Natural Gas Flaring, Respiratory Health, and Distributional Effects*

Wesley Blundell†

Anatolii Kokoza‡§

December 31, 2020

Abstract

Although there is strong evidence that oil and natural gas development lead to decreases in local ambient air quality, there is less evidence of a causal link between these activities and human health. We estimate the effect of flared natural gas on respiratory health by using quasi-random variation in upwind flaring generated by the interaction of wind patterns and natural gas processing capacity. We combine data on well location, flaring, weather, natural gas processing facilities, and patient-level hospital visits with the five digit zip code and health diagnostic codes for each patient in North Dakota. We estimate the causal effect of increased upwind flaring on the monthly respiratory-related hospital visitation rate for a zip code by using the number of upwind wells that are connected to a natural gas processing facility with limited processing capacity as an instrument for monthly upwind flaring. We find that (1) a 1% increase in the amount of flared natural gas in North Dakota would increase the respiratory-related hospital visitation rate by 0.0012 (0.7%), and (2) zip codes that were exposed to more than half of all flared natural gas extracted less than 20% of all resource wealth during the sample time period. These results inform current policy debates on the benefits of restrictions on natural gas flaring, the externalities associated with oil extraction, and the distribution of those externalities. (JEL Q53, Q35, Q51, I18, L71)

Keywords: flaring, air pollution, shale development, respiratory health, distributional implications.

*We thank Mark Agerton, Andy Boslett, Mario Bravo, Jed DeVaro, Tim Fitzgerald, Gautam Gowrisankaran, Elaine Hill, Ryan Kellogg, Keith Joiner, Ryan Lampe, Ashley Langer, Derek Lemoine, Greg Upton, and Tiemen Woutersen, as well as seminar participants at the Environmental Defense Fund, Rochester Department of Public Health Sciences, the 2017 Western Economic Association International Summer Conference, and the 2020 Workshop on Economic and Environmental Effects of Oil and Gas for helpful comments. Declarations of interest: none.

†Washington State University

‡USAA

§Disclaimer: The views expressed in this paper do not represent those of USAA.

1 Introduction

Since 2008 the United States has experienced tremendous growth in oil production as a result of shale development. In addition to oil, shale rock also contains natural gas and other impurities. Because it is costly to capture, transport and process, this natural gas is often burned at the well in a process called flaring. Flaring produces local air pollutants that are detrimental to local respiratory health.

Recently, state and federal policymakers have been divided over whether – and to what extent – to allow flaring. At the federal level, previous restrictions on the venting and flaring of natural gas have been rolled back (Environmental Protection Agency, 2019a). In contrast, the large quantity of flared natural gas – valued at more than $1 billion in 2014 for North Dakota – has led to public pressure and regulatory responses at the state level. Specifically, the North Dakota Industrial Commission (NDIC) passed Commission Order 24665 to reduce flaring, which resulted in a reduction in flaring (Lade and Rudik, 2019). In Texas, where the value of flared natural gas exceeds $300 million annually, the Railroad Commission[1] has instructed staff to investigate flaring at wells already connected to natural gas pipelines (Holland, 2019). These recent policy changes, as well as the continued upward trend in the level of flaring in the United States, indicate that United States (U.S.) flaring policy and practices are likely be revisited in the coming years.

In this paper, we provide evidence of a causal link between natural gas flaring and human health. We take advantage of a unique dataset on well locations, flaring, weather, natural gas processing facilities, and patient-level hospital visits with the five digit zip code and health diagnostic codes for each patient in North Dakota. Using an instrumental variables design, we estimate the impact of upwind flaring on the respiratory-related hospital visitation rate for a zip code. We find that each unit of upwind flaring significantly increases the monthly respiratory-related hospital visitation rate for zip codes up to 60 miles downwind from the source. We further find that the damages from the flaring of natural gas are not concentrated amongst the zip codes most likely to benefit from oil and natural gas activity, but are distributed across zip codes with no oil and natural gas extraction.

We estimate that a one percent increase in upwind natural gas flaring causes a 0.0012 (0.7%) increase in the downwind respiratory-related hospital visitation rate. This estimate indicates that if the 88% gas capture rate of Order 24665 had been in place prior to 2007, health costs from respiratory-related hospital visits in North Dakota would be reduced by up to $296.3 million USD (in 2018 dollars) over a nine year period. Although this magnitude appears large, we show that it is consistent with engineering estimates from the literature on the level and cost of the hazardous pollutants produced by flaring (Giwa et al. (2019), Holland et al. (2016)). Furthermore, using satellite evidence, we show that our measure of exposure to upwind flared

[1]The Texas Railroad Commission is the regulator of the oil and gas industry in the state.

natural gas is tied to measurements of local air pollutants. North Dakota is an ideal setting given the rapid expansion in flared natural gas observed in Figure 1. Our results extend to other settings where flaring is a concern. For example, the Permian Basin and Eagle Ford in Texas have a combined population three times greater in this study and, in recent years, a higher level of flaring.

Identifying a causal relationship between the flaring of natural gas and respiratory health poses several challenges. First, there could be measurement error since exposure to flared natural gas could be endogenous to an individual's avoidance behavior (Neidell (2009); Chay and Greenstone (2005)). Second, the oil activity that coincides with flaring also corresponds to the presence of other activities that impact local ambient air quality, such as vehicle traffic (Fershee, 2012). Finally, increases in oil and natural gas extraction, which are associated with higher levels of flaring, may result in the migration of younger and healthier individuals for employment opportunities to the area.

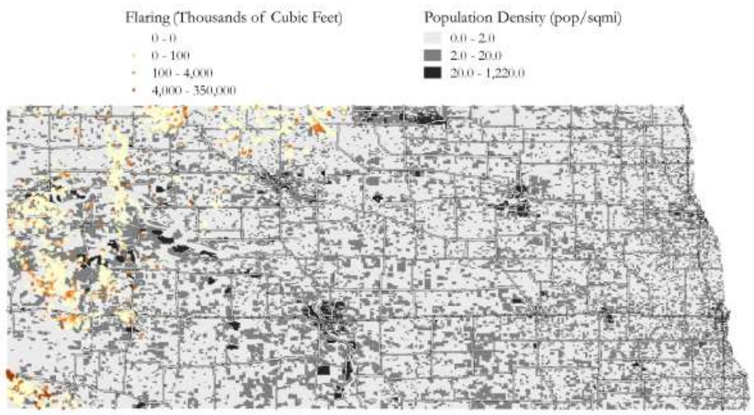
Our instrumental variables approach addresses these concerns by using the capacity at nearby gas processing plants as an instrument for flaring. If a plant is near its processing capacity, then nearby wells are more likely to flare, but it should not affect other sources of pollution that impact local health. A number of considerations point to this being a good instrument. First, once a facility hits its processing capacity, the process which determines which wells are forced to flare (gathering line pressure) is quasi-random. Second, although wind directions have a seasonal pattern, there remains substantial month to month variation and variation in the wind directions for each month across years. This variation, for the purposes of this study, is effectively random. Third, differences in the planning and construction horizons for natural gas processing and oil drilling infrastructure makes expansions in natural gas processing capacity exogenous to contemporaneous drilling and extraction activity.

Estimates of the impact of flared natural gas on respiratory health prove robust in sign and significance across a range of specifications. First, we report IV estimates based on a wide variety of controls included in different combinations and across various geographic ranges. We also report the results of further robustness and falsification tests. (1) Stronger effects for repeated patient hospital visits by long term North Dakota residents as well as infants, which addresses concern about endogenous migration (Banzhaf and Walsh, 2008). (2) To counter concerns that our instrument may correspond to contemporaneous oil activity and changes in the composition of the downwind population, we examine the outcomes of car accidents, trauma incidents, or the birth rate, and find no effect. (3) To address concerns that the instrumental variables strategy itself may be generating the apparent causal effects, we report results for two different difference-in-differences estimators based on geography and the timing of Order 24665.

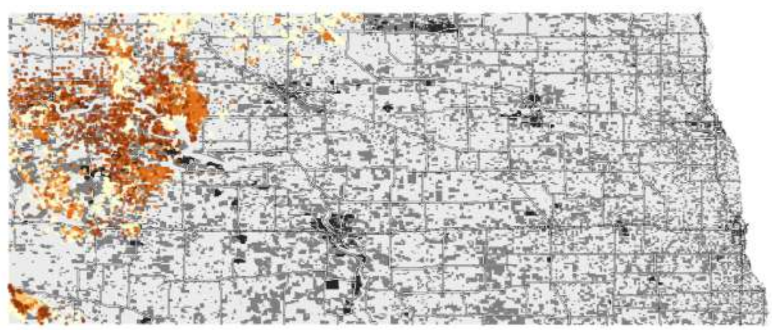
In recent years, the literature on the externalities from oil and natural gas production has been expanding. These include effects on education (Cascio and Narayan, 2015), public good provision due to the resource curse (Weber, 2014), and the impact of air, light, and water pollution (Boxall et al. (2005), Mason et al. (2015), Muehlenbachs et al. (2015), Boslett et al. (2019)). While the public health literature has considered the impact of air pollution from oil and natural gas development on health (e.g. McKenzie et al. (2012), Hill (2018), Willis et al. (2018)), to our knowledge, ours is the first study to causally identify the impact of flaring on health. In addition, our comparison of the distributional effects of local damages from flaring to the local benefits from oil royalties is similar in spirit to recent work by Hausman and Kellogg (2015), Black et al. (2018), and Bartik et al. that consider the distributional impacts of the costs and benefits of shale development. Overall, our results regarding the impact of natural gas flaring on health and the distribution of the associated damages, are timely and policy relevant in light of recent trends in the level of flared natural gas in the United States.

The paper proceeds as follows. Section 2 describes the causes of flaring, its associated pollutants, and engineering-based estimates of its health costs. Section 3 reports data sources, our definition of flaring exposure, and satellite based association between harmful pollutants and flaring. Sections 4 and 5 discuss our empirical strategy and results. Section 6 computes the health costs from flaring and examines distributional implications and section 7 concludes.

Figure 1: Increased Flaring in North Dakota



(a) Map of Flaring in North Dakota 2006



(b) Map of Flaring in North Dakota 2011

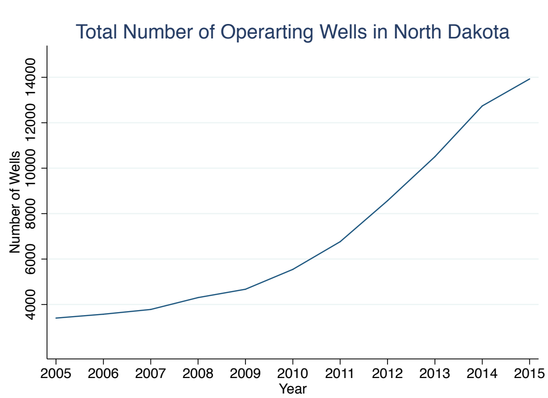
2 Background

This section provides an overview of 1) natural gas flaring and reasons why operators may choose to flare a valuable resource, 2) the pollutants associated with natural gas flaring, and 3) the potential health effects and costs from these associated pollutants. We also highlight some key strands of the economics and public health literature that motivate our study

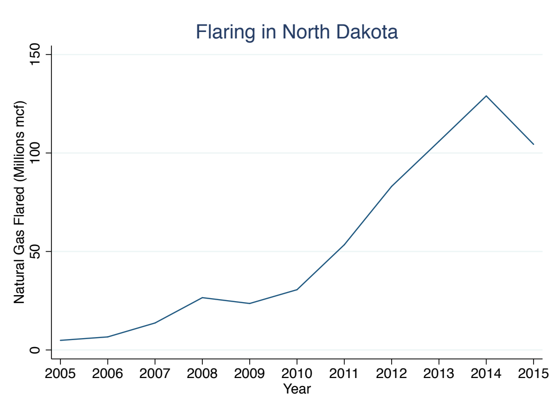
2.1 Natural Gas Flaring

Flaring is the practice of burning the unprocessed "wet" natural gas byproduct from oil-well production. This is typically done when the well operator is unable to transport the wet gas to a processing facility that separates the methane from other natural gas liquids. The increased profitability of unconventional oil development has resulted in substantial growth in oil production in North Dakota (Figure 2a), and a corresponding increase in flared natural gas (Figure 2b). However, the practice of flaring is inefficient due to the significant value of the foregone gas, which exceeded $1 billion dollars per year in North Dakota in 2014 (Fitzgerald and Stiglbauer, 2015). Critically, flaring also leads to noise, light and air pollution; the latter has the greatest impact on human health and is the central focus of this study.

Figure 2: Oil Production Activity in North Dakota 2005 - 2015



(a) Number of Wells



(b) Aggregate Flaring

There are several institutional reasons for why well operators choose to burn, rather than sell, a valuable resource. First, oil is significantly more profitable to well operators than natural gas. The value of the oil deposits far exceeds the value of the natural gas deposits within a typical North Dakota field, therefore decisions are made with respect to oil (Kellogg, 2014). Additionally, the transportation cost for unprocessed natural gas is significantly greater than that for crude oil. While oil may be transported to a refinery via a truck or pipeline, unrefined natural gas must be transported from the well to a processing facility via gas-gathering pipelines. Second, oil wells produce more gas early in the well's life cycle when a well is less likely to be connected to a gas gathering system. Thus, many operators delay or forgo the installation of systems to capture the well's natural gas.[2] Furthermore, alternatives to the use of gas gathering systems, such as re-injection or on-site electricity generation are infeasible or cost prohibitive for most wells. Finally, a lack of capacity at nearby processing facilities may force a well that is already connected to gas gathering systems to flare regardless. It is this constrained natural gas processing capacity that provides the exogenous variation in flaring that forms the basis of our analysis in section 4.

The third alternative to flaring or processing natural gas is to release it into the atmosphere without combustion – a process termed venting in the industry. However, the warming potential of methane (a major constituent of unprocessed natural gas) far exceeds that of the carbon dioxide emitted from flaring (Ford et al., 2015). The release of methane may also impact local health as in the Aliso Canyon natural gas leak (Barboza, 2015). Due to these health and environmental concerns, flaring of natural gas is preferable to venting.

In an effort to reduce natural gas flaring, several states have implemented policies targeted at oil producers. The North Dakota Industrial Commission passed Commission Order 24665 which established gas capture goals. These were initially set at 74% for October 2014 and progressively increasing to 91% for 2020. Lade and Rudik (2019) find that regulation to restrict flaring in North Dakota reduced flaring by 4-17% in the first year of operation. Nonetheless, oil well operators in North Dakota struggle to meet the standards set by the regulation. In March 2019 oil producers flared about 20% of the natural gas produced in the state – considerably above the contemporaneous limit of 12% (MacPherson, 2019).

2.2 Pollutants from Natural Gas Flaring

In addition to the lost sales revenue and increased greenhouse gas emissions, flaring is also responsible for a number of local externalities. The most visible of which is light pollution that has made the western plains of North Dakota comparable to the eastern seaboard at night (Krulwich, 2013; Boslett et al., 2019). Wang et al. (2014) also find that the combustion of natural gas creates significant noise pollution distinct from other well-site activity. Flaring also produces a number of criteria air pollutants that are known to be detrimental to human health. These include particulate matter (PM), carbon monoxide (CO), and nitrogen oxides (NO_x). It is health damages from these air pollutants that comprise the main focus of our analysis.

