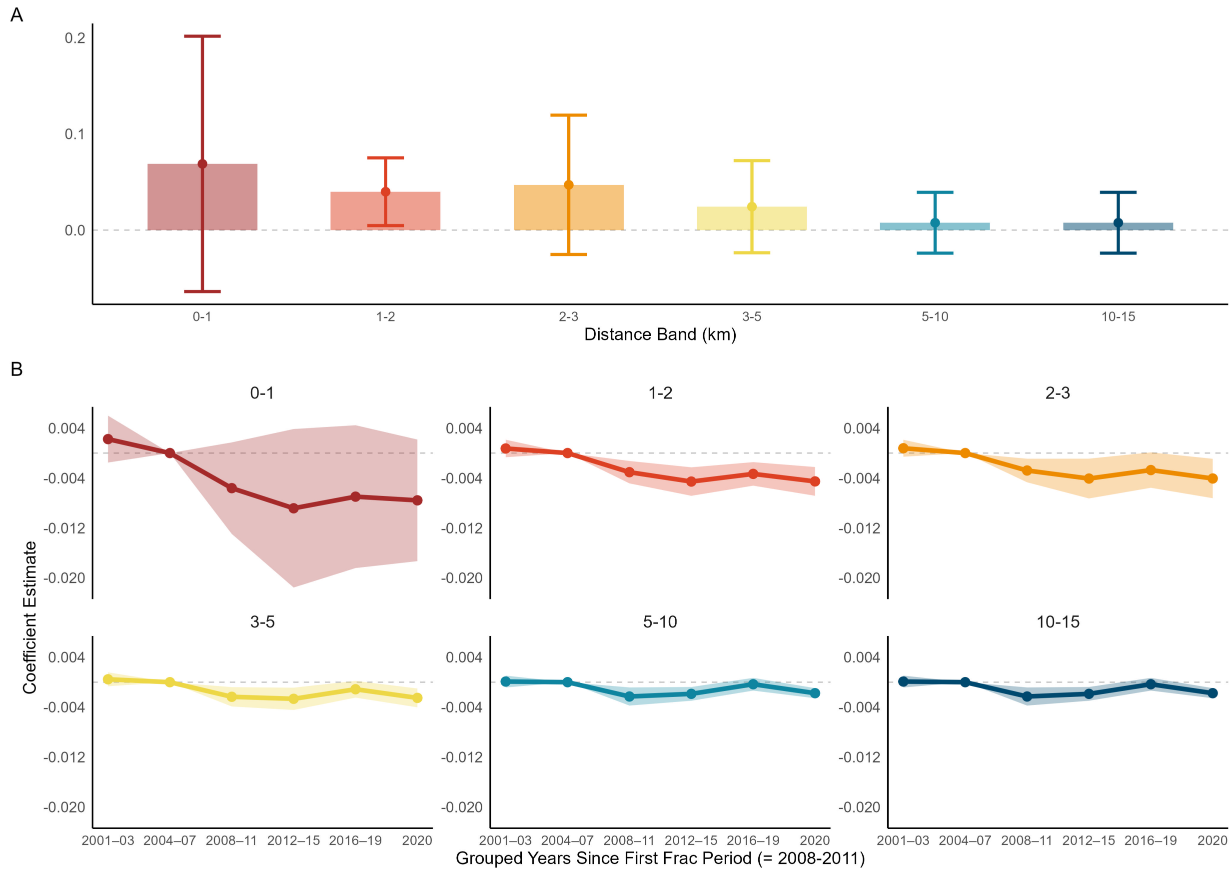

Figure 10: Dynamic Air Quality Effects of Oil and Gas Extraction

Notes: This figure presents the dynamic effects of oil and gas extraction across distance bands on air quality. Shaded areas represent 95% confidence intervals. Panel A shows baseline air quality levels in 2004–2007, the reference period prior to the onset of fracking in the region. Panel B displays interaction effects relative to this baseline, capturing both pre- and post-fracking changes. Years are grouped in four-year intervals to reduce computational burden, with shorter first and last groups (2001–2003 and 2020). Results are shown for five distance bands from each receptor cell: 0–1 km, 2–3 km, 3–5 km, 5–10 km, and 10–15 km.

References

[1] John L. Adgate, Bernard D. Goldstein, and Lisa M. McKenzie. "Potential public health hazards, exposures and health effects from unconventional natural gas development". In: *Environmental Science & Technology* 48.15 (2014), pp. 8307–8320. ISSN: 1520-5851. DOI: 10.1021/es404621d.

[2] Hunt Allcott and Daniel Keniston. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America". In: *The Review of Economic Studies* 85.2 (Apr. 1, 2018), pp. 695–731. ISSN: 0034-6527. DOI: 10.1093/restud/rdx042. URL: https://doi.org/10.1093/restud/rdx042 (visited on 09/02/2025).

[3] Zoya Banan and Jeremy M. Gernand. "Emissions of particulate matter due to Marcellus Shale gas development in Pennsylvania: Mapping the implications". In: *Energy Policy* 148 (Jan. 1, 2021), p. 111979. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2020.111979. URL: https://www.sciencedirect.com/science/article/pii/S030142152030690X.

[4] Alexander W. Bartik et al. "The Local Economic and Welfare Consequences of Hydraulic Fracturing - American Economic Association". In: *American Economic Review* 11.4 (2019), pp. 105–155. URL: https://www.aeaweb.org/articles?id=10.1257/app.20170487.

[5] Trinidad Beleche and Inna Cintina. "Fracking and risky behaviors: Evidence from Pennsylvania". In: *Economics and Human Biology* 31 (Sept. 2018), pp. 69–82. ISSN: 1873-6130. DOI: 10.1016/j.ehb.2018.08.001.

[6] Marc F. Bellemare and Casey J. Wichman. "Elasticities and the Inverse Hyperbolic Sine Transformation". In: *Oxford Bulletin of Economics and Statistics* 82.1 (2020). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12325, pp. 50–61. ISSN: 1468-0084. DOI: 10.1111/obes.12325. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/obes.12325 (visited on 08/28/2025).

[7] Jason P. Brown, Timothy Fitzgerald, and Jeremy G. Weber. "Capturing rents from natural resource abundance: Private royalties from U.S. onshore oil & gas production". In: *Resource and Energy Economics* 46 (Nov. 1, 2016), pp. 23–38. ISSN: 0928-7655. DOI: 10.1016/j.reseneeco.2016.07.003. URL: https://www.sciencedirect.com/science/article/pii/S0928765516300355 (visited on 10/02/2025).

[8] Jason P. Brown, Timothy Fitzgerald, and Jeremy G. Weber. "Does Resource Ownership Matter? Oil and Gas Royalties and the Income Effect of Extraction". In: *Journal of the Association of Environmental and Resource Economists* 6.6 (Nov. 2019). Publisher: The University of Chicago Press, pp. 1039–1064.

ISSN: 2333-5955. DOI: 10.1086/705505. URL: https://www.journals.uchicago.edu/doi/10.1086/705505.

[9] Jonathan J Buonocore et al. “Air pollution and health impacts of oil & gas production in the United States”. In: *Environmental Research: Health* 1.2 (May 2023). Publisher: IOP Publishing, p. 021006. ISSN: 2752-5309. DOI: 10.1088/2752-5309/acc886. URL: https://dx.doi.org/10.1088/2752-5309/acc886 (visited on 09/02/2025).

[10] US Census Bureau. *Design Comparison of LODES and ACS Commuting Data Products*. Census.gov. Section: Government. Sept. 2014. URL: https://www.census.gov/library/working-papers/2014/adrm/ces-wp-14-38.html (visited on 09/02/2025).

[11] US Census Bureau. *Gridded Environmental Impacts Frame (Gridded EIF)*. Census.gov. Section: Government. URL: https://www.census.gov/data/experimental-data-products/gridded-eif.html (visited on 07/19/2025).

[12] Jiafeng Chen and Jonathan Roth. “Logs with Zeros? Some Problems and Solutions*”. In: *The Quarterly Journal of Economics* 139.2 (May 1, 2024), pp. 891–936. ISSN: 0033-5533. DOI: 10.1093/qje/qjad054. URL: https://doi.org/10.1093/qje/qjad054 (visited on 09/04/2025).

[13] Janet Currie, Michael Greenstone, and Katherine Meckel. “Hydraulic fracturing and infant health: New evidence from Pennsylvania”. In: *Science Advances* 3.12 (Dec. 2017), e1603021. ISSN: 2375-2548. DOI: 10.1126/sciadv.1603021.

[14] Luke R. Dennin and Nicholas Z. Muller. “Funding a Just Transition Away from Coal in the U.S. Considering Avoided Damage from Air Pollution”. In: *Journal of Benefit-Cost Analysis* 16.1 (Mar. 2025), pp. 79–106. ISSN: 2194-5888, 2152-2812. DOI: 10.1017/bca.2024.20. URL: https://www.cambridge.org/core/journals/journal-of-benefit-cost-analysis/article/funding-a-just-transition-away-from-coal-in-the-us-considering-avoided-damage-from-air-pollution/5F8819C174300AA155A1726597F45759 (visited on 09/06/2025).

[15] Aaron van Donkelaar et al. “Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty”. In: *Environmental Science & Technology* 55.22 (Nov. 16, 2021). Publisher: American Chemical Society, pp. 15287–15300. ISSN: 0013-936X. DOI: 10.1021/acs.est.1c05309. URL: https://doi.org/10.1021/acs.est.1c05309 (visited on 12/19/2022).

[16] US Census Bureau Center for Economic Studies. *US Census Bureau Center for Economic Studies Publications and Reports Page*. 2022. URL: https://lehd.ces.census.gov/data/ (visited on 12/11/2022).

[17] *Enverus — Creating the future of energy together.* July 3, 2025. URL: https://www.enverus.com/ (visited on 07/19/2025).

[18] Ozkan Eren and Emily Owens. *Economic Booms and Recidivism — Journal of Quantitative Criminology*. May 16, 2023. URL: https://link.springer.com/article/10.1007/s10940-023-09571-2 (visited on 09/08/2025).

[19] James Feyrer, Erin T. Mansur, and Bruce Sacerdote. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution - American Economic Association". In: *American Economic Review* 107.4 (2017), pp. 1313–1334. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20151326.

[20] R. Kaj Gittings and Travis Roach. *Who Really Benefits from a Resource Boom? Evidence from the Marcellus and Utica Shale Plays*. Rochester, NY, Feb. 19, 2019. DOI: 10.2139/ssrn.3347229. URL: https://papers.ssrn.com/abstract=3347229 (visited on 12/11/2022).

[21] David J. X. Gonzalez et al. "Upstream oil and gas production and ambient air pollution in California". In: *Science of The Total Environment* 806 (Feb. 1, 2022), p. 150298. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2021.150298. URL: https://www.sciencedirect.com/science/article/pii/S0048969721053754.

[22] David J. X. Gonzalez et al. "Historic redlining and the siting of oil and gas wells in the United States". In: *Journal of Exposure Science & Environmental Epidemiology* 33.1 (Jan. 2023). Publisher: Nature Publishing Group, pp. 76–83. ISSN: 1559-064X. DOI: 10.1038/s41370-022-00434-9. URL: https://www.nature.com/articles/s41370-022-00434-9 (visited on 09/02/2025).

[23] David J. X. González et al. "Temporal Trends of Racial and Socioeconomic Disparities in Population Exposures to Upstream Oil and Gas Development in California". In: (Mar. 23, 2023). DOI: 10.1029/2022GH000690. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GH000690 (visited on 09/02/2025).

[24] Corbett Grainger and Andrew Schreiber. "Discrimination in Ambient Air Pollution Monitoring?" In: *AEA Papers and Proceedings* 109 (May 2019), pp. 277–282. ISSN: 2574-0768. DOI: 10.1257/pandp.20191063. URL: https://www.aeaweb.org/articles?id=10.1257/pandp.20191063 (visited on 09/02/2025).

[25] Mouhcine Guettabi and Alexander James. "Who benefits from an oil boom? Evidence from a unique Alaskan data set". In: *Resource and Energy Economics* 62 (Nov. 1, 2020), p. 101200. ISSN: 0928-7655.

DOI: 10.1016/j.reseneeco.2020.101200. URL: https://www.sciencedirect.com/science/article/pii/S0928765519301113 (visited on 12/18/2022).

[26] Catherine Hausman and Ryan Kellogg. *Welfare and Distributional Implications of Shale Gas*. Rochester, NY, Apr. 1, 2015. URL: https://papers.ssrn.com/abstract=2599381 (visited on 07/19/2025).