Although natural gas is considered a relatively clean source of energy, relative to other fossil fuels such as coal, this is so only for processed natural gas thats consists solely of methane. In contrast, unprocessed

[2] Although oil production at a well also decreases over time, the infrastructure for transporting oil from the well to processing facilities is much more flexible than for natural gas. This, combined with oil being much more valuable than natural gas, results in well operators having a strong incentive to capture all of the oil produced at the well.

natural gas includes other hydrocarbons (natural gas liquids), as well as water vapor, carbon dioxide, hydrogen sulfide, nitrogen, oxygen, and helium (EIA, 2006). Natural gas processing separates the various hydrocarbons (methane, ethane, propane, and butane) and removes the other contaminants. In contrast, flaring of unprocessed natural gas combusts the hydrocarbons and other contaminants to produce several air pollutants. Combustion of the hydrocarbons generates carbon dioxide, and incomplete combustion produces carbon monoxide and soot, which contributes to particulate matter (EPA, 1983). Combustion of hydrogen sulfide and other sulfur compounds creates sulfur dioxide, and combustion of nitrogen produces nitrogen oxides (EPA, 2018).

In order to assess the health damages caused by a pollution source, it is necessary to first understand how pollution is dispersed from that source. There is a vast discussion both within the atmospheric sciences literatures and economics literature on the dispersion of – and distribution of damages from – local air pollutants. Using integrated assessment modeling, Tong et al. (2006) demonstrate that surface level emission of NO_x can change ambient air quality between 60-120 miles downwind. Mauzerall et al. (2005) demonstrate a similar phenomenon, where air monitors up to 108 miles away from the source detect changes in ambient air quality. In the economics literature, a recent paper by Holland et al. (2016), finds that 57% of damages associated with the pollution from gasoline automobiles (NO_x, CO, and CO_2) occur outside the county in which the automobile is driven. Johnsen et al. (2019) consider a geographic range of 70 miles for assessing the damages from electricity generation. These studies consider a variety of sources, ranging from vehicles that are at ground level to coal-fired electricity generators whose flue stacks are over 500 feet tall (GAO, 2011). Natural gas flare stacks at oil wells, on the other hand, range from 30-450 feet tall (EPA, 2019b). We therefore focus our analysis on a range of 30 and 60 miles from oil wells, and consider 90- and 120-mile ranges for additional sensitivity analysis.

2.3 Air Pollution and Potential Health Costs from Flaring

There is substantial evidence within the economics and medical literatures that PM, NO_x, CO, and volatile organic compounds (VOCs) are detrimental to human health. Moretti and Neidell (2011) find significant detrimental effects on adult respiratory health from exposure to ozone, which is mainly created from the interaction of NO_x with VOCs. Currie et al. (2009) also find pollution from CO, ozone and PM10 negatively impact infant health. Currie and Walker (2011) and Anderson (2015) link vehicle emissions, which include CO, NO_x and PM, to negative health outcomes for infants and the elderly, respectively. There is also a number of papers that find exposure to particulate matter results in increased infant mortality (Chay and Greenstone, 2003; Knittel et al., 2016). Several studies have investigated the health impacts of unconventional natural gas development (Adgate et al., 2014; Werner et al., 2015). McKenzie et al. (2012), Whitworth et al. (2018), Hill (2018) and Tustin et al. (2017) investigate the impact of air pollution from unconventional natural gas development on cancer risk, infant health and respiratory health, respectively. Willis et al. (2020) investigate pediatric asthma hospitalizations in Texas due to decreased air quality from flaring and other aspects of natural gas development. However, they find "associations with flaring volumes are inconsistent." One explanation for this is that Texas data has significant discrepancies in reported flaring as compared to North Dakota, making it difficult to isolate the impact of flaring (Collins, 2018). To our knowledge, this paper provides the first estimate of the pollution and health costs from flaring separate from other sources of pollution associated with unconventional natural gas development.

To further motivate our primary analysis, we provide an estimate of the potential magnitude of the health costs from flaring in North Dakota. We combine the estimated emissions factors from Giwa et al. (2019) with the county-specific marginal damage estimates for PM, NO_x, SO_2, and VOCs from Holland et al. (2016).[3] The emissions factors are based on evidence derived from air quality measures and reported flaring levels in the Niger Delta,[4] while the pollution cost estimates are derived from various sources in the academic health literature. Specifically, we first take the emissions factor per mcf of flared natural gas for CO, PM, NO_x, SO_2, and VOCs and multiply them by the total amount flared in each county to get an estimate of the total increase in these pollutants. Next, we combine these estimated county-specific pollution increases with the county-specific marginal damages (in 2018 dollars) for each of these pollutants to get the total estimated damages from flaring in North Dakota during our sample period. Column one of Table 1 provides estimates of the total increase for each pollutant. Column two of Table 1 provides estimates of the potential dollar costs from these pollution increases, with a total estimated potential health cost of $400 million USD over our sample time period, or $44 million USD annually (in 2018 dollars).

These merely provide a rough estimate for the magnitude of the health costs from the flaring of natural gas in North Dakota. However, the combustion efficiency of a flare determines the severity of its impact on respiratory health. At 100% efficiency, all of the flared gas is converted into innocuous carbon dioxide and water. However, wells generally flare at lower efficiencies, producing carbon monoxide and particulate matter that adversely affect local respiratory health (Kleinberg, 2019). Thus, the emission factors in North Dakota likely differ from those in Nigeria. In addition, the estimates from Holland et al. (2016) provide a snap shot of the damages in 2011 and therefore fail to capture the demographic changes that occurred in North Dakota during our sample period. Nonetheless, this exercise indicates the potential health costs

[3]To obtain county-level marginal damage estimates for CO we use the scale of local damage from CO to PM found in Litman (2015).

[4]Other estimates of the emissions factors for the pollutants from flaring are from other countries or based on engineering models.

Table 1: Preliminary Estimates of the Health Costs from Flaring in North Dakota from 2007-2015

Pollutant	Amount (Tons)	Health Cost ($USD)
$PM_{2.5}$	11,404	266,606,342
NO_x	19,843	100,049,564
VOC	25,659	27,832,355
SO_2	182	5,029,539
CO	89,528	1,360,446

Notes: This tables shows engineering estimates of the amount of harmful pollutants produced by flaring (column 1) and the projected health costs of those pollutants (column 2). Marginal damage estimates for $PM_{2.5}$, NO_x, VOC, and SO_2 are from Holland et al. (2016). Marginal damage estimates for CO are from Litman (2015). Flaring from all wells in North Dakota between 2007 to 2015 is considered.

from flaring in North Dakota are significant. Additionally, Cushing et al. (2018) and Franklin et al. (2019) leverage infrared satellite data to establish the extent of flaring in the Eagle Ford Shale region of Texas. Given Texas's substantially larger population and similar level of flaring, it is reasonable to expect the magnitude of health costs from flaring is comparable to North Dakota's. Therefore, this paper's primary contribution is through the investigation of a causal relationship between the flaring of natural gas and health outcomes. The estimates provided by this analysis allow for an improved calculation of the overall cost of flaring, in terms of the increase in medical expenses. This provides valuable information for policy makers and contributes to the literature measuring the externalities associated with oil and natural gas extraction.

3 Data, Flaring Exposure, and Satellite Evidence

Estimating the effect of natural gas flaring on human health requires sufficient information on health outcomes and extraction activity near affected populations. For our analysis, we construct a novel dataset of extraction activity and individual health outcomes by combining a number of proprietary and publicly available datasets. We use monthly well-level production data to determine both how much wells produce and how much natural gas they flare. We combine this with information about nearby natural gas plant processing capacity to establish each well's ability to avoid flaring. We also use weather data to control for monthly trends in wind direction; this allows us to aggregate a household's monthly exposure to air pollution from flaring, as well as other production activities, based on the frequencies of the prevailing winds. Finally, we utilize NASA's Ozone Monitoring Instrument satellite data to examine the relationship between our measure of exposure and measurable air quality. We now explain our use of these datasets and provide a correlation between our measure of exposure to flared natural gas and air pollution measured by the Ozone Monitoring Instrument.

3.1 Data

Our well-site production and gas facility processing data comes from the North Dakota Department of Mineral Resources (DMR) from the period of January 2005 to December 2015. In addition, we obtain a panel of natural gas plant processing capacity over this time period from the North Dakota Pipeline Authority. The DMR's well data contains monthly information on oil production, natural gas production and sales, flaring, and water use for 16,906 wells, giving us a total of 986,880 well-month observations. In addition, the DMR's "Gas Plant Volumes" database contains information on the level of wet natural gas received by each natural gas processing plant for each month of our sample period. Since the majority of flared natural gas comes from connected wells, a processing plant's excess processing capacity – the difference between wet natural gas received and gas processing plant's capacity – will determine the amount of flaring that occurs at any given time. Similar to Fitzgerald and Stiglbauer (2015), we assign connected wells to gas plants based on smallest great circle distance. During the time period of January 2005 to December 2015, the number of active wells in North Dakota increased from 2,000 to 14,000 (Figure 2a). This increase in active wells corresponds to a 586% increase in annual processing capacity from 75,920,000 to 520,490,000 mcf, a 668% increase in the amount of wet natural gas received by plants from 55,367,931 to 425,022,026 mcf, and a 708% increase in the amount of dry processed natural gas sold from 45,699,000 to 369,242,000 mcf.

3.1.1 Hospital Data

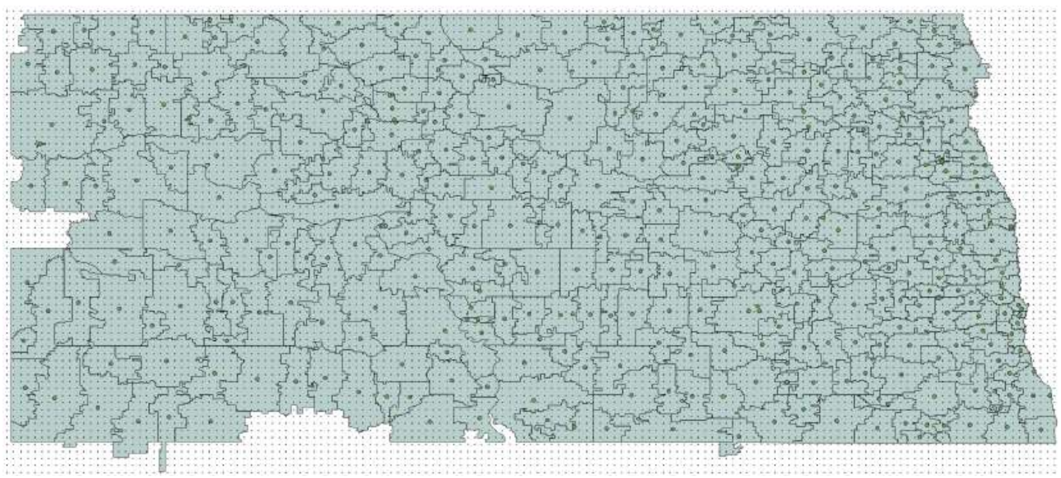
To measure local health outcomes we use patient-level hospital data from January 2007 to October 2015 at sixteen ND hospitals. These data include all hospital visits at these hospitals which comprises over 90% of hospital visits in North Dakota during the sample period. Each patient is assigned a unique identifier, allowing us to track repeat visits by the same patient. We also observe the five digit zip code of residence for each patient, as well as their age, sex, and time of visit. For each visit we observe the primary diagnosis code, which represents the cause of the visit. The data also include the outcome of each patient's visit: whether they are discharged or deceased. These data allow us to calculate hospital visits and mortalities for different diagnoses. We focus on hospital visits for specific respiratory diagnoses that we expect to be exacerbated by air pollution. In particular, we focus on ICD9 codes associated with respiratory conditions due to external agents, chemicals and fumes, pneumonia and influenza, respiratory infections, bronchitis, emphysema, and other chronic obstructive pulmonary disease.[5] We also use visits for pregnancy, vehicle trauma, and broken bones as placebo tests. We aggregate these outcomes to the zip-month level for our primary analysis.

[5]Specifically we classify these visits according to primary or additional diagnostic codes: 460 - 466, 480-488, 490 - 494, 496, 503, 504, and 506.

3.1.2 Satellite Data

To provide a measure of air-quality in each zip code we incorporate a satellite dataset on pollution levels from the Berkeley High Resolution Group (BEHR, (Laughner et al., 2018)). The use of satellite data allows us to circumvent the lack of ground-level air quality monitors noted in the literature (Krupnick and Echarte, 2017). The BEHR data is based on the NASA Ozone Monitoring Instrument (OMI) satellite data and contains NO_2 estimates in .05x.05 degree cells (about 10 square miles) for the contiguous United States. Figure 3 displays the fine level of detail provided by the satellite data; the small black dots are the OMI grid, and the larger green dots are the centroid of each zip code in North Dakota.

Figure 3: OMI grid over North Dakota



The atmospheric sciences literature indicates that this data is a good approximation for ground-level NO_2 (Bechle et al., 2013). We focus on NO_2 since NO_x, which includes NO_2, are common byproducts of natural gas combustion when the gas is flared at less than peak efficiency, which is the case at many wells (EPA, 2019b). The data presented significant computational challenges due to its 600GB size; we aggregate two measures of monthly average NO_2 readings for each North Dakota zip code from the daily data, removing any observations with error codes. The main measure is the monthly average of all readings across grid cells within a zip code. We also perform a robustness check using the monthly NO_2 reading for the single cell closest to a zip code's centroid.

One additional challenge posed by the use of this satellite data is that NO_2 measurements are presented as column densities (molecules/cm2), while ground level pollution is usually measured as parts per billion (ppb). To avoid converting the column densities to ground level units, which would require an atmospheric model and data on local topography, we follow the literature (Grainger et al., 2016) and use localized z-scores. Since we are interested in making local comparisons, observing how pollution in a given zip code compares to the surrounding region, a localized z-score provides a valid estimate of pollution for our analysis. Specifically, the z-score for zip code i in month t is defined as $z_{it} = (b_{it} - \mu_{rit})/\sigma_{rit}$, where b_{it} is the monthly average NO_2 across zip code i in month t. μ_{rit} is the average value of b_{it} across all zip codes in the county r and σ_{rit} is the standard deviation of the b_{it} values for the same county. For robustness, we also conduct a supplemental analysis using the atmospheric column densities for each zip code, b_{it}, as the pollution outcome, since these provide an adequate approximation of ground level NO_2 measurements according to the atmospheric sciences literature (Goldberg et al., 2017).

3.1.3 Other Datasets

We supplement our analysis with data from the American Community Survey (ACS) and the U.S. Energy Information Administration (EIA). We obtain zip code level demographic data for 2007, 2010, and 2013 from the ACS and interpolate between surveys to obtain measures for the intervening years. In addition, we collect monthly average oil prices for North Dakota from the EIA's First Purchase Price by Area dataset.