[27] Elaine L. Hill. "Shale gas development and infant health: Evidence from Pennsylvania". In: *Journal of Health Economics* 61 (Sept. 2018), pp. 134–150. ISSN: 1879-1646. DOI: 10.1016/j.jhealeco.2018.07.004.

[28] Elaine L. Hill. "The Impact of Oil and Gas Extraction on Infant Health". In: *American Journal of Health Economics* 10.1 (Jan. 2024). Publisher: The University of Chicago Press, pp. 68–96. ISSN: 2332-3493. DOI: 10.1086/724218. URL: https://www.journals.uchicago.edu/doi/full/10.1086/724218 (visited on 09/02/2025).

[29] Elaine L. Hill and Lala Ma. "Drinking water, fracking, and infant health". In: *Journal of Health Economics* 82 (Mar. 2022), p. 102595. ISSN: 1879-1646. DOI: 10.1016/j.jhealeco.2022.102595.

[30] Daniel J. Jacob. *Introduction to atmospheric chemistry*. Online-Ausg. Princeton, N.J: Princeton University Press, 1999. 266 pp. ISBN: 978-1-4008-4154-7 978-0-691-00185-2.

[31] Grant D. Jacobsen and Dominic P. Parker. "The Economic Aftermath of Resource Booms: Evidence from Boomtowns in the American West". In: *The Economic Journal* 126.593 (June 1, 2016), pp. 1092–1128. ISSN: 0013-0133. DOI: 10.1111/ecoj.12173. URL: https://doi.org/10.1111/ecoj.12173.

[32] Thomas Jemielita et al. "Unconventional Gas and Oil Drilling Is Associated with Increased Hospital Utilization Rates". In: *PloS One* 10.7 (2015), e0131093. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0131093.

[33] Brenna C. Kelly et al. "Racial and Ethnic Disparities in Regulatory Air Quality Monitor Locations in the US". In: (Dec. 4, 2024). URL: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2827225 (visited on 09/02/2025).

[34] Tim Komarek and Attila Cseh. "Fracking and public health: Evidence from gonorrhea incidence in the Marcellus Shale region". In: *Journal of Public Health Policy* 38.4 (Nov. 2017), pp. 464–481. ISSN: 1745-655X. DOI: 10.1057/s41271-017-0089-5.

[35] Timothy M. Komarek. "Labor market dynamics and the unconventional natural gas boom: Evidence from the Marcellus region". In: *Resource and Energy Economics* 45 (Aug. 1, 2016), pp. 1–17. ISSN: 0928-7655. DOI: 10.1016/j.reseneeco.2016.03.004. URL: https://www.sciencedirect.com/science/article/pii/S092876551630077X.

[36] Daniel Krewski et al. "Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality". In: *Research Report (Health Effects Institute)* 140 (May 2009), 5–114, discussion 115–136. ISSN: 1041-5505.

[37] Ian Lange and Michael Redlinger. "Effects of stricter environmental regulations on resource development". In: *Journal of Environmental Economics and Management* 96 (July 2019). Publisher: Elsevier BV, pp. 60–87. ISSN: 0095-0696. DOI: 10.1016/j.jeem.2019.04.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0095069618303437 (visited on 07/19/2025).

[38] Johanna Lepeule et al. "Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009". In: *Environmental Health Perspectives* 120.7 (July 2012), pp. 965–970. ISSN: 1552-9924. DOI: 10.1289/ehp.1104660.

[39] Aviva Litovitz et al. "Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania". In: *Environmental Research Letters* 8.1 (Jan. 2013). Publisher: IOP Publishing, p. 014017. ISSN: 1748-9326. DOI: 10.1088/1748-9326/8/1/014017. URL: https://dx.doi.org/10.1088/1748-9326/8/1/014017.

[40] Erin N. Mayfield et al. "Cumulative environmental and employment impacts of the shale gas boom". In: *Nature Sustainability* 2.12 (Dec. 2019). Number: 12 Publisher: Nature Publishing Group, pp. 1122–1131. ISSN: 2398-9629. DOI: 10.1038/s41893-019-0420-1. URL: https://www.nature.com/articles/s41893-019-0420-1.

[41] John Mullahy and Edward C. Norton. "Why Transform Y? The Pitfalls of Transformed Regressions with a Mass at Zero". In: *Oxford Bulletin of Economics and Statistics* 86.2 (2024). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12583, pp. 417–447. ISSN: 1468-0084. DOI: 10.1111/obes.12583. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/obes.12583 (visited on 09/15/2025).

[42] *Multiple Cause of Death Data on CDC WONDER*. URL: https://wonder.cdc.gov/mcd.html (visited on 07/19/2025).

[43] *Natural Gas Dry Production*. URL: https://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_FPD_mmcf_a.htm (visited on 07/19/2025).

[44] Richard G. Newell, Brian C. Prest, and Ashley B. Vissing. "Trophy Hunting versus Manufacturing Energy: The Price Responsiveness of Shale Gas". In: *Journal of the Association of Environmental and Resource Economists* 6.2 (Mar. 2, 2019). Publisher: University of Chicago Press, pp. 391–431. ISSN:

2333-5955, 2333-5963. DOI: 10.1086/701531. URL: https://www.journals.uchicago.edu/doi/10.1086/701531 (visited on 07/19/2025).

[45] Richard G. Newell and Daniel Raimi. "The fiscal impacts of increased U.S. oil and gas development on local governments". In: *Energy Policy* 117 (June 1, 2018), pp. 14–24. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2018.02.042. URL: https://www.sciencedirect.com/science/article/pii/S0301421518301198 (visited on 09/02/2025).

[46] Richard G. Newell and Daniel Raimi. "US state and local oil and gas revenue sources and uses". In: *Energy Policy* 112 (Jan. 1, 2018), pp. 12–18. ISSN: 0301-4215. DOI: 10.1016/j.enpol.2017.10.002. URL: https://www.sciencedirect.com/science/article/pii/S0301421517306250.

[47] Edward C. Norton. "The inverse hyperbolic sine transformation and retransformed marginal effects". In: *The Stata Journal* 22.3 (Sept. 1, 2022). Publisher: SAGE Publications, pp. 702–712. ISSN: 1536-867X. DOI: 10.1177/1536867X221124553. URL: https://doi.org/10.1177/1536867X221124553 (visited on 09/04/2025).

[48] Barry G. Rabe. "Shale Play Politics: The Intergovernmental Odyssey of American Shale Governance". In: *Environmental Science & Technology* 48.15 (Aug. 5, 2014). Publisher: American Chemical Society, pp. 8369–8375. ISSN: 0013-936X. DOI: 10.1021/es4051132. URL: https://doi.org/10.1021/es4051132.

[49] Sara G. Rasmussen et al. "Association Between Unconventional Natural Gas Development in the Marcellus Shale and Asthma Exacerbations". In: *JAMA internal medicine* 176.9 (Sept. 1, 2016), pp. 1334–1343. ISSN: 2168-6114. DOI: 10.1001/jamainternmed.2016.2436.

[50] *Reconsideration of the National Ambient Air Quality Standards for Particulate Matter*. Federal Register. Mar. 6, 2024. URL: https://www.federalregister.gov/documents/2024/03/06/2024-02637/reconsideration-of-the-national-ambient-air-quality-standards-for-particulate-matter (visited on 09/07/2025).

[51] Yusuf H. Roohani et al. "Impact of natural gas development in the Marcellus and Utica shales on regional ozone and fine particulate matter levels". In: *Atmospheric Environment* 155 (Apr. 1, 2017), pp. 11–20. ISSN: 1352-2310. DOI: 10.1016/j.atmosenv.2017.01.001. URL: https://www.sciencedirect.com/science/article/pii/S1352231017300018.

[52] Jeffrey Rous et al. "Evaluating determinants of shale gas well locations in an urban setting". In: *The Annals of Regional Science* 65.3 (Dec. 1, 2020), pp. 645–671. ISSN: 1432-0592. DOI: 10.1007/s00168-020-00998-0. URL: https://doi.org/10.1007/s00168-020-00998-0 (visited on 09/02/2025).

[53] Anirban A. Roy, Peter J. Adams, and Allen L. Robinson. “Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas”. In: *Journal of the Air & Waste Management Association* 64.1 (Jan. 2, 2014), pp. 19–37. ISSN: 1096-2247. DOI: 10.1080/10962247.2013.826151. URL: https://doi.org/10.1080/10962247.2013.826151 (visited on 12/11/2022).

[54] John H. Seinfeld and Spyros N. Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 3rd Edition — Wiley*. Wiley.com. 2016. URL: https://www.wiley.com/en-us/Atmospheric+Chemistry+and+Physics%3A+From+Air+Pollution+to+Climate+Change%2C+3rd+Edition-p-9781118947401 (visited on 07/19/2025).

[55] Brian Sergi et al. “Regional and county flows of particulate matter damage in the US”. In: *Environmental Research Letters* 15.10 (Oct. 2020). Publisher: IOP Publishing, p. 104073. ISSN: 1748-9326. DOI: 10.1088/1748-9326/abb429. URL: https://dx.doi.org/10.1088/1748-9326/abb429.

[56] *SOI tax stats - Migration data — Internal Revenue Service*. URL: https://www.irs.gov/statistics/soi-tax-stats-migration-data (visited on 07/19/2025).

[57] Tammy M. Thompson et al. “Modeling to Evaluate Contribution of Oil and Gas Emissions to Air Pollution”. In: *Journal of the Air & Waste Management Association* (Apr. 3, 2017). Publisher: Taylor & Francis. ISSN: 1096-2247. URL: https://www.tandfonline.com/doi/abs/10.1080/10962247.2016.1251508 (visited on 09/02/2025).

[58] Huy Tran et al. “Air Quality and Health Impacts of Onshore Oil and Gas Flaring and Venting Activities Estimated Using Refined Satellite-Based Emissions”. In: (Mar. 6, 2024). DOI: 10.1029/2023GH000938. URL: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2023GH000938 (visited on 09/02/2025).

[59] Michelle C. Turner et al. “Long-Term Ozone Exposure and Mortality in a Large Prospective Study”. In: *American Journal of Respiratory and Critical Care Medicine* 193.10 (May 15, 2016). Publisher: American Thoracic Society, pp. 1134–1142. ISSN: 1073-449X, 1535-4970. DOI: 10.1164/rccm.201508-1633oc. URL: https://www.atsjournals.org/doi/10.1164/rccm.201508-1633OC (visited on 07/19/2025).

[60] Aaron W. Tustin et al. “Associations between Unconventional Natural Gas Development and Nasal and Sinus, Migraine Headache, and Fatigue Symptoms in Pennsylvania”. In: *Environmental Health Perspectives* (Aug. 25, 2016). Publisher: National Institute of Environmental Health Sciences. DOI: 10.1289/EHP281. URL: https://ehp.niehs.nih.gov/doi/10.1289/EHP281 (visited on 09/02/2025).

[61] U.S. Bureau of Economic Analysis. *What is rental income of persons? — U.S. Bureau of Economic Analysis (BEA)*. 2022. URL: https://www.bea.gov/help/faq/64 (visited on 12/11/2022).

[62] U.S. Bureau of Economic Analysis BEA. *Regional Economic Accounts — U.S. Bureau of Economic Analysis (BEA)*. 2022. URL: https://www.bea.gov/data/economic-accounts/regional (visited on 12/11/2022).

[63] U.S. Bureau of Economic Analysis BEA. *Local Area Personal Income and Employment: Concepts and Methods — U.S. Bureau of Economic Analysis (BEA)*. Apr. 21, 2025. URL: https://www.bea.gov/resources/methodologies/local-area-personal-income-employment (visited on 09/07/2025).

[64] U.S. Bureau of Labor Statistics. *Oil and Gas Extraction - May 2020 OEWS Industry-Specific Occupational Employment and Wage Estimates*. 2022. URL: https://www.bls.gov/oes/2020/may/naics4_211100.htm (visited on 12/11/2022).

[65] U.S. Energy Information Administration. *Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma*. FRED, Federal Reserve Bank of St. Louis. Publisher: FRED, Federal Reserve Bank of St. Louis. 2022. URL: https://fred.stlouisfed.org/series/ACOILWTICO (visited on 12/11/2022).

[66] U.S. Energy Information Administration. *Henry Hub Natural Gas Spot Price*. FRED, Federal Reserve Bank of St. Louis. Publisher: FRED, Federal Reserve Bank of St. Louis. 2022. URL: https://fred.stlouisfed.org/series/AHHNGSP (visited on 12/11/2022).

[67] U.S. Energy Information Administration. *Maps: Oil and Gas Exploration, Resources, and Production - Energy Information Administration*. 2022. URL: https://www.eia.gov/maps/maps.htm (visited on 12/11/2022).