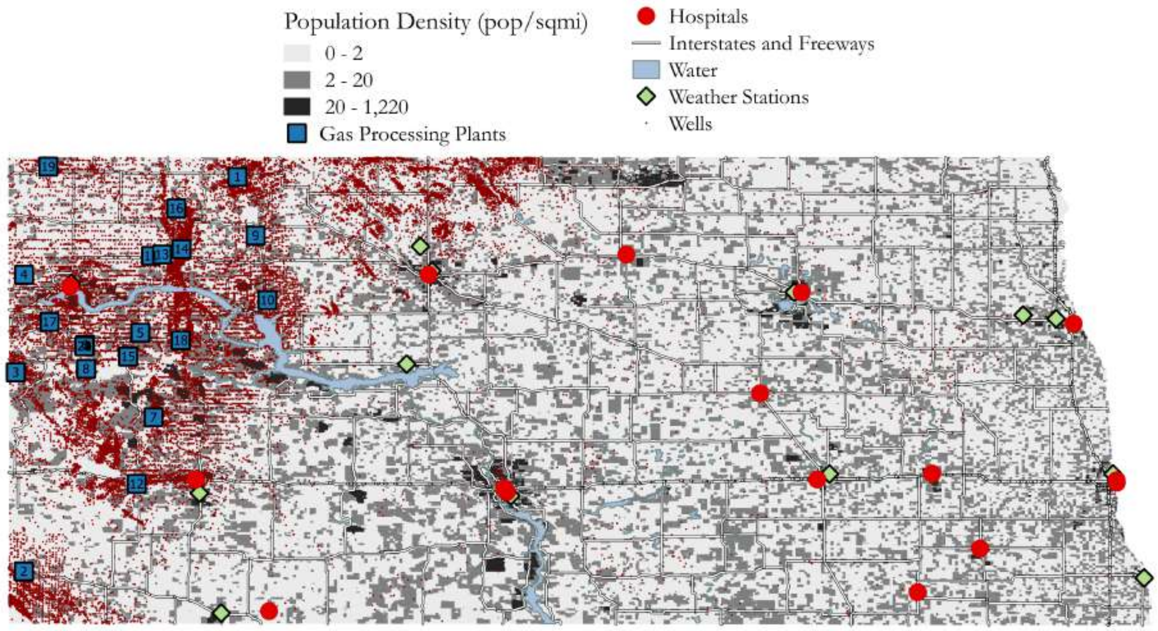
3.1.4 Weather

In order to calculate a measure of exposure to pollutants for an area, we use hourly weather data from Weather Underground from January 2005 to December 2015. These data are from thirteen stations throughout North Dakota, shown as green diamonds in Figure 4. The weather stations are at airports that tend to be in the more densely populated regions of the state. We use the hourly data for wind direction to calculate the percent of time the wind is blowing from each octant for each month. Next, we use ordinary kriging to interpolate the percent of time the wind is blowing from each octant for a grid spanning the entire state. We use this interpolated map to determine the percent of time the wind blows from each octant for every well and zip code centroid. We provide additional details as well as a discussion of seasonal wind variation in the appendix. In the following subsection we discuss how this data is combined with our oil well data to define a zip code's exposure to flared natural gas.

3.2 Measuring Exposure

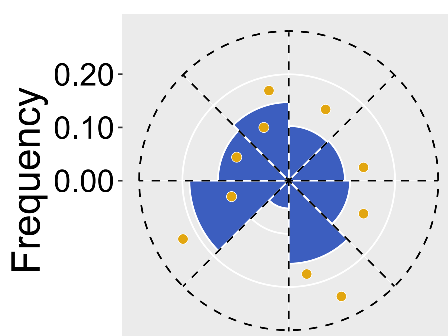
We use the kriged weather data together with well data to calculate a measure of exposure to flaring and air pollution from other well activities at the zip code level. For each zip code centroid we calculate a weighted average of the monthly upwind flaring at wells within 30, 60, 90 and 120 miles, weighted by the percent of time the wind is blowing from each octant. Figure 5 shows a simple example where the black dashed lines create the eight octants within the circle of a given radius, e.g. 30 miles. Each point represents a well within

Figure 4: Location of gas processing plants (blue squares) weather stations (green diamonds) and oil wells (small red dots) in North Dakota



the specified distance from the zip code centroid; for our measure of exposure, we consider the flaring from all wells contained in the circle. The blue shaded area represents the proportion of time (frequency) that the wind blew from each octant to the zip code centroid. The farther from the centroid that the area extends, the more often the wind was blowing from that direction in that month.

Figure 5: Measuring Exposure to Pollution Sources



We use the frequency as weights, $Wind_{iqt}$, to compute a weighted average of the zip code's total exposure to upwind flaring in a given month as follows:

$$UpwindFlaring_{it} = \sum_{q=1}^{8} Wind_{iqt} Flaring_{iqt}.$$

$Flaring_{iqt}$, is the total amount of natural gas flared from all wells in octant, q, within a radius of 30, 60, 90, or 120 miles from zip code i's centroid in month t. We create similar measures, $UpwindOil_{it}$, $UpwindDrilling_{it}$ and $UpwindWells_{it}$ for upwind oil production, wells that are being drilled, and operating wells, respectively. These variables are meant to control for the health effects of air pollution from the oil production process, the drilling of new wells and other well operations, respectively, separate from the healths effects of pollution from natural gas flaring. We also compute the total oil production, total number of wells and the total number of new wells drilled within the same radius, but without weighting by the frequencies, $Wind_{iqt}$.

3.3 Estimation Data

Table 2 presents a summary of the combined zip-month level data by exposure distance. Columns 1 and 2 present the mean and sample standard error for zip codes with at least one flaring well within 30 miles of the zip code centroid. Columns 3 and 4 correspond to zip codes with a well within 60 miles, columns 5 and 6 extend this range to 90 miles, and columns 7 and 8 extend it to 120 miles. There are 11,550 zip-month observations within 30 miles of an active well, 15,225 observations within 60 miles, 19,845 observations within 90 miles, and 23,835 observations within 120 miles. The average zip code is downwind from about 41,704 mcf of monthly flared natural gas at a range of 30 miles and 240,952 mcf when the range is extended to 120 miles. The average zip code is within range of 596 to 3,408 active wells. Each zip code has an average population between 1,459 residents and 1,816 residents depending on the distance to an active well considered. Across all geographic ranges, the average zip code's hospital visitation rate is between 0.785 and 0.828 with roughly a fifth of those visits classified as respiratory-related.[6] Overall, it is clear that we have a significant population size and number of respiratory cases with a total population of 180,000 and more than 32,000 total respiratory-related hospital visits for the 30 mile range alone. With the larger geographic ranges encompassing both a larger sample population and total respiratory visits.

To show a suggestive relationship between upwind flaring and respiratory illness, Figure 6 plots the zip code monthly average flaring exposure and respiratory-related hospitalization rate across years. There is an evident upward movement between these two series. However, this relationship is only suggestive given the evidence of possible confounding variables presented in panels B and C of Table 2.

[6] The visitation rate is one hundred times the number of hospital visits divided by the zip code's population. Thus a value of one corresponds to one percent of a zip code's population visiting a hospital every month.

Panels B and C of Table 2 break up the observations based on whether the zip-month observation has a monthly level of flaring exposure that is in the top or bottom quartile of the distribution. Focusing on observations within a range of 60 miles, we can see that patients from zip codes in the top quartile of exposure were younger and more likely to be female than patients from areas in bottom quartile of exposure. This set of healthier demographics likely explains why the high exposure zip-month observations exhibit a slightly lower rate of respiratory-related hospital visits. We do not expect areas with higher levels of exposure to have a higher mean rate of respiratory-related hospital visits if those high exposure areas have healthier patients. However, we can see that zip codes in the top quartile of exposure have slightly higher levels of atmospheric NO_2 than observations in the bottom quartile, consistent with increased pollution from flared natural gas.

In the following subsection, we further examine this relationship between flaring exposure and satellite NO_2 readings. In the following section, we present our instrumental variables strategy which addresses the apparent endogeneity between the level of flaring exposure and local demographics, as well as other concerns.

Figure 6: Flaring and Respiratory Illness

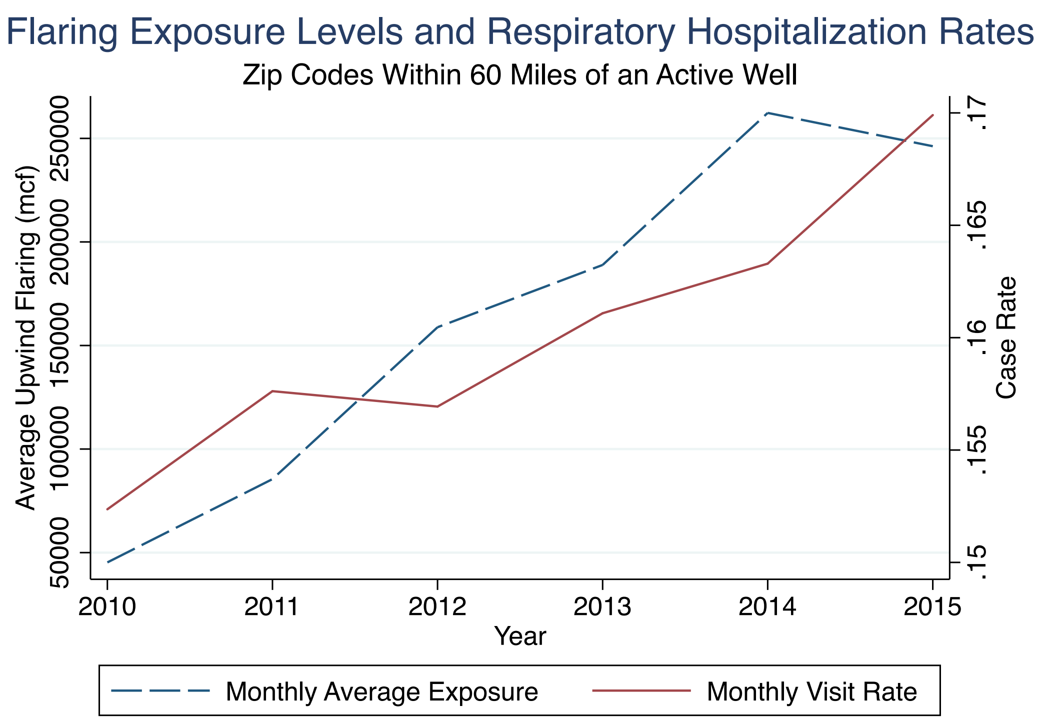


Table 2: Summary Statistics

	30 Miles		60 Miles		90 Miles		120 Miles	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Panel A: Whole Sample								
NO_2	34.77	0.0041	34.79	0.0035	34.79	0.0031	34.78	0.0028
Respiratory Visit Rate*	0.166	0.0025	0.164	0.0022	0.168	0.0021	0.173	0.0020
Hospital Visit Rate*	0.788	0.0060	0.785	0.0056	0.808	0.0054	0.828	0.0051
Upwind Flaring (mcf)†	41,704.0	924.6	121,686.1	1,854.9	191,727.1	2207.2	240,952.0	2,281.4
Upwind Constrained Wells	39.52	0.54	104.0	1.13	159.8	1.40	210.8	1.54
Upwind Wells	73.16	0.83	201.4	1.92	309.4	2.49	399.2	2.71
Upwind Oil (bbls)†	144,763.0	2,782.5	439,835.5	5,901.2	686,067.7	7,221.0	851,837.0	7,525.8
Upwind Drilled Wells†	4.632	0.091	13.49	0.19	20.78	0.22	26.09	0.23
Wells in Range	596.5	6.47	1,667.0	14.8	2,609.2	19.4	3,408.5	21.6
Oil in Range (bbls)†	115,7370.9	21,658.2	3,588,319.6	45,158.7	5,752,105.9	55,676.9	7,295,336.8	59,415.1
Drilled Wells in Range	37.10	0.71	110.4	1.46	176.3	1.75	226.5	1.86
Population	1711.3	41.3	1459.4	31.7	1816.8	35.6	1613.5	29.8
Income	5.613	0.016	5.447	0.013	5.399	0.012	5.324	0.011
Observations	11,550		15,225		19,845		23,835	
Panel B: Bottom Quartile of Exposure								
NO_2	34.83	0.0076	34.81	0.0072	34.76	0.0063	34.70	0.0060
Respiratory Visit Rate*	0.187	0.0051	0.158	0.0055	0.176	0.0052	0.190	0.0049
Hospital Visit Rate*	0.813	0.010	0.785	0.014	0.850	0.014	0.907	0.013
Avg. Patient Age	50.49	0.32	54.52	0.31	51.41	0.30	55.37	0.27
Patient is Male	0.464	0.0049	0.480	0.0055	0.465	0.0050	0.471	0.0049
Observations	2,888		3,807		4,962		5,959	
Panel C: Top Quartile of Exposure								
NO_2	34.78	0.0081	34.83	0.0074	34.87	0.0059	34.88	0.0052
Respiratory Visit Rate*	0.149	0.0043	0.149	0.0036	0.155	0.0033	0.166	0.0034
Hospital Visit Rate*	0.745	0.011	0.760	0.0099	0.776	0.0085	0.793	0.0080
Avg. Patient Age	48.15	0.37	48.50	0.33	47.83	0.28	49.14	0.25
Patient is Male	0.460	0.0063	0.474	0.0056	0.473	0.0047	0.473	0.0042
Observations	2,887		3,806		4,961		5,958	

Notes: This tables shows the summary statistics of the zip code month observations of our analysis. Columns (1) - (2) present the sample mean and standard error for zip codes that are within 30 miles of an active well at least once during the sample period. Columns (3) - (4) expand this range to 60 miles. Columns (5) - (6) expand this range to 90 miles. Columns (7) - (8) expand this range to 120 miles. *Hospital visitation rates are per capita, where a value of 1.0 corresponds to 1% of the zip code's population visiting the hospital. †Flaring and oil production measures are monthly.

3.4 Association Between Flaring and Pollution

We begin by documenting a positive correlation between monthly upwind flaring and satellite NO_2 readings, which is consistent with the atmospheric sciences and engineering literature that nitrogen oxides are a common byproduct of natural gas flaring. The use of satellite data allows us to circumvent the lack of ground-level air quality monitors noted in the literature (Krupnick and Echarte, 2017). These regressions suggest a link between the flaring of natural gas and the spread of pollutants harmful to human health.

For this analysis, we analyze two measures of a zip code's monthly NO_2 as the outcome of interest, as discussed in section 3.1.2. The first monthly measure of NO_2 from satellite readings is the local z-score z_{it}. This measure mitigates concerns regarding mismatch between atmospheric column density and ground level NO_2 readings by focusing on within-county variation in NO_2. The second is an average monthly measure of NO_2 from satellite readings across all grid cells within a zip code. Thus, we are including all of the readings from each of the cells illustrated as black dots in Figure 3. Grainger et al. (2016) find that atmospheric NO_2 readings explain a significant portion of the variation in ground level NO_2 estimates, particularly in small areas such as zip codes. This indicates that the second measure should provide a valid robustness check for our analysis.

In both cases, we estimate the following specification:

$$z_{it} = \beta_0 + \beta_1 \cdot log(UpwindFlaring_{it}) + \gamma_i + \delta_t + \varepsilon_{it},$$

where z_{it} is the measure of NO_2 pollution for zip code i in month t. $UpwindFlaring_{it}$ is the monthly flaring exposure, γ_i is the zip code fixed effect and δ_t is the month of sample fixed effect. Standard errors are clustered at the zip code level to account for zip code specific shocks.

The results are shown in Table 3. Each column corresponds to a different range, starting with zip codes within 30 miles of an active well, and extending to zip codes within 120 miles of an active well. Panel A of Table 3 presents results from regressions using the local z-score z_{it} as the outcome. Panel B shows results for the regressions using average NO_2 readings across all grid points within a zip code as the outcome. In both cases, there is a positive and statistically significant correlation between flaring 30 to 60 miles upwind and atmospheric NO_2 readings. The Panel B results in Column 3 for zip codes within 90 miles of an active well indicate a smaller, but still positive, correlation between flaring and atmospheric NO_2 that is only statistically significant at the 10% level. Once the range is extended to 120 miles in Column 4, the correlation is not statistically significant. Overall, there is a positive correlation between atmospheric NO_2 and our measure of a zip code's exposure to upwind flaring from wells within 30 to 60 miles. Furthermore, the positive correlation disappears at farther distances, consistent with findings in the atmospheric sciences literature.