[68] U.S. Environmental Protection Agency. *Estimating $PM_{2.5}$- and Ozone-Attributable Health Benefits: Technical Support Document*. Technical Support Document. Research Triangle Park, NC: U.S. Environmental Protection Agency, June 2024. URL: https://www.epa.gov/system/files/documents/2024-06/estimating-pm2.5-and-ozone-attributable-health-benefits-tsd-2024.pdf.

[69] U.S. Environmental Protection Agency US EPA. *2017 National Emissions Inventory (NEI) Data*. June 30, 2017. URL: https://www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-data (visited on 12/11/2022).

[70] *U.S. Field Production of Crude Oil (Thousand Barrels per Day)*. URL: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=mcrfpus2&f=m (visited on 07/19/2025).

[71] *U.S. Natural Gas Gross Withdrawals (Million Cubic Feet per Day)*. URL: https://www.eia.gov/dnav/ng/hist/ngm_epg0_fgw_nus_mmcfdm.htm (visited on 07/19/2025).

[72] OP US EPA. *Mortality Risk Valuation*. Apr. 20, 2014. URL: https://www.epa.gov/environmental-economics/mortality-risk-valuation (visited on 07/19/2025).

[73] Karn Vohra et al. "The health burden and racial-ethnic disparities of air pollution from the major oil and gas lifecycle stages in the United States". In: *Science Advances* (Aug. 22, 2025). Publisher: American Association for the Advancement of Science. URL: https://www.science.org/doi/10.1126/sciadv.adu2241 (visited on 09/02/2025).

[74] Ellen Webb et al. "Potential hazards of air pollutant emissions from unconventional oil and natural gas operations on the respiratory health of children and infants". In: *Reviews on Environmental Health* 31.2 (June 1, 2016), pp. 225–243. ISSN: 2191-0308. DOI: 10.1515/reveh-2014-0070.

[75] Jeremy G. Weber. "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming". In: *Energy Economics* 34.5 (Sept. 1, 2012), pp. 1580–1588. ISSN: 0140-9883. DOI: 10.1016/j.eneco.2011.11.013. URL: https://www.sciencedirect.com/science/article/pii/S0140988311002878.

[76] Jeremy G. Weber. "A decade of natural gas development: The makings of a resource curse?" In: *Resource and Energy Economics* 37 (Aug. 1, 2014), pp. 168–183. ISSN: 0928-7655. DOI: 10.1016/j.reseneeco.2013.11.013. URL: https://www.sciencedirect.com/science/article/pii/S0928765513000882 (visited on 12/11/2022).

[77] Amanda L. Weinstein, Mark D. Partridge, and Alexandra Tsvetkova. "Follow the money: Aggregate, sectoral and spatial effects of an energy boom on local earnings". In: *Resources Policy* 55 (Mar. 1, 2018), pp. 196–209. ISSN: 0301-4207. DOI: 10.1016/j.resourpol.2017.11.018. URL: https://www.sciencedirect.com/science/article/pii/S0301420717302490 (visited on 09/02/2025).

[78] Mary D. Willis et al. "Unconventional natural gas development and pediatric asthma hospitalizations in Pennsylvania". In: *Environmental Research* 166 (Oct. 1, 2018), pp. 402–408. ISSN: 0013-9351. DOI: 10.1016/j.envres.2018.06.022. URL: https://www.sciencedirect.com/science/article/pii/S001393511830183X (visited on 09/02/2025).

[79] Mary D. Willis et al. "Associations between Residential Proximity to Oil and Gas Drilling and Term Birth Weight and Small-for-Gestational-Age Infants in Texas: A Difference-in-Differences Analysis". In: *Environmental Health Perspectives* 129.7 (July 2021). Publisher: Environmental Health Perspectives, p. 077002. DOI: 10.1289/EHP7678. URL: https://ehp.niehs.nih.gov/doi/full/10.1289/ehp7678 (visited on 09/02/2025).

[80] Douglas H. Wrenn, Timothy W. Kelsey, and Edward C. Jaenicke. "Resident vs. Nonresident Employment Associated with Marcellus Shale Development — Agricultural and Resource Economics Review — Cambridge Core". In: *Agricultural and Resource Economics Review* 44.2 (2015), pp. 1–19. URL: https://www.cambridge.org/core/journals/agricultural-and-resource-economics-review/article/abs/resident-vs-nonresident-employment-associated-with-marcellus-shale-development/01028163455FDCEE774C91BFFC06ECDC.

[81] Ruohao Zhang et al. *Air Quality Impacts of Shale Gas Development in Pennsylvania1*. Publisher: The University of Chicago Press. Mar. 2023. DOI: 10.1086/721430. URL: https://www.journals.uchicago.edu/doi/abs/10.1086/721430.

Appendix

Appendix A: Supporting Figures and Tables

Appendix A provides additional background materials. Figure A.1 maps the three major shale plays, Marcellus, Utica, and Devonian, that underlie PA, OH, and WV, while Table A.1 reports the principal point sources of air pollution from oil and gas production and their associated pollutants.

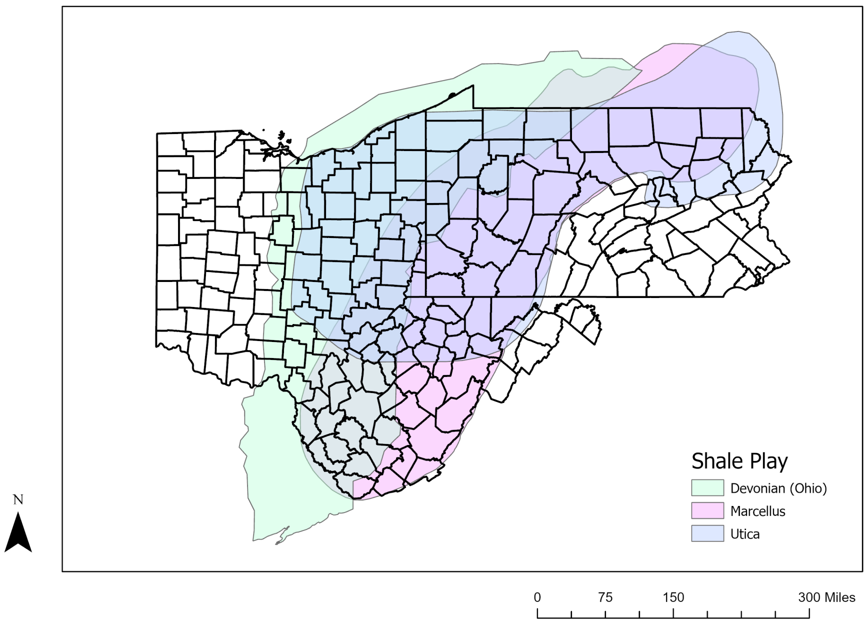

Figure A.1: Major Shale Plays in Pennsylvania, Ohio, and West Virginia
Notes: This figure shows the oil and shale gas reservoirs in PA, OH, and WV. Marcellus, Utica, and Devonian shale plays underlie the three states. The shapefiles come from the Energy Information Administration (2022) [67].

Table A1: Major point sources of air pollution from oil and gas production

	CO	NOX	VOC	Methane	$PM_{2.5}$
NAICS 211120: Crude Petroleum Extraction					
Boiler	Y	Y	Y	Y	Y
Chemical Reactor	Y	Y	Y		
Flare	Y	Y	Y	Y	Y
Gasoline Loading Rack or Arm	Y				
Other combustion	Y	Y	Y		
Process Equipment Fugitive Leaks	Y	Y			
Process Heater	Y	Y	Y	Y	Y
Reciprocating IC Engine	Y	Y	Y	Y	Y
Silo	Y				
Storage Tank	Y	Y			
Turbine	Y	Y	Y	Y	Y
Other fugitive	Y				
NAICS 211130: Natural gas Extraction					
Boiler	Y	Y	Y	Y	Y
Distillation Column/Stripper	Y	Y	Y	Y	
Flare	Y	Y	Y	Y	Y
Gasoline Loading Rack or Arm	Y				
Incinerator	Y	Y	Y	Y	
Open Air Fugitive Source	Y				
Other process equipment	Y	Y	Y	Y	Y
Oxidation Unit	Y	Y	Y	Y	
Process Equipment Fugitive Leaks	Y	Y	Y	Y	Y
Process Heater	Y	Y	Y	Y	Y
Reciprocating IC Engine	Y	Y	Y	Y	Y
Storage Tank	Y	Y	Y	Y	
Transfer Point	Y				
Turbine	Y	Y	Y	Y	Y
Other combustion	Y	Y	Y	Y	
Other evaporative sources	Y				
Other fugitive	Y	Y	Y	Y	

Notes: Table A.1 reports the air pollutants associated with each point source of oil and gas–related air pollution. A "Y" indicates that the source emits the corresponding pollutant. Data are drawn from the U.S. EPA's National Emissions Inventory [69].

Appendix B: First-stage Results

Appendix B presents first-stage regression results to demonstrate instrument relevance. Table B.1 shows the first-stage results for my county-level economic regressions, and Table B.2 reports the corresponding results for the cell-level air quality regressions. Both sets of regressions confirm that the instrumental variables satisfy the relevance condition,[30] as evidenced by strong first-stage F-statistics, a standard diagnostic for instrument strength that tests the null hypothesis that the instrument is uncorrelated with the endogenous variable.

B.1. First-stage Results: County-Level Economic Regressions

	Oil and Gas Production	
	(2002–2020 Sample)	(2001–2020 Sample)
	(1)	(2)
Predicted Oil and Gas Production	0.9780***	0.9796***
	(0.1616)	(0.1627)
Observations	3,990	4,200
R^2	0.93156	0.93001
Within R^2	0.15265	0.15386
F-test	61.249	63.116

Notes: Table B.1 reports first-stage regression results corresponding to the economic regressions in Tables 2–11. The dependent variable is county-level oil and gas production, and the instrument is predicted county-level production from equation (2). Column (1) reports first-stage results for regressions using LODES data (2002–2020), while Column (2) reports first-stage results for regressions using other economic outcomes (2001–2020). All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. Robust standard errors, clustered by county and year, are reported in parentheses. The F-statistics correspond to the excluded instrument test. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

[30] The relevance condition requires that the instrument is correlated with the endogenous variable.

B.2. First-stage Results: Cell-Level Air Quality Regressions

	Oil and Gas Production Within:						
	0-1 km	1-2 km	2-3 km	3-5 km	5-10 km	5-10 km	5-10 km
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Oil and Gas Production Within 0-1 km	0.2507***						
	(0.0141)						
Predicted Oil and Gas Production Within 1-2 km		0.2029***					
		(0.0144)					
Predicted Oil and Gas Production Within 2-3 km			0.1281***				
			(0.0136)				
Predicted Oil and Gas Production Within 3-5 km				0.1366***			
				(0.0142)			
Predicted Oil and Gas Production Within 5-10 km					0.1647***		
					(0.0146)		
Predicted Oil and Gas Production Within 10-15 km						0.1783***	
						(0.0165)	
Predicted Oil and Gas Production Within 15-20 km							0.3770***
							(0.0227)
Observations	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280
R^2	0.88747	0.92053	0.93393	0.94673	0.96003	0.96553	0.95809
Within R^2	0.22822	0.44903	0.52194	0.57350	0.62506	0.65752	0.57407
F-test	37.460	44.018	44.759	56.277	76.058	106.44	108.60

Notes: Table B.2 reports first-stage regression results corresponding to the air quality regressions in Table 12. Each column corresponds to production within a specific distance band (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10–15 km, and 15–20 km) as indicated in the column headers. The dependent variable is oil and gas production within the given band, and the instrument is predicted production from equation (4). All regressions control for the inverse hyperbolic sine of cell population and production in the four adjacent bands, and include county and year fixed effects. Robust standard errors, clustered by county and year, are reported in parentheses. F-statistics correspond to the excluded instrument test. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix C: Robustness Checks

Appendix C presents results under alternative model specifications for my main regression results, specifically those reported in Tables 3 and 12. Appendix C.1 examines the robustness of the results across specifications without covariates. Appendix C.2 reports results using a conventional log–log transformation, with a +1 adjustment to independent variables that take zero values; all dependent variables are strictly positive and thus, do not require + 1 adjustment. Appendix C.3 presents findings from ordinary least squares (OLS) regressions, as a comparison to the instrumental variable specifications.