As a robustness check, Panels C and D present donut specifications that consider exposure from flared natural gas from the last 15 to 30 miles of the geographic range. The natural gas flared in the inner portion of the donut is included as an additional control. Results from Column 1 indicate that the correlation between NO_2 and natural gas flared 15 to 30 miles from the zip code centroid is both statistically significant and of similar magnitude to the results in Panels A and B. Column 2 demonstrates a similar result, with the estimated correlation between atmospheric NO_2 and natural gas flared from wells 30 to 60 miles upwind from a zip code being positive and statistically significant for average zip code NO_2 readings. Finally, there is no evidence of a statistically significant relationship between average zip code NO_2 readings and flared natural gas from wells 60 to 90 or 90 to 120 miles from the zip code centroid. In the appendix we provide the same table using NO_2 readings from the centroid of each zip code, rather than average readings across zip codes, as a robustness check. The results are qualitatively similar.

Table 3: OLS Estimates of Flaring and NO_2 Readings

	(1) 30 Miles b/se	(2) 60 Miles b/se	(3) 90 Miles b/se	(4) 120 Miles b/se
Panel A: Z-Score NO_2 Reading				
Log(UpwindFlaring)	0.027***	0.021**	0.002	-0.005
	(0.009)	(0.010)	(0.012)	(0.011)
N	9,498	12,528	16,344	19,911
R^2	0.006	0.004	0.002	0.002
Panel B: Average Zip Code NO_2 Reading				
Log(UpwindFlaring)	0.018***	0.017***	0.009*	0.006
	(0.003)	(0.004)	(0.005)	(0.004)
N	9,509	12,545	16,362	19,930
R^2	0.445	0.437	0.427	0.433
	15-30 Miles	30-60 Miles	60-90 Miles	90-120 Miles
Panel C: Donut Specification Z-Score NO_2 Reading				
Log(UpwindFlaring)	0.027**	0.006	-0.012	-0.007
	(0.010)	(0.011)	(0.012)	(0.013)
N	9,498	12,528	16,344	19,911
R^2	0.006	0.005	0.003	0.002
Panel D: Donut Specification Average Zip Code NO_2 Reading				
Log(UpwindFlaring)	0.014***	0.006*	-0.003	0.001
	(0.003)	(0.004)	(0.004)	(0.004)
N	9,509	12,545	16,362	19,930
R^2	0.445	0.439	0.430	0.434
Zip Code Fixed Effects	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y

Notes: This table reports OLS regressions of upwind natural gas flaring on both the zip code NO_2 z-score and average monthly atmospheric column density NO_2 readings (molecules/cm2) for zip codes in North Dakota from 2007 - 2015. Standard errors are clustered at the zip code level. The panel is not balanced due to gaps in satellite coverage during the sample period. Panels C and D include the log of the total upwind natural gas flared from wells within the inner portion of the donut as an additional control. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level. **Sample:** Column (1) considers monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Column (2) considers monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period. Column (3) considers monthly observations from 189 zip codes that were within 90 miles of at least one active well during the sample period. Column (4) considers monthly observations from 227 zip codes that were within 120 miles of at least one active well during the sample period.

4 Methodological Challenges and our Empirical Approach

The goal of this study is to determine the impact of natural gas flaring on individuals' health. However, there are a number of challenges associated with estimating the causal impact of flaring on human health.

First, there is an issue of omitted variable bias as many other activities associated with natural gas flaring can negatively impact ambient air quality. For instance, unconventional oil and natural gas development (UONGD) can lead to an increase in vehicle traffic (Graham et al., 2015) and its corresponding pollutants. In addition, activities such as drilling or accidents such as methane leaks could also worsen air quality (Hausman and Muehlenbachs, 2019).

Second, there may be omitted variable bias related to changes in the population that coincide with natural gas flaring. Job opportunities from UONGD could incentivize younger, healthier people to migrate to areas with increased natural gas flaring (Cascio and Narayan, 2015). In addition, royalty payments or increased incomes from UONGD could also induce people to spend more on their health.

A third concern arises about the potential endogeneity in individuals' exposure to the pollution from flared natural gas. People who are the most susceptible to the pollution may engage in avoidance behavior to mitigate their level of exposure (Neidell, 2009). This may include staying indoors or even migrating away from areas with high levels of pollution (Banzhaf and Walsh, 2008).

The following sections elaborate on how our main specification deals with these challenges in three ways: (i) the use of natural gas processing capacity as an instrument for flaring, (ii) the use of a wide variety of controls related to UONGD and demographics, (iii) estimation subsamples restricted to individuals who never migrated or who are infants.

4.1 Empirical Approach

To address the challenges in the previous section we use 2SLS and estimate the following:

$$log(UpwindFlaring_{it}) = \alpha_0 + \alpha_1 \cdot log(ConstrainedWells_{it}) + X'_{it}\alpha_2 + W'_{it}\alpha_3 + \theta_i + \phi_t + \eta_{it}, \quad (1)$$

$$Health_{it} = \beta_0 + \beta_1 \cdot log(UpwindFlaring_{it}) + X'_{it}\beta_2 + W'_{it}\beta_3 + \gamma_i + \delta_t + \varepsilon_{it}. \quad (2)$$

Equation 2 is the second-stage and equation 1 is the first-stage regression. The dependent variable $Health_{it}$ is the rate of hospital visits in zip code i in month t for a given condition, e.g. respiratory-related visits. Focusing on the rate of respiratory-related hospital visits has several advantages. By using the per capita rate we do not have to make any assumptions about zero valued observations. In addition, by focusing on a broad range of respiratory-related ICD9 codes, the data provide sufficient variation for identification.

As we discussed in section 3.4, our main independent variable of interest is a zip code's exposure to upwind flared natural gas. However, there are a number of other pollution generating activities that are associated with higher levels of flaring, such as well drilling and vehicle traffic. Regarding these confounding sources of pollution, the vector X'_{it} includes a set of controls related to oil production and drilling to capture trends in shale development and the resulting co-pollutants over time. In particular, we include both the number of wells being drilled upwind, the total amount of oil produced upwind, and the number of upwind wells within the specified geographic range. We also include the total number of wells being drilled, oil produced, and active wells within range of a zip code.

With respect to increased incomes from UONGD, it is possible that individuals may utilize significantly more health care after their incomes increase either from royalties or improved employment opportunities. Ignoring this could lead us to under-estimate the impact of natural gas flaring on health. The vector W'_{it} includes controls for average income in the zip code as well as housing values, which can be a proxy for royalties (Muehlenbachs et al., 2015).

We also include zip code and month fixed effects: θ_i and γ_i, and ϕ_t and δ_t. The zip code fixed effects control for time invariant characteristics within a zip code and the month fixed effects capture seasonal trends that affect health and upwind flaring. Finally, ε_{it} and η_{it} are the error terms that include all of the other health and upwind flaring shocks in zip code i, month t that we have not controlled for.

4.2 First Stage Specification: Constrained Wells Upwind as an Instrument

Our coefficient of interest is β_1, the coefficient on our measure of a zip code's exposure to upwind natural gas flaring. As discussed in the previous sections there are a number of challenges in using observational data to estimate a causal relationship between flaring and human health. Our ideal experiment would randomly assign different levels of flaring to different areas in order to be able to separate the effect of flaring from other aspects of oil production. Of course this is not feasible, so we use an instrumental variables approach.

To instrument for $log(UpwindFlaring_{it})$ in equation 1, we use a weighted sum of the number of wells whose nearest gas processing plant is near its processing capacity:

$$Log(ConstrainedWells_{it}) = Log(\sum_{q=1}^{8} ConstrainedWells_{iqt} \cdot Wind_{iqt}). \tag{3}$$

$ConstrainedWells_{iqt}$ is the number of wells in octant q within a radius of zip code i's centroid in month t whose nearest gas processing plant is above 60% of its processing capacity.[7] $Wind_{iqt}$ is the percent of time

[7] As an alternative specification, in the appendix we raise the threshold to whether the gas processing plant is above 70% of its processing capacity.

in month t that the wind in zip code i's centroid is blowing from octant q, as shown in Figure 5.

The influence of natural gas processing capacity on flaring is well recognized. To quote Dave (2009), "Even after the oil and gas well has been connected to a gas-gathering pipeline and processing facility, there are still times when the well may flare intermittently, this is generally due to excess line pressures" from a lack of processing capacity at the facility to which the well is connected. During our sample period, the majority of the natural gas flared came from wells connected to a natural gas processing facility. In addition, wells that were connected to a facility that we define as constrained had significantly higher levels of flared natural gas and were 30% more likely to flare all of the gas they produced. There are a number of characteristics that make the number of upwind wells connected to a constrained natural gas processing plant valuable for our analysis.

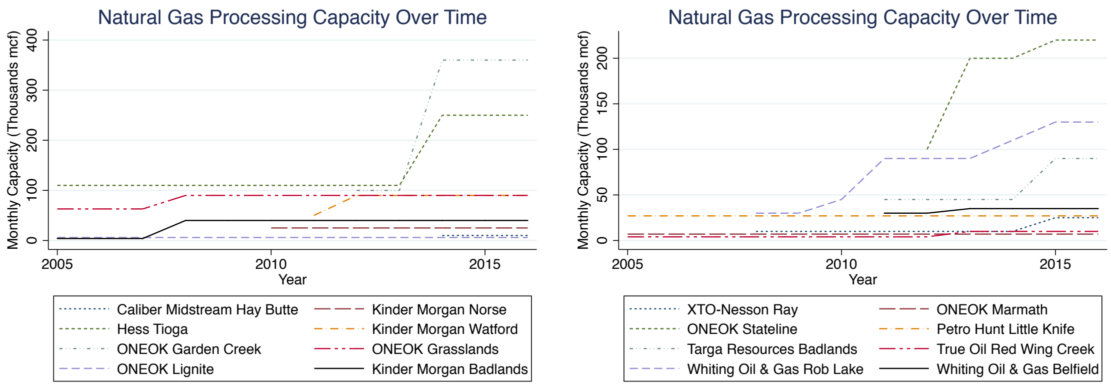
First, we argue that for a given constrained processing plant, the determination of which wells are forced to flare is quasi-random. The line pressure that forces a well to flare is determined by a combination of whether there is pooled natural gas liquid in the portion of the line near the well, a function of when the well operator last cleared the line, and the productivity or pressure from other wells connected to the constrained processing facility.

Second, there is significant variation in the maximum capacity of natural gas processing facilities over time. Figures 7a and 7b demonstrate how many natural gas processing facilities either opened or expanded their capacity during our sample period. It is important to note that the timing of these openings and expansions should be random conditional on the overall level of drilling activity. This is because drilling decisions are made with respect to oil because "producers see the associated gas as a less valuable byproduct of crude oil production, which can be sacrificed" (Ehrman, 2014). Furthermore, the planning and timing of these facilities is significantly longer than that of wells. For example, the Hess Tioga expansion completed in 2014 was originally announced in 2011 OGJ (2011). This makes it unlikely that the owners of processing facilities would time the expansion of processing capacity with the production of new wells, which are typically drilled in two to three months.

Third, our instrument for flaring exposure is a function of the number of these wells connected to a constrained plant that are *upwind* from a zip code. North Dakota is subject to significant seasonal variation in wind direction; the wind blows predominantly from the west to north and from the south during summer (Enz, 2003). Any such seasonality in weather patterns is of course captured by our month fixed effects, making the identifying variation in our instrument the variation in wind direction within a month interacted with the expansions in processing capacity. Figure 8 demonstrates that there is substantial variation in wind direction within a month across years.

The validity of our instrument requires that, after controlling for UONGD activities and demographics,

Figure 7: Natural Gas Processing Capacity Over Time

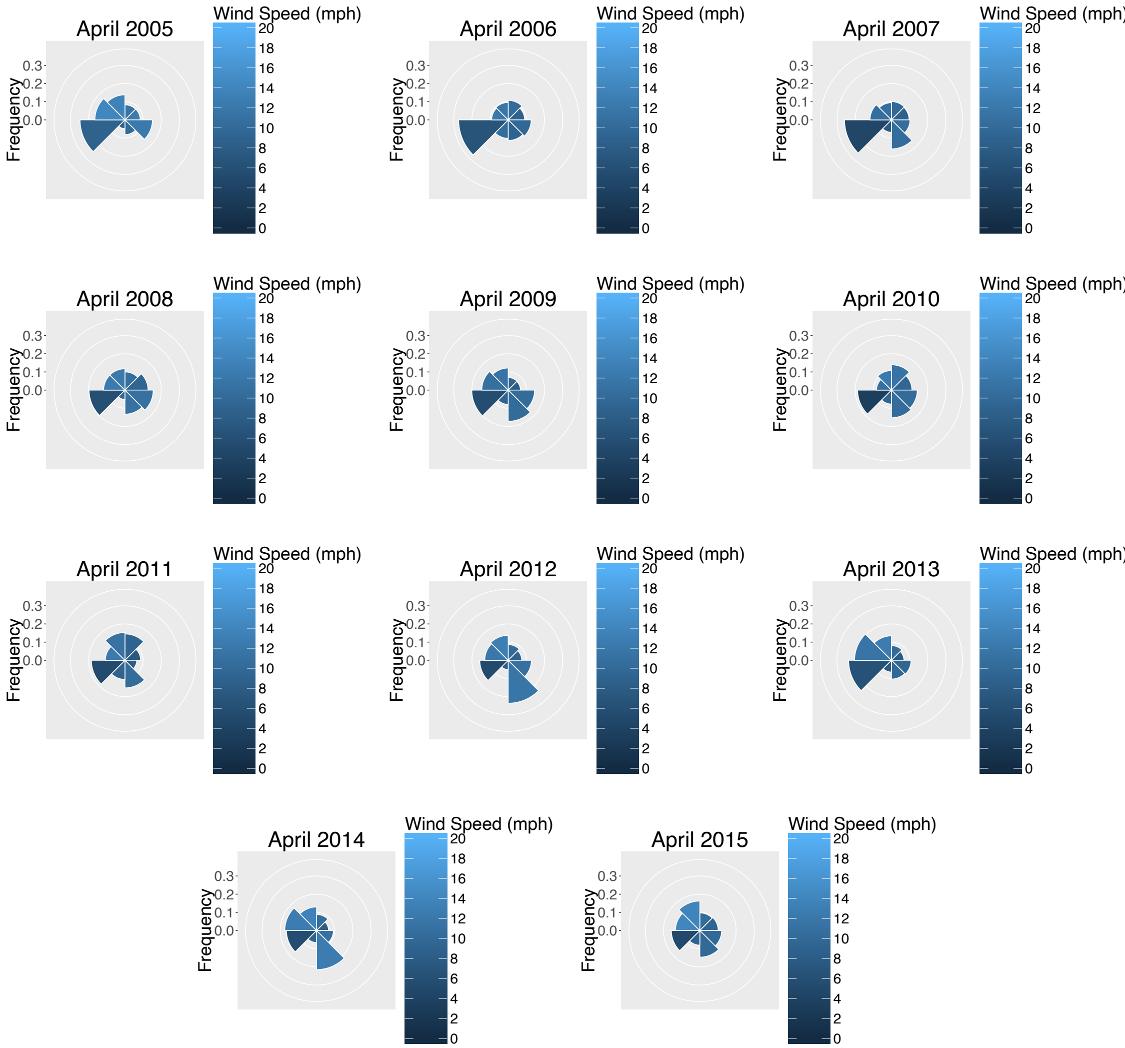


(a) Processing Capacity Plants 1 - 8

(b) Processing Capacity Plants 9 - 16

the only link between the number of constrained upwind wells and a zip code's respiratory health is via flared natural gas. After controlling for the overall level of UONGD activity in the region, the variation in wind direction interacted with the expansion of natural gas processing capacity and the manner in which connected wells are forced to flare should be quasi-random. Given our rich set of controls for upwind production, total production in the zip code's geographic region, and other zip code characteristics, it is difficult to think of any additional link to either co-pollutants from oil production activities or changes in zip code demographics. Therefore, we consider monthly variation in the number of upwind wells connected to a constrained natural gas processing facility as plausibly exogenous to other short-run determinants of respiratory health in a zip code.