C.1. Without Covariates

In this section, I examine the sensitivity of my main results to specifications that omit covariates. Table C.1.1 reports the main labor market outcomes from Table 3 without including county-level control variables (i.e., the inverse hyperbolic sine of population, temperature, and precipitation). These results show that controlling for covariates increases both the magnitude and statistical significance of the estimated effects, suggesting that variation in observable conditions would otherwise bias the estimates downward. Further analysis reveals that omitting population as a control accounts for most of this difference. Nevertheless, even in specifications without covariates, the results indicate that the primary beneficiaries of oil and gas production are cross-county workers, particularly those who are non-local.

Table C.1.2 reports the main air quality results from Table 12 without controlling for cell-level covariates, such as the inverse hyperbolic sine of population and production in adjacent distance bands. In these specifications, the effects of oil and gas production are detectable across the full 1–20 km range. This contrasts with the main results in Table 12, which remain robust primarily within the 1–3 km bands. The broader detectability in the no-covariate specification may capture a combination of factors: (i) omission of relevant local controls, which allows correlations with more distant production to deflate estimated effects; (ii) correlation of production across neighboring distance bands, making it difficult to isolate truly local effects; and (iii) physical dispersion of $PM_{2.5}$, whereby production in nearby bands can contribute to pollution in the focal cell. Including covariates in the main specification helps isolate the local effects and reduces attenuation from these confounding influences.

Table C.1.1. Impact of Oil and Gas on Decomposed Job and Earnings to Three Worker Groups — Without Covariates

	Local Jobs	Job Inflows	Job Outflows	Local Earnings	Earnings Inflows	Earnings Outflows
	(1)	(2)	(3)	(4)	(5)	(6)
Oil and Gas Production	-0.0069	0.0119	0.0183*	0.0032	0.0167*	0.0473***
	(0.0053)	(0.0072)	(0.0088)	(0.0078)	(0.0081)	(0.0146)
Pred. Lvl Effect / 100,000 BOE	-3.9	6.6	8.5	158K	422K	1.05M
Observations	3,990	3,990	3,990	4,200	4,200	4,200
R^2	0.99634	0.98780	0.99212	0.99415	0.99440	0.98574
Within R^2	-0.01761	-0.03002	-0.04508	0.00114	-0.10403	-0.14682

Notes: This table reports regression results evaluating the robustness of the main estimates in Table 3 without covariates. The dependent variables are the inverse hyperbolic sine of county-level jobs and earnings for three worker groups: homecounty workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work). The main explanatory variable is the inverse hyperbolic sine of county-level oil and gas production. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports predicted effects, representing the predicted level change in the dependent variable for a 100,000 BOE increase in oil and gas, computed using the estimated elasticity at the mean of the dependent variable and oil and gas. All regressions include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.1.2. Impact of Oil and Gas on Particulate Matter 2.5 — Without Covariates

	Receptor-cell $PM_{2.5}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Oil and Gas Production Within 0-1 km	0.0111						
	(0.0067)						
Predicted Oil and Gas Production Within 1-2 km		0.0071**					
		(0.0027)					
Predicted Oil and Gas Production Within 2-3 km			0.0062**				
			(0.0029)				
Predicted Oil and Gas Production Within 3-5 km				0.0049*			
				(0.0024)			
Predicted Oil and Gas Production Within 5-10 km					0.0040*		
					(0.0020)		
Predicted Oil and Gas Production Within 10-15 km						0.0039*	
						(0.0019)	
Predicted Oil and Gas Production Within 15-20 km							0.0035*
							(0.0019)
Pred. Lvl Effect / 100,000 BOE	1.126	0.236	0.122	0.030	0.005	0.003	0.002
Observations	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280
R^2	0.97766	0.97918	0.97958	0.98012	0.98054	0.98060	0.98071
Within R^2	-0.15730	-0.07879	-0.05831	-0.02990	-0.00853	-0.00523	0.00070

Notes: This table reports regression results evaluating the robustness of the main estimates in Table 12 without covariates. The dependent variable is the inverse hyperbolic sine of annual average $PM_{2.5}$ in the receptor cell. The main explanatory variable is the inverse hyperbolic sine of oil and gas production within concentric distance bands (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10–15 km, and 15–20 km) from each receptor cell. Coefficients can be interpreted as the percent change in $PM_{2.5}$ given a 1% change in oil and gas production within a distance band of the cell. All regressions include cell and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by cell and year, are shown in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.2. Log-log Specification

In this section, I present results using a conventional log–log transformation for the main regressions, with a +1 adjustment applied only to independent variables that contain zero values. The estimated effects in Tables C.2.1 and C.2.2 are broadly consistent with the main results in Tables 3 and 12 using the inverse hyperbolic sine transformation, demonstrating that the findings are robust to alternative functional forms.

Table C.2.1. Impact of Oil and Gas on Decomposed Job and Earnings to Three Worker Groups — Log-Log Specification

	Job outcomes			Earnings outcomes		
	Home-county	Inflows	Outflows	Home-county	Inflows	Outflows
	(1)	(2)	(3)	(4)	(5)	(6)
Oil and Gas Production + 1	-0.0027	0.0159*	0.0264***	0.0114	0.0226**	0.0585***
	(0.0046)	(0.0080)	(0.0086)	(0.0069)	(0.0082)	(0.0149)
Pred. Lvl Effect / 100,000 BOE	-1.5	8.7	12.2	591K	569K	1.29M
Observations	3,990	3,990	3,990	4,200	4,200	4,200
R^2	0.99678	0.98792	0.99277	0.99524	0.99471	0.98636
Within R^2	0.10689	-0.02008	0.04083	0.18643	-0.04266	-0.09654

Notes: This table reports regression results evaluating the robustness of the main estimates in Table 3 using a conventional log–log specification instead of the inverse hyperbolic sine (asinh–asinh) transformation. The dependent variables are the log of county-level jobs and earnings for three worker groups: home-county workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work). The main independent variable is the log of county-level oil and gas production + 1. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports predicted effects, representing the predicted level change in the dependent variable for a 100,000 BOE increase in oil and gas, computed using the estimated elasticity at the mean of the dependent variable and oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.2.2. Impact of Oil and Gas on Particulate Matter 2.5 — Log-Log Specification