Figure 8: Frequency of time the wind blew toward† and the average speed in each octant for April 2005-2015



†Note that this is reversed from the standard convention of the direction from which the wind is blowing.

5 Results

5.1 OLS Results

Table 4 presents the OLS results from estimating equation (2). The four columns correspond to estimation for zip codes within the specified distance from an active well during the sample period. The four geographic ranges considered are circles with a radius of 30, 60, 90, and 120 miles. Focusing on zip codes within 30 miles of an active well in Column (1), the coefficient of interest is the log of upwind flared natural gas from a zip code with a value of 0.0019. The interpretation of this coefficient is as follows: other things equal - a 1% increase in the amount of flared natural gas exposure will result in 0.000019 additional monthly respiratory hospital visits per capita. In all four specifications the coefficient is positive but not statistically significant. We attribute this lack of statistical significance to the presence of endogeneity as discussed in section 4. Specifically, the presence of younger, healthier individuals in areas with higher levels of flaring, noted in Panel C of Table 2, likely biases us from finding positive correlation in this OLS analysis. Overall, this preliminary OLS analysis necessitates our instrumental variables research design estimated in the following subsection.

Table 4: OLS Estimates of Impact on Respiratory Health

	(1) 30 Miles b/se	(2) 60 Miles b/se	(3) 90 Miles b/se	(4) 120 Miles b/se
Panel A: OLS Analysis				
Log(UpwindFlaring)	0.0019	0.0012	0.0005	0.0022
	(0.002)	(0.002)	(0.002)	(0.002)
Zip Code Fixed Effects	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y
Additional Controls	-	-	-	-
N	11,550	15,225	19,845	23,835

Notes: This table reports OLS regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2007 - 2015. Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level, and + at the 15% level.
Sample: Column (1) considers monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Column (2) considers monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period. Column (3) considers monthly observations from 189 zip codes that were within 90 miles of at least one active well during the sample period. Column (4) considers monthly observations from 227 zip codes that were within 120 miles of at least one active well during the sample period.

5.2 Instrumental Variables Analysis

5.2.1 First Stage Results

Panel A of Table 5 reports the first stage estimates, shown in equation (1) for the geographic ranges of 30, 60, 90, and 120 miles. The impact of the number of wells connected to a constrained processing facility on the amount of natural gas flared is statistically significant for all geographic distances with the expected sign. This confirms our understanding of natural gas flaring in North Dakota: the majority of flared natural gas comes from wells that are connected to a processing facility that is capacity constrained. The estimates in columns (1) to (4) of Table 5 indicate that a 1% increase in the number of upwind constrained wells from a zip code will result in a 0.338% - 0.398% increase in flared natural gas exposure. Across all four specifications, the first stage F statistic exceeds 10, which indicates that the number of wells connected to a constrained natural gas processing facility meets the criteria for a strong instrument outlined in Staiger and Stock (1994).

5.2.2 Second Stage Results

Panel B of Table 5 reports the second stage results. Column (1) reports the estimated effect of exposure to flaring from upwind wells within 30 miles on the zip code's respiratory-related hospital visitation rate. The estimated coefficient indicates that a 1% increase in exposure to flared natural gas will correspond to a 0.0001 unit (0.0006%) increase in respiratory-related hospital visitation rate for a zip code. Column (2) reports a similar estimate for the effect of exposure to flaring from upwind wells within 60 miles of a zip code. In both of these cases, the results are statistically significant and differ from the OLS estimates. The IV estimates are an order of magnitude greater than the OLS estimates which is consistent with two of our endogeneity concerns in the OLS estimation: individual avoidance behavior (Neidell, 2009) and the migration of younger and healthier individuals to areas with significant oil activity (Cascio and Narayan, 2015).

Columns (3) and (4) of Panel B in Table 5 report the second stage estimates of the effect of exposure to flaring from 90 and 120 miles from a zip code. In both specifications, the estimated effect of upwind flaring is statistically indistinguishable from zero. This finding is consistent with Johnsen et al. (2019) who found the pollutants associated with natural gas flaring, but in the context of coal-fired generators, do not travel beyond 70 miles. Due to the lower height of natural gas flare stacks relative to the flue stacks of coal-fired generators, we expect pollutants from natural gas flaring to not travel as far (GAO, 2011; EPA, 2019b). Therefore, we focus on results for flaring within 30 and 60 miles of a zip code for the rest of this analysis.

Table 6 presents the IV results at the geographic ranges of 30 and 60 miles with an extensive set of control variables. Columns (1) and (4) add the number of currently active wells in the range to the more

Table 5: IV Estimates of Impact on Respiratory Health for All Ranges

	(1) 30 Miles b/se	(2) 60 Miles b/se	(3) 90 Miles b/se	(4) 120 Miles b/se
Panel A: First Stage				
Log(ConstrainedWells)	0.381***	0.338***	0.398***	0.393***
	(0.052)	(0.044)	(0.045)	(0.040)
Panel B: Second Stage				
Log(UpwindFlaring)	0.010*	0.011*	0.002	-0.004
	(0.006)	(0.006)	(0.006)	(0.005)
Zip Code Fixed Effects	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y
Additional controls	-	-	-	-
N	11,550	15,225	19,845	23,835

Notes: This table reports IV regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2007 - 2015. Panel A corresponds to the first stage of the estimation, which uses the log of the number of wells connected to a constrained processing plant upwind from a zip code as an instrument for that zip codes exposure to flared natural gas. Panel B corresponds to the second stage estimates of the impact of flared natural gas on the rate of respiratory-related hospital visits in a zip code. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. **Sample:** Column (1) considers monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Column (2) considers monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period. Column (3) considers monthly observations from 189 zip codes that were within 90 miles of at least one active well during the sample period. Column (4) considers monthly observations from 227 zip codes that were within 120 miles of at least one active well during the sample period.

parsimonious specification in Table 5. The parameter estimates are larger than the results in Table 5, but the differences are not statistically significant at the 5% level. Columns (2) and (5) also control for the number of active wells, oil extracted, and wells drilled both upwind and the total number of wells drilled within range of the zip code. These additional controls account for all pollution sources from oil development activity near the zip code. Finally, columns (3) and (6) of Table 6, which is our preferred specifications, also control for demographics including the average house size, home value, and household income. The Kleibergen-Paap first stage Wald F statistics presented in Table 6 indicate that despite the inclusion of these additional controls, the instrument still meets the criteria for being a strong instrument (Staiger and Stock, 1994). In addition, the relative stability of our coefficient of interest across the specifications in Table 6 provides evidence that we have sufficiently controlled for economic and demographic factors that could impact the respiratory health for a zip code.

Overall, the instrumental variables results in Tables 5 and 6 indicate that upwind flared natural gas has a significant effect on the respiratory health of individuals who reside within 60 miles of wells. In the following subsection we explore the robustness of these results for sub-sample populations and perform a number of falsification tests that examine the validity of our research design.

Table 6: Primary IV Results of Impact on Respiratory Health

	30 Miles			60 Miles		
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Panel A: First Stage						
Log(ConstrainedWells)	0.122***	0.102***	0.099***	0.065***	0.058***	0.059***
	(0.024)	(0.018)	(0.017)	(0.019)	(0.015)	(0.015)
Panel B: Second Stage						
Log(UpwindFlaring)	0.031	0.049+	0.052*	0.076*	0.111**	0.117**
	(0.025)	(0.031)	(0.031)	(0.045)	(0.052)	(0.052)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	-	Y	Y	-	Y	Y
Demographic Controls	-	-	Y	-	-	Y
K.P. Wald F Statistic	26.443	31.803	32.231	11.973	14.480	15.178
N	11,550	11,550	11,550	15,225	15,225	15,225

Notes: This table reports IV regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2007 - 2015. Panel A corresponds to the first stage of the estimation, which uses the log of the number of wells connected to a constrained processing plant upwind from a zip code as an instrument for that zip codes exposure to flared natural gas. K.P. refers to the Kleibergen-Paap first stage Wald F statistic. Panel B corresponds to the second stage estimates of the impact of flared natural gas on the rate of respiratory-related hospital visits in a zip code. Columns (1) and (4) add controls for the number of active wells within the specified range to capture the trend in shale development over time. Columns (2) and (5) also add total oil extracted upwind, oil extracted within range, the number of upwind wells drilled, total wells drilled within the specified range, and total upwind wells of the zip code as additional controls. Columns (3) and (6) further add controls related to the average house size, value, and income for the zip code as demographic controls. Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level, and + at the 15% level. **Sample:** Columns (1) - (3) consider monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Columns (4) - (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

5.3 Robustness and Falsification Tests

The estimated effects of exposure to flared natural gas on respiratory health are conditional on choices about the impacted population and the appropriate medical outcome of consideration. In this section, we explore the estimates' sensitivity to these sample choices and conduct a series of falsification exercises. The purpose of these exercises is to determine if our primary results are driven by: (i) residential sorting from UONGD; (ii) our definition of exposure to flared natural gas being a proxy for other sources of air pollution; or, as intended, (iii) our research design.

5.3.1 Robustness

We estimate equation 2 for a number of alternative sub-populations and outcomes in order to examine whether there is something specific about our results for respiratory-related hospital visits from the general population of a zip code. Columns (1) and (4) of Table 7 report the results of our analysis when restricting the sample to patients who are long-term residents – defined as being observed prior to 2010 and after. The impact of exposure to flared natural gas is positive and statistically significant at both geographic ranges. Point estimates are smaller than in our primary specifications, which is consistent with attenuation due to the subset of long-term residents being smaller than the general population. These results indicate that our results are not explained by residential sorting; residents who lived in North Dakota prior to the oil boom were still susceptible to health impacts from exposure to flared natural gas.

Next, we explore the sensitivity of our results to the choice of respiratory-related hospital visits as the outcome variable of interest. Previous studies have found an association between air pollution and a number of outcomes beyond respiratory health. These include infant mortality, cardiovascular health, and other cancers (Schlenker and Walker (2015), Currie et al. (2009)). Given the rarity of these outcomes and the limited size of our population, we instead focus on the aggregate rate of hospital visits with the understanding that it will reflect the accumulation of these additional health outcomes from flared natural gas. Columns (2) and (5) report the results of our preferred specifications when considering the rate of all hospital visits from a zip code. The effects are positive and statistically significant for both geographic ranges. This suggests that exposure to flared natural gas leads to a general decrease in health outcomes and that our primary results are not entirely explained by the choice to focus on respiratory-related outcomes.

Finally, we examine the respiratory health of infants, which has been studied before in the context of UONGD (Hill, 2018). Focusing on infant health has a number of advantages in terms of identification. First, there is less of a concern about sorting and avoidance behavior by infants due to their inherent immobility. Furthermore, mothers with respiratory health complications, whose children are more likely to be sensitive to air pollution, are prone to relocate away from areas with unconventional oil and natural gas development. Additionally, it is unlikely that unobserved variation in mothers' socio-economic characteristics is tied to within month variation in the number of upwind wells connected to a capacity constrained natural gas processing plant. Columns (3) and (6) of Table 7 report the results of our preferred specifications when considering the percentage of all respiratory-related infant hospital visits (those with a primary Major Diagnostic Category code of four) from a zip code. The expansion of the definition of respiratory-related hospital visits in these specifications is due to the fact that, unlike with the general population, we do not need to distinguish between short term and long term respiratory-related issues for infants. The ICD9 codes used in our primary analysis better ensure the respiratory-related visit observed was due to an external agent and not an issue that developed internally over a long period of time. We believe it is not necessary to make this distinction for infants since such concerns do not apply. Both point estimates are positive and statistically significant, providing additional evidence that our primary results are not explained by unobserved demographic changes or residential sorting.

Table 7: IV Results: Robustness

	30 Miles			60 Miles		
	(1) Long-term b/se	(2) All Visits b/se	(3) Infants b/se	(4) Long-term b/se	(5) All Visits b/se	(6) Infants b/se
Panel A: Second Stage						
Log(UpwindFlaring)	0.026*	0.191***	0.017**	0.042*	0.274**	0.030*
	(0.014)	(0.072)	(0.008)	(0.024)	(0.117)	(0.017)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	Y	Y	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y	Y	Y
N	11,550	11,550	11,550	15,225	15,225	15,225

Notes: This table reports IV regressions of upwind natural gas flaring on various outcomes to examine the robustness of our main results. Columns (1) and (4) show results for the rate of respiratory-related hospital visits per zip code by long term residents in North Dakota. Columns (2) and (5) show results for the hospital visit rate per zip code. Columns (3) and (6) show the percentage of all respiratory-related infant hospital visits per zip code. All specifications include both the number of active wells, oil extracted, and wells drilled both upwind and total within the specified range of the zip code, along with average house size, value, and income as additional controls. Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. **Sample:** Columns (1) to (3) consider monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Columns (4) to (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

5.3.2 Sensitivity Analysis and Falsification Tests

Identification in this study hinges on the assumption that, conditional on controlling for the local oil related economic activity and demographics, the number of upwind wells from a zip code that are connected to a constrained facility is as good as randomly assigned. There are two ways in which this assumption could fail. One is if there are still changes in population demographics that correspond with the variation in the number of constrained upwind wells. In practice, this would result in our estimation strategy predicting health outcomes associated with different population groups that are unrelated to air pollution. The second is if the number of upwind wells connected to a constrained processing facility, conditional on our controls for oil development, still predict increases in other sources of air pollution. Previous work has indicated that oil development in North Dakota has led to increases in vehicle traffic (Fershee, 2012) and thus decreasing local ambient air quality (Knittel et al., 2016). In this case, it is important to examine our primary result with a robust set of controls regarding local UONGD, as well as to test our empirical framework with an outcome associated with non-flaring sources of emissions. Table 8 thus estimates relationships between outcomes related to changes in demographics and non-flaring sources of air pollution to exposure to flared natural gas.