	Receptor-cell $PM_{2.5}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Oil and Gas Production Within 0-1 km + 1	0.0435 (0.0265)						
Oil and Gas Production Within 1-2 km + 1		0.0390*** (0.0130)					
Oil and Gas Production Within 2-3 km + 1			0.0428* (0.0224)				
Oil and Gas Production Within 3-5 km + 1				0.0267 (0.0173)			
Oil and Gas Production Within 5-10 km + 1					0.0153 (0.0118)		
Oil and Gas Production Within 10-15 km + 1						0.0141 (0.0099)	
Oil and Gas Production Within 15-20 km + 1							0.0059 (0.0044)
Pred. Lvl Effect / 100,000 BOE	4.386	1.277	0.835	0.163	0.020	0.011	0.003
Observations	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280
R^2	0.93859	0.94567	0.94010	0.96441	0.97594	0.97705	0.98025
Within R^2	-2.1781	-1.8115	-2.0997	-0.84169	-0.24527	-0.18747	-0.02198

Notes: This table reports regression results evaluating the robustness of the main estimates in Table 12 using a conventional log–log specification instead of the inverse hyperbolic sine (asinh–asinh) transformation. The dependent variable is the inverse hyperbolic sine of annual average $PM_{2.5}$ in the receptor cell. The main explanatory variable is the inverse hyperbolic sine of oil and gas production + 1 within concentric distance bands (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10–15 km, and 15–20 km) from each receptor cell. Coefficients can be interpreted as the percent change in $PM_{2.5}$ given a 1% change in oil and gas production within a distance band of the cell. All regressions control for the inverse hyperbolic sine of cell population + 1 and production + 1 in the four adjacent bands, and include cell and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by cell and year, are shown in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.3. Ordinary Least Squares

In this section, I report results using OLS regressions instead of the IV approach employed in the main analysis. These specifications use the same covariates and distance-band constructions as in the primary models. Comparing OLS and IV results allows assessment of the sensitivity of the estimated effects to the identification strategy and shows how potential endogeneity in oil and gas production may bias the observed estimates.

The OLS results in this section generally yield substantially smaller estimated effects compared with the IV specifications reported in the main text. This pattern aligns with Weber (2012), who also find that wage and income effects are larger when using an IV approach compared to OLS [75]. The difference suggests that oil and gas production is likely endogenous: factors that influence production are correlated with local labor and environmental outcomes, and the IV approach corrects for this endogeneity.

In particular, Table C.3.1 shows that labor market gains for non-local workers are smaller but remain robust. In contrast, OLS estimates for local workers differ from Table 3. Specifically, commuting local workers do not experience detectable changes in total jobs or earnings, whereas home-county workers observe meaningful earnings gains. The OLS estimates for air quality in Table C.3.2 indicate that $PM_{2.5}$ effects are statistically significant only for production located beyond 5 km from the receptor cell. Nonetheless, the effect sizes remain economically negligible across all distance bands. Overall, the attenuation of OLS estimates suggests that endogeneity and omitted variable bias lead to underestimation of the true causal effects of oil and gas production on both labor market outcomes and local air quality, demonstrating the importance of the IV approach for identifying these effects.

Table C.3.1. Impact of Oil and Gas on Decomposed Job and Earnings to Three Worker Groups — OLS Estimates

	Job outcomes			Earnings outcomes		
	Home-county (1)	Inflows (2)	Outflows (3)	Home-county (4)	Inflows (5)	Outflows (6)
Oil and Gas Production	0.0014	0.0010	0.0065**	0.0051***	0.0028	0.0150***
	(0.0013)	(0.0023)	(0.0026)	(0.0017)	(0.0025)	(0.0033)
Pred. Lvl Effect / 100,000 BOE	0.8	0.6	3.0	265K	70K	331K
Observations	3,990	3,990	3,990	4,200	4,200	4,200
R^2	0.99681	0.98847	0.99342	0.99529	0.99545	0.98899
Within R^2	0.11451	0.02600	0.12727	0.19556	0.10411	0.11477

Notes: This table reports regression results evaluating the robustness of the main estimates in Table 3 using OLS instead of the IV approach. The dependent variables are the log of county-level jobs and earnings for three worker groups: home-county workers (i.e., locals working in their home county), inflow workers (i.e., locals working outside their home county), and outflow workers (i.e., non-locals commuting into the producing county to work). The main independent variable is the log of county-level oil and gas production + 1. Coefficients can be interpreted as the percent change in the outcome given a 1% change in oil and gas production. The table also reports predicted effects, representing the predicted level change in the dependent variable for a 100,000 BOE increase in oil and gas, computed using the estimated elasticity at the mean of the dependent variable and oil and gas. All regressions control for inverse hyperbolic sine–transformed county population, temperature, and precipitation, and include county and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by county and year, are reported in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C.3.2. Impact of Oil and Gas on Particulate Matter 2.5 — OLS Estimates

	Receptor-cell $PM_{2.5}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Oil and Gas Production Within 0-1 km	-6.4×10^{-5} (0.0001)						
Oil and Gas Production Within 1-2 km		6.96×10^{-5} (0.0001)					
Oil and Gas Production Within 2-3 km			5.1×10^{-5} (0.0001)				
Oil and Gas Production Within 3-5 km				0.0002 (0.0001)			
Oil and Gas Production Within 5-10 km					0.0006^{***} (0.0002)		
Oil and Gas Production Within 10-15 km						0.0005^{***} (0.0002)	
Oil and Gas Production Within 1520 km							0.0010^{**} (0.0004)
Pred. Lvl Effect / 100,000 BOE	-0.006	0.002	0.001	0.001	0.001	0.000	0.001
Observations	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280	4,728,280
R^2	0.98075	0.98080	0.98086	0.98090	0.98092	0.98092	0.98092
Within R^2	0.00279	0.00491	0.00837	0.01010	0.01138	0.01137	0.01135

This table reports regression results evaluating the robustness of the main estimates in Table 12 using OLS instead of the IV approach. The dependent variable is the inverse hyperbolic sine of annual average $PM_{2.5}$ in the receptor cell. The main explanatory variable is the inverse hyperbolic sine of oil and gas production + 1 within concentric distance bands (0–1 km, 1–2 km, 2–3 km, 3–5 km, 5–10 km, 10–15 km, and 15–20 km) from each receptor cell. Coefficients can be interpreted as the percent change in $PM_{2.5}$ given a 1% change in oil and gas production within a distance band of the cell. All regressions control for the inverse hyperbolic sine of cell population + 1 and production + 1 in the four adjacent bands, and include cell and year fixed effects. See Table 3 note for interpretation of "within R^2." Robust standard errors, clustered by cell and year, are shown in parentheses. Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.