Table 8 presents 2SLS results for estimating equation 2 with outcomes that should be unrelated to exposure to flared natural gas if our research design is valid. Columns (1) and (4) estimate our preferred 2SLS specification in which the dependent variable is a zip code's vehicle-related hospital visitation rate. The estimates for both geographic ranges are not statistically significant at conventional levels. This indicates

Table 8: IV Results: Falsification Tests

	30 Miles			60 Miles		
	(1) Vehicle Visit Rate	(2) Pregnancy Visit Rate	(3) Trauma Visit Rate	(4) Vehicle Visit Rate	(5) Pregnancy Visit Rate	(6) Trauma Visit Rate
Panel A: Second Stage						
Log(UpwindFlaring)	0.001	0.023	0.002	0.001	0.016	0.007
	(0.001)	(0.017)	(0.003)	(0.001)	(0.031)	(0.005)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	Y	Y	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y	Y	Y
N	11,550	11,550	11,550	15,225	15,225	15,225

Notes: This table reports IV results of various falsification tests of our primary analysis using alternative outcomes. Columns (1) and (4) show results for the rate of vehicle related hospital visits per zip code. Columns (2) and (5) show results for the number of pregnancy related hospital visits per zip code. Columns (3) and (6) show results for the rate of trauma related hospital visits per zip code. All specifications include both the number of active wells, oil extracted, and wells drilled both upwind and total within the specified range of the zip code, along with average house size, value, and income as additional controls. Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. **Sample:** Columns (1) to (3) consider monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Columns (4) to (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

that there is no evidence that our measure of exposure to flared natural gas corresponds to increased vehicle traffic. Columns (2) and (5) report estimates from our preferred specification using a zip code's rate of pregnancy related hospital visits as the outcome variable. The value of examining this outcome is that any increases would be associated with a younger and healthier population, those of child bearing age, moving to where there are high levels of oil related activity. The coefficients for this test are statistically insignificant, which provides evidence that our research design is not picking up any demographic changes in the population. Columns (3) and (6) report point estimates using the rate of trauma related hospital visits for a zip code, a health outcome that should be completely unrelated to exposure to flared natural gas, but would correspond to a trend toward riskier activities in the population. In both cases, we fail to reject the null hypothesis that our research design is not picking up other health trends unrelated to the air pollution from flared natural gas.

5.3.3 Additional Analaysis

We provide additional results regarding falsification tests and the robustness of our main result in the appendix. These include (i) testing the statistical significance of the main IV results using heteroskedastic consistent standard errors and standard errors clustered at the county level; (ii) testing for whether the respiratory health effect is the result of cumulative exposure to flared natural gas or contemporaneous exposure to flared natural gas; (iii) testing for the consistency of the main result with both leads and lags of monthly exposure to flared natural gas; (iv) testing the sensitivity of the main results to the definition of a constrained natural gas plant; (v) examining the importance of being downwind from newly producing wells, (vi) testing the importance of the functional form by estimating a log-log specification; and (vii) estimating two different difference-in-differences estimators based on geography and the timing of Order 24665. These analyses confirm both the nature of our primary result and the validity of our research design.

6 Economic Costs and the Distribution of Damages

6.1 Economic Costs: Back of the Envelope Calculation

We now use our regression estimates of the impact of exposure to upwind flaring on respiratory health to compare two back of the envelope calculations over our sample period. Here, we describe the scenarios and discuss our findings.

Our first back of the envelope calculation considers the health costs that could have been avoided if the 2018 flaring standards of Commission Order 24655 were in place at the beginning of our sample period. The order mandated that 88% of all natural gas produced must be captured as of November 2018. In reality only 72.6% of all gas produced was captured during the sample period. Leaving the total amount of natural gas produced during our sample time frame constant, meeting this standard would have required a reduction in the amount of natural gas flared by about 310 million thousand cubic feet (mcf). The preferred IV estimates at a range of 60 miles (Column 6 of Table 6) show that if this reduction occurred during the sample period, the rate of respiratory-related hospital visits would decrease by 5.23 (31.96%).[8] This corresponds to roughly 11,554 fewer respiratory-related hospital visits over the nine year sample period.

Putting a dollar value to an average respiratory-related hospital visit is a challenging task. Schlenker and Walker (2015) use an average Medicare reimbursement for respiratory-related diagnoses of $2,702 in 2006 dollars, which corresponds to $3,377 in 2018 dollars. Using this estimate of an average respiratory-related hospital visit, the hypothetical scenario of having an 88% capture rate in place prior to 2007 would have resulted in a $39 million reduction in health care costs over the sample period – roughly $4.3 million per year. Pfuntner et al. (2006) put the cost of a respiratory-related hospital stay at $25,422 in 2018 dollars. Because not all hospital visits result in a hospital stay, this serves as an upper bound on the total decrease in health costs. Under the 88% capture rate, this would have reduced health care costs by $296.3 million or roughly $32.9 million per year.

Our second back of the envelope calculation limits this analysis to only the respiratory health impacts

[8] Since this is a level-log specification a $p\%$ change in flaring corresponds to a $\hat{\beta}_1 \cdot log(1 + p/100)$ level change in the hospital visitation rate.

for zip codes within 30 miles of an active well during our sample period. Using the smaller IV estimates at a range of 30 miles (Column 3 of Table 6), we predict that the rate of respiratory-related hospital visits would decline by 2.32 (13.98%). This corresponds to 4,553 fewer respiratory-related hospital visits and a health cost reduction of $15.4 million over the entire sample period based on the respiratory-related hospital visit cost of $3,377.

Focusing on the $39 million estimate for a range of 60 miles, this result is economically significant for a few reasons. First, using estimates from Lade and Rudik (2019), our counterfactual reduced respiratory health costs from flaring is equivalent to 2.1% of the entire amount spent in North Dakota on gas gathering lines and well connections in the first 18 months that Commission Order 24665 was in place. Since many wells would have been profitable to connect to gas processing infrastructure in the absence of the regulation, this percentage is a lower bound on the percentage comparison of the health costs to the additional connection costs from Commission Order 24665. Second, these health cost estimates represent a significant proportion of the engineering based cost estimates of flaring described in section 2.3. Third, these estimates are based on a capture rate of 88%, thus if the 91% capture rate set to take effect by Commission Order 24665 in November 2020 were in effect at the start of our sample, the reduction in health costs would be greater. Fourth, as indicated previously the $39 million estimate doesn't account for the higher costs from a respiratory related hospital stay.

6.2 Geographic Distribution of Damages

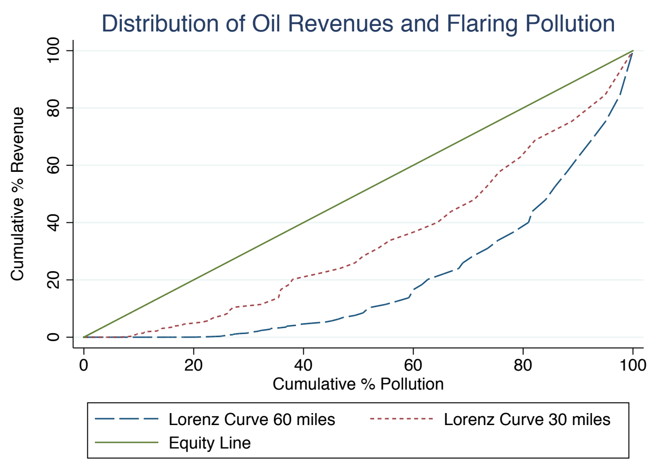
The literature on the welfare consequences of shale development is concerned with heterogeneity in the benefits for the local population (Bartik et al.), while the literature on environmental inequality often considers inequality in the exposure to various pollutants (Boyce et al., 2016). In this section, we ask whether the areas with the highest levels of extracted oil revenue also experienced the largest levels of exposure to flared natural gas. Our previous results indicate that exposure to flaring causes an increase in health costs for exposed individuals. We now investigate whether the health cost burden is felt more or less by areas with low levels of local resource development.

Figure 9 displays the disparity between extracted oil revenues and exposure to pollution from flaring for the geographic ranges of 30 and 60 miles. Zip code oil revenue is defined as the total amount of oil extracted from wells within the zip code multiplied by the monthly oil price, over the sample period.[9] The percentage of revenue is then determined by dividing the total revenue for a zip code by the combined sum of revenue across all sample zip codes. A zip code's proportion of pollution is defined in a similar manner,

[9] We use the monthly average price at the wellhead across North Dakota from the EIA's "First Purchase Prices by Area" data.

where pollution is exposure to upwind flared natural gas during the sample period.

Figure 9: Inequality in the Distribution of Pollution from Flaring



Results from Figure 9 indicate a modest amount of inequality in exposure to flared natural gas at a range of 30 miles with a Gini coefficient value of 0.663. This measure of exposure inequality increases to 0.796 once we expand the geographic range to include zip codes that were within 60 miles of an active well at any time during the sample period.[10] Focusing on the least productive zip codes, that combined to extract 20% of the oil revenues during the sample period, this group was exposed to more than 31.8% of the total upwind flared at a range of 30 miles and 62% of the total at a range of 60 miles. Moreover, using the $3,377 cost of a respiratory-related hospital visit and our preferred IV estimates (Column 6 of Table 6), we find that the estimated reduction in health costs under the 88% capture rate of Commission Order 24655 exceeds the oil revenues for 30.9% of zip codes at a range of 30 miles and 48.3% of zip codes at a range of 60 miles. Therefore, there is evidence that many zip codes exposed to flared natural gas did not extract significant amounts of oil during the sample period.

There are a number of caveats to this analysis. First, it is important to note that our use of a Gini coefficient is only a measure of inequality on the basis of oil revenues and exposure, not on any other measure such as minority status. Second, there is significant variation in oil prices across wells in North Dakota that could impact our revenue estimates. However, we obtain similar results if we restrict our analysis to the distribution of oil barrels extracted. Third, due to high rates of absentee mineral ownership in the Bakken, there is no guarantee that the residents of a zip code that extracts a significant amount of oil receive substantial benefits (Brown et al., 2019). Therefore, this analysis only shows a disparity in geographic flaring exposure and oil extraction, not necessarily the distribution of royalties. Nonetheless, this analysis is

[10] The Gini coefficient lies in the interval between zero and one, with higher values denoting greater levels of inequality.

important because the value of the oil extracted represents an upper bound on the local benefits of resource development and for many areas this value is less than the estimated health costs from flaring exposure.

7 Conclusion

The primary contribution of this paper is to identify a causal relationship between flared natural gas exposure and respiratory health. Our approach exploits variation in exposure to flared natural gas induced by naturally occurring changes in wind direction interacted with the upwind well's constrained natural gas processing capacity. In the quasi-experimental design, treatment is defined as a zip code being downwind from a well that is constrained in the amount of natural gas it can sell. The analysis provides suggestive evidence of a causal link between the amount of flared natural gas up to 60 miles upwind from a zip code and the percent of that zip code's population that experiences a respiratory-related hospital visit within a given month. As a secondary contribution, we document that a significant portion of the damages from flaring are distributed to zip codes with little to no shale development.

The size of the estimated effect of flared natural gas exposure and respiratory health is substantial. In our preferred specification, a 1% increase in upwind flared natural gas within 60 miles leads to an increase of 0.12 in the the zip code's respiratory-related hospital visitation rate for a given month. If flaring had been reduced by 54% during our nine year sample period, this would reduce the total number of respiratory-related hospital visits for individuals who live within 60 miles of an active oil well by 11,554, which is roughly valued between $39 million and $296.3 million USD.

While we believe that this analysis provides substantial evidence that the flaring of natural gas had a significant impact on respiratory health in North Dakota, the damages from flaring could be greater in other settings. Specifically, the Permian Basin and Eagle Ford in Texas flared roughly 160 million mcf of natural gas in 2018 (OGJ, 2019). This is more than double the annual average observed in North Dakota during our sample period. Furthermore, the population of the Permian Basin and Eagle Ford is roughly triple the 180,000 North Dakota residents that lived within 60 miles of an active well during our sample (Emerging, 2017). This indicates that the costs of flaring could be much greater in other settings.

To our knowledge, this is the first paper to document a causal relationship between exposure to flared natural gas and changes in respiratory health. From a policy standpoint, the analysis points to an additional social cost from oil and natural gas extraction that may be potentially mitigated by firms' response to regulation. More generally, our work contributes to the growing recognition that there are significant negative externalities associated with shale development that are difficult to quantify, but economically significant.

References

Adgate, J. L., Goldstein, B. D., and McKenzie, L. M. (2014). Potential public health hazards, exposures and health effects from unconventional natural gas development. *Environmental Science & Technology*, 48(15):8307–8320.

Anderson, M. L. (2015). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*.

Banzhaf, S. H. and Walsh, R. P. (2008). Do people vote with their feet? an empirical test of tiebout's mechanism. *The American Economic Review*, 98(3):843–863.

Barboza, T. (2015). Natural gas leak that's sickening valley residents could take months to fix. In *Los Angeles Times, November 20, 2015*. Los Angeles Times.

Bartik, A. W., Currie, J., Greenstone, M., and Knittel, C. R. The local economic and welfare consequences of hydraulic fracturing. *American Economic Journal: Applied Economics*.

Bechle, M. J., Millet, D. B., and Marshall, J. D. (2013). Remote sensing of exposure to no2: Satellite versus ground-based measurement in a large urban area. *Atmospheric Environment*, 69:345–353.

Black, K. J., McCoy, S. J., and Weber, J. G. (2018). When externalities are taxed: The effects and incidence of pennsylvania?s impact fee on shale gas wells. *Journal of the Association of Environmental and Resource Economists*, 5(1):107–153.

Boslett, A., Hill, E., Ma, L., , and Zhang, L. (2019). The impact of shale development on light pollution. Working Paper.

Boxall, P., Chan, W., and Mcmillan, M. (2005). The impact of oil and natural gas facilities on rural residential property values: A spatial hedonic analysis. *Resource and Energy Economics*, 27:248–269.

Boyce, J. K., Zwickl, K., and Ash, M. (2016). Measuring environmental inequality. *Ecological Economics*, 124:114–123.

Brown, J. P., Fitzgerald, T., and Weber, J. G. (2019). Does resource ownership matter? oil and gas royalties and the income effect of extraction. *Journal of the Association of Environmental and Resource Economists*, 6(6):1039–1064.

Cascio, E. U. and Narayan, A. (2015). Who needs a fracking education? the educational response to low-skill biased technological change. Working Paper 21359, National Bureau of Economic Research.

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

Chay, K. Y. and Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, 113(2):376–424.

Collins, B. (2018). Are some shale producers under-reporting gas flaring to keep oil flowing. Report, S and P Global.

Currie, J., Neidell, M., and Schmieder, J. F. (2009). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics*, 28(3):688–703.

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

Cushing, L., Johnston, J., Franklin, M., and Chau, K. (2018). OP III 3 Using satellite observations to estimate exposure to flaring: implications for future studies of the health impacts of unconventional oil and gas operations. volume 75.

Dave, H. (2009). So why are all these gas flares burning in the oil fields? *Department of Mineral Resources Newsletter*.

Ehrman, M. U. (2014). Lights Out in the Bakken: A Review and Analysis of Flaring Regulation and Its Potential Effect on North Dakota Shale Oil Production. *W. Va. L. Rev.*, 117:549.

Emerging, T. (2017). Midland–odessa. *Population*, pages 327–762.

Energy Information Administration (2006). Natural gas processing: The crucial link between natural gas production and its transportation to market. *Energy Information Administration, Office of Oil and Gas*.

Environmental Protection Agency (1983). Flare efficiency study. EPA-600/2-83-052. https://www3.epa.gov/ttn/chief/old/ap42/ch13/s05/reference/ref_01c13s05_jan1995.pdf.

Environmental Protection Agency (2018). *Industrial Flares*, volume 1, chapter 13.5. Fifth edition. https://www3.epa.gov/ttn/chief/ap42/ch13/final/C13S05_02-05-18.pdf.

Environmental Protection Agency (2019a). 40 CFR Part 60. https://www.epa.gov/sites/production/files/2019-08/documents/frn_oil_and_gas_review_2060-at90_nprm_20190828revised_d.pdf.

Environmental Protection Agency (2019b). EPA Pollution Control Cost Manual. Section 3 – VOC Controls. https://www.epa.gov/sites/production/files/2019-08/documents/flarescostmanualchapter7thedition_august2019vff.pdf.

Enz, J. W. (2003). North Dakota Topographic, Climate, and Agricultural Overview. *North Dakota State University. Available online at http://www.ndsu.edu/ndsco/publication/ndsco/ndclimate.pdf.*

Fershee, J. P. (2012). The Oil and Gas Evolution: Learning from the Hydraulic Fracturing Experiences in North Dakota and West Virginia. *Tex. Wesleyan L. Rev.*, 19:23.

Fitzgerald, T. and Stiglbauer, C. (2015). Flaring of associated natural gas in the bakken shale. In *The Dynamic Energy Landscape, 33rd USAEE/IAEE North American Conference, Oct 25-28, 2015*. International Association for Energy Economics.

Ford, M., Wilczewski, W., and Nülle, G. (2015). North Dakota natural gas flaring targets challenged by rapid production growth. Article 23752, U.S Energy Information Administration.

Franklin, M., Chau, K., Cushing, L. J., and Johnston, J. E. (2019). Characterizing Flaring from Unconventional Oil and Gas Operations in South Texas Using Satellite Observations. *Environmental Science & Technology*, 53(4):2220–2228. PMID: 30657671.

Giwa, S. O., Nwaokocha, C. N., Kuye, S. I., and Adama, K. O. (2019). Gas flaring attendant impacts of criteria and particulate pollutants: A case of niger delta region of nigeria. *Journal of King Saud University-Engineering Sciences*, 31(3):209–217.

Goldberg, D. L., Lamsal, L. N., Loughner, C. P., Swartz, W. H., Lu, Z., and Streets, D. G. (2017). A high-resolution and observationally constrained omi no2 satellite retrieval. *Atmospheric Chemistry and Physics (Online)*, 17(18).

Government Accountability Office (2011). Information on tall smokestacks and their contribution to interstate transport of air pollution. https://www.gao.gov/assets/320/318175.pdf.

Graham, J., Irving, J., Tang, X., Sellers, S., Crisp, J., Horwitz, D., Muehlenbachs, L., Krupnick, A., and Carey, D. (2015). Increased traffic accident rates associated with shale gas drilling in Pennsylvania. *Accident Analysis & Prevention*, 74:203–209.

Grainger, C., Schreiber, A., and Chang, W. (2016). How states comply with federal regulations: Strategic ambient pollution monitoring. Technical report, mimeo.

Hausman, C. and Kellogg, R. (2015). Welfare and distributional implications of shale gas. *Brookings Papers on Economic Activity*, 46(1 (Spring)):71–139.

Hausman, C. and Muehlenbachs, L. (2019). Price regulation and environmental externalities: Evidence from methane leaks. *Journal of the Association of Environmental and Resource Economists*, 6(1):73–109.

Hill, E. L. (2018). Shale gas development and infant health: evidence from pennsylvania. *Journal of health economics*, 61:134–150.

Holland, B. (2019). Prioritizing oil, Texas Railroad Commission rejects challenge to gas flaring permit. *S and P Global*.

Holland, S. P., Mansur, E. T., Muller, N. Z., and Yates, A. J. (2016). Are there environmental benefits from driving electric vehicles? the importance of local factors. *American Economic Review*, 106(12):3700–3729.

Johnsen, R., LaRiviere, J., and Wolff, H. (2019). Fracking, coal, and air quality. *Journal of the Association of Environmental and Resource Economists*, 6(5):1001–1037.

Kellogg, R. (2014). The effect of uncertainty on investment: evidence from Texas oil drilling. *The American Economic Review*, 104(6):1698–1734.

Kleinberg, R. (2019). Greenhouse gas footprint of oilfield flares accounting for realistic flare gas composition and distribution of flare efficiencies. *Earth and Space Science Open Archive*, page 14.

Knittel, C. R., Miller, D. L., and Sanders, N. J. (2016). Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, 98(2):350–366.

Krulwich, R. (2013). A Mysterious Patch Of Light Shows Up In The North Dakota Dark. *Krulwich Wonders*.

Krupnick, A. J. and Echarte, I. (2017). Health impacts of unconventional oil and gas development. *Resources for the Future (RFF). June*.

Lade, G. E. and Rudik, I. (2019). Costs of inefficient regulation: Evidence from the bakken.

Laughner, J. L., Zhu, Q., and Cohen, R. C. (2018). The berkeley high resolution tropospheric no 2 product. *Earth System Science Data*, 10(4):2069–2095.

Litman, T. (2015). Transportation cost and benefit analysis ii–air pollution costs. *Victoria Transport Policy Institute, Victoria BC, Canada*.

MacPherson, J. (2019). North Dakota oil producers are wasting billions of cubic feet of natural gas. *The Los Angeles Times*.

Mason, C. F., Muehlenbachs, L. A., and Olmstead, S. M. (2015). The economics of shale gas development. *Annual Review of Resource Economics*, 7(1):269–289.

Mauzerall, D. L., Sultan, B., Kim, N., and Bradford, D. F. (2005). Nox emissions from large point sources: variability in ozone production, resulting health damages and economic costs. *Atmospheric Environment*, 39(16):2851–2866.

McKenzie, L. M., Witter, R. Z., Newman, L. S., and Adgate, J. L. (2012). Human health risk assessment of air emissions from development of unconventional natural gas resources. *Science of The Total Environment*, 424:79 – 87.

Moretti, E. and Neidell, M. (2011). Pollution, health, and avoidance behavior evidence from the ports of Los Angeles. *Journal of Human Resources*, 46(1):154–175.

Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, 105(12):3633–59.

Neidell, M. (2009). Information, avoidance behavior, and health the effect of ozone on asthma hospitalizations. *Journal of Human Resources*, 44(2):450–478.

OGJ, E. (2011). Hess Awards Design for North Dakota Gas Plant Expansion. *Oil and Gas Journal*.

OGJ, E. (2019). Permian gas flaring, venting reaches record high. *Oil and Gas Journal*.

Pfuntner, A., Wier, L. M., and Steiner, C. (2006). Costs for hospital stays in the united states, 2010: statistical brief# 146.

Schlenker, W. and Walker, W. R. (2015). Airports, air pollution, and contemporaneous health. *The Review of Economic Studies*, page rdv043.

Staiger, D. and Stock, J. H. (1994). Instrumental variables regression with weak instruments. Technical report, National Bureau of Economic Research.

Tong, D. Q., Muller, N. Z., Mauzerall, D. L., and Mendelsohn, R. O. (2006). Integrated assessment of the spatial variability of ozone impacts from emissions of nitrogen oxides.

Tustin, A. W., Hirsch, A. G., Rasmussen, S. G., Casey, J. A., Bandeen-Roche, K., and Schwartz, B. S. (2017). Associations between unconventional natural gas development and nasal and sinus, migraine headache, and fatigue symptoms in pennsylvania. *Environmental Health Perspectives*, 125(2):189–197.

Wang, Q., Chen, X., Jha, A. N., and Rogers, H. (2014). Natural gas from shale formation–the evolution, evidences and challenges of shale gas revolution in united states. *Renewable and Sustainable Energy Reviews*, 30:1–28.

Weber, J. G. (2014). A decade of natural gas development: The makings of a resource curse? *Resource and Energy Economics*, 37:168 – 183.

Werner, A. K., Vink, S., Watt, K., and Jagals, P. (2015). Environmental health impacts of unconventional natural gas development: a review of the current strength of evidence. *Science of the Total Environment*, Journal of Political Economy:1127–1141.

Whitworth, K. W., Marshall, A. K., and Symanski, E. (2018). Drilling and production activity related to unconventional gas development and severity of preterm birth. *Environmental Health Perspectives*, 126(3):037006.

Willis, M., Hystad, P., Denham, A., and Hill, E. (2020). Natural gas development, flaring practices, and pediatric asthma hospitalizations in texas. Working Paper.

Willis, M. D., Jusko, T. A., Halterman, J. S., and Hill, E. L. (2018). Unconventional natural gas development and pediatric asthma hospitalizations in pennsylvania. *Environmental research*, 166:402–408.

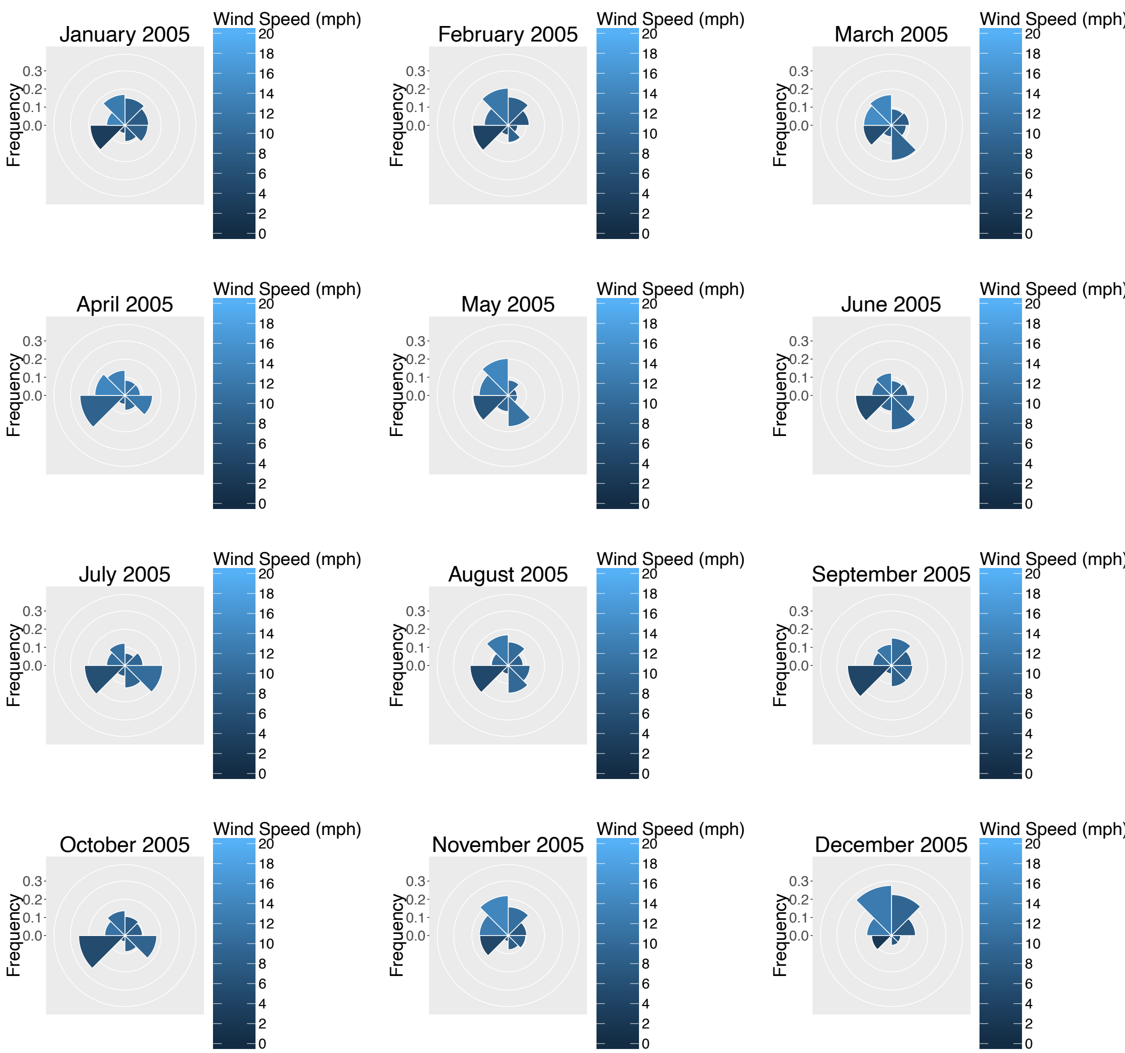
A Appendix

A.1 Supplementary Tables, Figures, and Analysis

A.1.1 Wind Patterns Across Months

Figure A1 shows the proportion of wind direction for each month in 2005. The prevailing wind is from the east, blowing to the west. However there is variation across months, with a southerly wind, blowing to the northwest in November and December. Overall, Figure A1 and Figure 8 indicate significant within and across month variation in wind direction.

Figure A1: Frequency of time the wind is blowing toward and the average speed in each octant for 2005



A.1.2 Robustness of Satellite Analysis

In Table A1 we re-estimate the specifications in Table 3 using an alternative measure of NO_2. Specifically, we use the NO_2 reading from the grid point closest to the zip code centroid, rather than the average of all satellite readings within the zip code. This alternative is consistent with the fact that exposure to flared natural gas for a zip code is measured at the zip code's centroid. Results are qualitatively similar to the results in Table 3, consistent with the understanding that our measure of exposure to flared natural gas corresponds to measures of harmful pollutants.

Table A1: OLS Estimates of Flaring and Zip Code Center NO_2 Readings

	(1) 30 Miles b/se	(2) 60 Miles b/se	(3) 90 Miles b/se	(4) 120 Miles b/se
Panel A: Z-Score Center NO_2 Reading				
Log(UpwindFlaring)	0.026***	0.019*	-0.002	-0.003
	(0.009)	(0.011)	(0.012)	(0.011)
N	8,545	11,329	14,876	18,175
R^2	0.004	0.003	0.002	0.001
Panel B: Zip Code Center NO_2 Reading				
Log(UpwindFlaring)	0.024***	0.023***	0.010*	0.005
	(0.003)	(0.004)	(0.006)	(0.004)
N	8,554	11,345	14,896	18,196
R^2	0.382	0.382	0.370	0.378
	15-30 Miles	30-60 Miles	60-90 Miles	90-120 Miles
Panel C: Donut Specification Z-Score Center NO_2 Reading				
Log(UpwindFlaring)	0.030***	0.005	-0.015	-0.001
	(0.010)	(0.011)	(0.013)	(0.013)
N	8,545	11,329	14,876	18,175
R^2	0.004	0.003	0.003	0.001
Panel D: Donut Specification Center Zip Code NO_2 Reading				
Log(UpwindFlaring)	0.020***	0.011***	-0.004	-0.001
	(0.004)	(0.004)	(0.005)	(0.004)
N	8,554	11,345	14,896	18,196
R^2	0.382	0.385	0.373	0.378
Zip Code Fixed Effects	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y

Notes: This table reports OLS regressions of upwind natural gas flaring on both the zip code centroid NO_2 z-score and monthly atmospheric column density NO_2 reading (molecules/cm2) for zip code centroids in North Dakota from 2007 - 2015. Standard errors are clustered at the zip code level. The panel is not balanced due to gaps in satellite coverage during the sample period. Panels C and D include the log of the total upwind natural gas flared from wells within the inner portion of the donut as an additional control. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level. **Sample:** Column (1) considers monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Column (2) considers monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period. Column (3) considers monthly observations from 189 zip codes that were within 90 miles of at least one active well during the sample period. Column (4) considers monthly observations from 227 zip codes that were within 120 miles of at least one active well during the sample period.

Table A2: Primary IV Results with Different Standard Error Estimates

	30 Miles			60 Miles		
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Log(UpwindFlaring)	0.031	0.049	0.052	0.076	0.111	0.117
Zip Code Clustered SE	(0.025)	(0.031)	(0.031)	(0.045)	(0.052)	(0.052)
County Code Clustered SE	(0.020)	(0.028)	(0.028)	(0.050)	(0.062)	(0.064)
Huber-White SE	(0.019)	(0.024)	(0.025)	(0.037)	(0.043)	(0.044)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	-	Y	Y	-	Y	Y
Demographic Controls	-	-	Y	-	-	Y
N	11,550	11,550	11,550	15,225	15,225	15,225

Notes: This table reports IV regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2006 - 2015. The estimation uses the log of the number of wells connected to a constrained processing plant upwind from a zip code as an instrument for that zip codes exposure to flared natural gas. The estimates are of the impact of flared natural gas on the rate of respiratory-related hospital visits in a zip code. Columns (1) and (4) add controls for the number of active wells within the specified range to capture the trend in shale development over time. Columns (2) and (5) also add total oil extracted upwind, oil extracted within range, the number of upwind wells drilled, total wells drilled within the specified range, and total upwind wells of the zip code as additional controls. Columns (3) and (6) further add controls related to the average house size, value, and income for the zip code as demographic controls. The first set of standard errors are clustered at the zip code level. The second set of standard errors are clustered at the county level. The third set of standard are White-Huber or Heteroskedastic consistent standard errors. **Sample:** Columns (1) - (3) consider monthly observations from 110 zip codes that were within 30 miles of at least one active well during the sample period. Columns (4) - (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

A.1.3 Alternative Standard Errors

In Table A2 we estimate alternative sets of standard errors for our primary point estimates in Table 6. In addition to estimating standard errors clustered at the zip code level, we estimate standard errors clustered at the county level, as well as Huber-White robust standard errors. For the geographic range of 30 miles, both sets of alternative standard errors are smaller than the original standard errors clustered at the zip code level. At a range of 60 miles, the standard errors clustered at the county level are slightly larger than the zip code level ones. However, our primary estimate in column (6) is still statistically significant at conventional levels. Overall, this analysis demonstrates that the statistical significance of our primary results is not specific to the choice of standard errors clustered at the zip code level.

A.1.4 New Wells

If newly producing wells have higher emissions factors for each unit of flared natural gas, then failing to account for these wells may bias our primary estimates. Column (1) of Table A3 adds the log of the number of wells in their first production upwind from the zip code as an additional control. With the additional control, the new point estimate of the impact of flared natural gas on respiratory health is larger than our primary result in Column (6) of Table 6.

Table A3: Primary IV Results 60 Miles: Robustness and Alternative Analysis

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Log(UpwindFlaring)	0.133**	0.328**	0.096**	0.104**	0.075*	0.142*
	(0.061)	(0.147)	(0.044)	(0.044)	(0.044)	(0.083)
Log(New Wells)	-0.043*					
	(0.024)					
Log(UpwindFlaring$_{i,t-1}$)		-0.221**				
		(0.100)				
Health$_{i,t-1}$			0.090***			
			(0.022)			
Log(CumulativeUpwindFlaring)				-0.052**		
				(0.026)		
Log(Zip Code Population)						0.126
						(0.081)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	Y	Y	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y	Y	Y
N	15,225	15,225	15,225	15,225	15,225	15,225

Notes: This table reports IV regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2006 - 2015. The estimation uses the log of the number of wells connected to a constrained processing plant upwind from a zip code as an instrument for that zip codes exposure to flared natural gas. The estimates are of the impact of flared natural gas on the rate of respiratory-related hospital visits in a zip code. All specifications include both the number of active wells, oil extracted, and wells drilled both upwind and total within the specified range of the zip code, along with average house size, value, and income as additional controls. Column (1) adds the log of the number of wells in their first month of production upwind as a control. Column (2) adds the lag of the log of upwind flaring from the previous month as a control. Column (3) adds the lag of the log of upwind flaring from the previous month as a control. Column (4) adds the log of cumulative upwind flaring exposure up to that point as a control. Column (5) raises the restriction for a natural gas processing plant to be considered constrained up to 70%. Column (6) changes the primary outcome to the log of one plus the count of respiratory-related hospital visits and the log of the zip code population as a control. Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. **Sample:** Columns (1) - (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

A.1.5 Leads, Lags, and Cumulative Exposure

To address concerns that there is serial correlation in respiratory health outcomes over time that is impacting our results, Columns (2) to (4) of Table A3 allow for lagged effects of flaring and respiratory health. As shown, although there is autocorrelation in health and flaring, the point estimates of the coefficient of interest is not disturbed. The negative correlation between cumulative (since the beginning of the sample period) exposure to flared natural gas and health is consistent with some harvesting of cases that would have occurred in subsequent months. However, the consistent magnitude of our coefficient of interest in these specifications indicates there is a contemporaneous relationship between flaring and health.

A.1.6 Sensitivity to Instrument and Outcome Definition

To examine the sensitivity of our main results to the choice of a 60% ratio between wet natural gas received and plant processing capacity to define a plant as constrained, we re-estimate our primary specification with a 70% threshold. Although the point estimate in Column (5) of Table A3 is smaller under this more restrictive definition, the difference between the two point estimates is not statistically significant. Therefore, we do not believe our primary results are driven by the choice of the 60% threshold for defining a plant as constrained. Column (6) presents the results of our primary specification, with the log of one plus the count of respiratory-related hospital visits, as the alternative outcome. We also include the log of the zip code population as an additional control in this specification. The estimated impact of exposure to upwind flared natural gas on respiratory health is similar under this alternative health outcome, indicating that our results are not driven by the choice of using a population rate as the primary outcome of interest.

A.1.7 Sensitivity to Wind Power and Monitor Choice

To test for the effect that wind speed also impacts the concentration of pollutants from upwind flared natural gas, we adjust our definition of exposure as follows:

$$UpwindFlaring_{it} = \sum_{q=1}^{8} Wind_{iqt} \cdot Flaring_{iqt} \cdot \frac{Speed_{iqt}}{Avg.Speed_{it}}.$$

This allows us to give a higher weight to flaring from octants with higher wind speeds, which may carry more pollutants to the zip code centroid. In Column (1) of Table A4 we re-estimate our primary specification using this speed adjusted measure of exposure for upwind flaring, constrained wells, oil production, drilling, and total wells. In addition, we test for the sensitivity of our main results to the construction of zip code weather from kriging between weather monitors by re-estimating our primary specification using only one

Table A4: Results 60 Miles: Additional Robustness

	IV				OLS	
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Log(UpwindFlaring)	0.106**	0.122**	0.118**	0.160**	-0.001	-0.001
	(0.048)	(0.054)	(0.052)	(0.080)	(0.004)	(0.004)
Log(Upwind Avg.Well - N.G. Distance)			0.034		0.004	
			(0.022)		(0.012)	
Upwind Avg.Well - N.G. Capacity				-0.028		0.028
				(0.039)		(0.021)
Zip Code Fixed Effects	Y	Y	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y	Y	Y
Active Wells Controls	Y	Y	Y	Y	Y	Y
Oil Production Controls	Y	Y	Y	Y	Y	Y
Demographic Controls	Y	Y	Y	Y	Y	Y
N	15,225	15,225	15,225	15,225	15,225	15,225

Notes: This table reports both IV and OLS regressions of upwind natural gas flaring on the rate of respiratory-related hospital visits for zip codes in North Dakota from 2007 - 2015. Columns (1) to (4) use the log of the number of wells connected to a constrained processing plant upwind from a zip code as an instrument for that zip codes exposure to flared natural gas. Columns (5) and (6) are OLS estimates. The estimates are of the impact of flared natural gas on the rate of respiratory-related hospital visits in a zip code. All specifications include both the number of active wells, oil extracted, and wells drilled both upwind and total within the specified range of the zip code, along with average house size, value, and income as additional controls. Column (1) incorporates wind speed in the weighting of upwind flaring exposure. Column (2) only uses wind data for a single monitor. Columns (3) and (5) add the log of the average distance between upwind wells and their nearest NG processing plant as an additional control. Columns (4) and (6) add the average processing capacity utilized by the NG plant upwind wells are connected to as an additional control (measured from 0 to 1, with 0.60 meaning 60%). Standard errors are clustered at the zip code level. *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level.
Sample: Columns (1) - (6) consider monthly observations from 145 zip codes that were within 60 miles of at least one active well during the sample period.

central monitor. Specifically, we use the weather data from the Garrison (KN60) weather station, which is central and proximate for many of our sample zip codes that have an active well within 60 miles of their centroid. Column (2) of Table A4 presents the result using this more limited weather data. As shown, our estimates of the impact of flared natural gas upwind on local respiratory health under these alternative weather data constructions are remarkably similar to our primary results.

A.1.8 Endogeneity of Natural Gas Plant Location and Capacity

It is important to note that our primary analysis relies on the assumption that the interaction of wind patterns and the number of upwind wells connected to a constrained plant is exogenous to local determinants of respiratory health, conditional on the inclusion of our extensive controls. As discussed in section 4, there are a number of reasons why we believe our instrumental variables approach identifies the impact of upwind flared natural gas on local respiratory health. In particular, after including our set of extensive upwind production controls, we believe that the location and capacity of the nearest natural gas plant to wells should be unrelated to any additional unobserved determinants of downwind respiratory health. Of course, it is impossible to test the validity of the exclusion restriction in an instrumental variables analysis, therefore we investigate whether changes in the distance of upwind wells to the nearest natural gas processing plant or changes in the average excess capacity of the nearest natural gas plant to upwind wells, correlates with changes in local respiratory health. Columns (3) and (4) of Table A4 include these additional controls with our primary IV specification, with qualitatively similar results for the impact of upwind flaring on local respiratory health. Since average capacity used is defined from zero to one we do not take the log of it. Columns (5) and (6) of Table A4 estimate an OLS regression of our primary specification with these additional upwind processing controls. In all four of these specifications, neither the average distance of upwind wells to the nearest natural gas plant or average capacity used by the nearest natural gas plant have a statistically significant relationship with local respiratory health.

A.1.9 Alternative Difference-in-Differences Estimation Strategy

To further test the robustness of our main results, we estimate difference-in-differences (DiD) specifications for two sub samples of our primary analysis data. In the first sub sample, we consider monthly observations from nine North Dakota zip codes that were above the latitude of 46°48'48" and at the longitude of 101°42'45" (along North Dakota Highway 28). As shown in Figure 1, there was a steady progression of flaring to the west of this area during our sample period and there is relatively little concern about flaring from other directions. In the second sub sample, we consider zip codes that were in the second and fifth quintile of total wet natural gas extracted from wells within 60 miles of the zip code centroid in October 2013 (when initial planning for Order 24665 was announced). For the second sub sample we focus only on the twelve months pre and post Order 24665. To limit the influence of outlier for these smaller sub samples, we omit observations above the 95th percentile for respiratory-related hospital visits. Specifically, we estimate the following formula:

$$Health_{it} = \beta_0 + \beta_1 \cdot Treat_{it} \cdot Post_t + \beta_2 \cdot Post_t + \beta_3 \cdot Treat_{it} + \gamma_i + \delta_t + \varepsilon_{it}.$$

For estimation with the first sub sample, $Post$ is the percentage of time the wind is blowing from the west and $Treat$ takes a value of one if the zip code shares a western border with zip codes that have flaring activity. In this first alternative specification, we essentially use observations from these zip codes when the wind is not blowing from the west, or when there is no adjacent flaring activity, as a control for when a zip code is exposed to flared natural gas. That is, given their close proximity, all of these zip codes should be experiencing similar trends in local respiratory health and the impact of flared natural gas on health is identified by slight differences in the timing of adjacent flaring activity. Columns (1) and (2) of Table A5 present the results for this first DiD analysis. These results indicate that when there is westward flaring activity, the respiratory-related hospital visitation rate increases with the rate that the wind is blowing from

Table A5: Alternative Diff-in-Diff Estimation Results

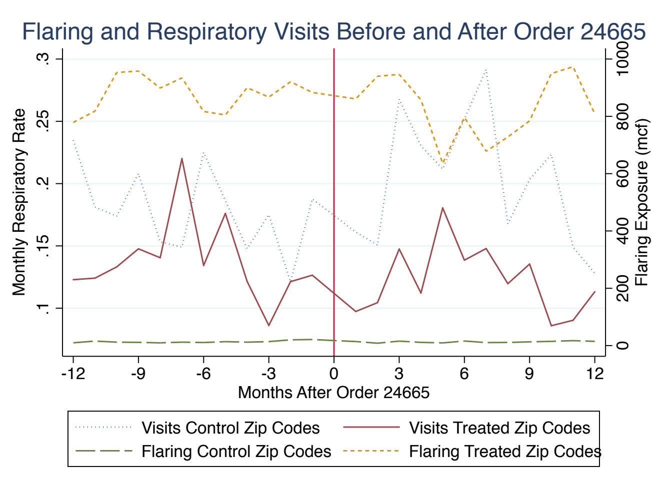
	East vs. West		2nd Quintile vs. 5th Quintile	
	(1) b/se	(2) b/se	(3) b/se	(4) b/se
Treat·Post	0.207* (0.099)	0.254 (0.151)	-0.039* (0.023)	-0.089*** (0.032)
Zip Code Fixed Effects	Y	Y	Y	Y
Month Fixed Effects	Y	Y	Y	Y
Active Wells Controls	-	Y	-	Y
Oil Production Controls	-	Y	-	Y
Demographic Controls	-	Y	-	Y
N	925	925	1,352	1,352

Notes: This table reports differences-in-differences regressions that measure the impact of exposure to flared natural gas on the respiratory-related hospital visitation rate for zip codes in North Dakota from 2007 - 2015. Standard errors are clustered at the zip code level. For the specifications in columns (1) and (2), Treat is defined as the percentage of time the wind is blowing from the west and post takes a value of one if there is well flaring to the west of the zip code. For the specifications in columns (3) and (4), Treat is defined as whether the zip code was in the top quintile for natural gas extracted within the geographic range and Post takes a value of one if the time period is after the implementation of Order 24665. Specifications in columns (2) and (4) include both the number of active wells, oil extracted, and wells drilled both upwind and total within the specified range of the zip code, along with average house size, value, and income as additional controls.*** indicates significance at the 1% level, ** at the 5% level, * at the 10% level. **Sample:** Columns (1) and (2) considers monthly observations from nine North Dakota zip codes that were above the latitude of 46°48'48" and at the longitude of 101°42'45". Columns (3) and (4) consider the zip codes in the second and fifth quintile of total wet natural gas extracted from wells within 60 miles of the zip code centroid in October 2013. The time period considered is October 2013 to October 2015.

the west. In addition, this result is statistically significant for the base specification, without additional controls.

For estimation with the second sub sample, we exploit the timing and implementation of North Dakota Order 24665, which reduced flaring, to estimate the impact of flared natural gas on respiratory health. Specifically, $Post$ takes a value of one if the month is in the post Order 24665 (October 2014) time period, while $Treat$ takes a value of one if the zip code was in the fifth quintile of total wet natural gas extracted from wells within 60 miles of the zip code centroid in October 2013. The control group for this analysis is zip codes in the second quintile of nearby natural gas extraction, as these zip codes are proximate to the economic activity from nearby shale development, thereby experiencing its impacts on local respiratory health but with little exposure to flared natural gas. Figure A2 demonstrates the respiratory-related hospital visitation rate and flaring exposure for these two groups before and after the implementation of Order 24665. Following the implementation of Order 24665, treated zip codes experienced both a decline in their exposure to upwind flared natural gas and their monthly respiratory-related hospital visitation rate. In contrast, control zip codes experienced little to no change in their exposure to flared natural gas and a slight increase in their respiratory-related hospital visitation rate. Columns (3) and (4) of Table A5 present the results for this second DiD analysis, with both specifications indicating a statistically significant and negative effect of Order 24665 on zip codes with large levels of nearby natural gas extraction activity.

Figure A2: Impact of Order 24665 on Flaring and Respiratory Health



Overall, the results of these alternative DiD analyses are consistent with our primary result that exposure to upwind flared natural gas increases the respiratory-related hospital visitation rate. Furthermore, these alternative estimation strategies indicate that our primary results are not driven by the use of an instrumental variables estimation strategy.