The Effect of Unconventional Oil and Gas Drilling on Healthcare Utilization: Evidence from Texas

Cara Haughey

6/10/2025

Abstract

This paper explores the relationship between adult healthcare utilization and hydraulic fracturing on the Eagle Ford Shale formation in Texas. Using a staggered difference-in-differences estimation strategy, I find that inpatient claims for conditions plausibly related to hydraulic fracturing increase by 15 percent in shale zip codes after the entry of hydraulic fracturing, relative to zip codes outside of the shale formation. However, this result could be explained by pollution exposure from hydraulic fracturing or compositional changes related to migration. I therefore construct a novel dataset that combines public water source data with healthcare claims and oil and natural gas well permit data to compare healthcare utilization in counties with unconventional drilling within one kilometer of a public groundwater source to those with drilling farther away from their public groundwater source. I find that the increases in healthcare utilization are 13 percent larger in counties with drilling near a water source when compared to counties with drilling farther away. This result not only suggests that the increases in utilization are not likely driven by compositional changes, but also that groundwater contamination may serve as a primary pathway for unconventional drilling pollution exposure.

JEL Codes– Q30, Q53, I12, H23

Keywords— Fracking, Unconventional Drilling, Drinking Water, Groundwater, Social Determinants of Health

1 Introduction

Hydraulic fracturing, horizontal drilling, and other unconventional gas and oil drilling methods[1] changed the landscape of natural resource extraction by allowing oil and natural gas producers to reach resources in previously inaccessible shale formations. Over the past few decades, these methods significantly increased and expanded oil and natural gas production in the United States. However, the use of unconventional methods in resource extraction is not without controversy; supporters tout the economic benefits from increased employment and higher wages, while opponents question the environmental and health hazards resulting from the new technology.

Hausman and Kellogg (2015) conduct a welfare analysis to measure the economic benefits of the hydraulic fracturing boom. The authors find significant increases in consumer surplus associated with lower natural gas prices, with a estimated $48 billion in increased, total surplus. However, the authors note that their welfare calculations exclude negative externalities, such as environmental damage and negative health effects, due to a lack of research. Although numerous studies identify potential pathways that could lead to poor health outcomes associated with these new technologies (Colborn et al. 2011; Epstein 2017; McKenzie et al. 2012; Shonkoff, Hays, and Finkel 2014), the research causally identifying negative health consequences resulting from unconventional oil and gas drilling is limited. A complete assessment of the welfare effects of unconventional drilling would consider health effects, however hydraulic fracturing is still a relatively new technology and any health-related ramifications may take time to develop. Therefore, continually updating research on the health effects of hydraulic fracturing will lead to a better understanding of the impact of hydraulic fracturing.

In this paper, I combine inpatient, healthcare claims data with oil and natural gas well locations to identify the effect of hydraulic fracturing on healthcare utilization. Prior work shows that unconventional drilling negatively affects infants whose mothers lived near drilling sites during pregnancy (Hill and Ma 2022). I extend this work by estimating the effect of hydraulic fracturing on inpatient healthcare utilization for a broad spectrum of conditions using variation in exposure to hydraulic fracturing on the Eagle Ford Shale formation in southern Texas. I find an 15 percent increase in inpatient healthcare utilization after the entry of hydraulic fracturing for zip codes located on the Eagle Ford Shale formation compared to zip codes outside of the formation. Further, I find larger increases in inpatient utilization for specific conditions, including cancer, conditions of the central nervous system, gastrointestinal conditions, urological conditions, dermatological conditions, and blood diseases.

These effects could be a result of exposure to hydraulic fracturing or compositional changes due to

1. Hydraulic fracturing or the combination of hydraulic fracturing and horizontal drilling are more commonly known as “fracking.”

migration related to hydraulic fracturing. For example, oil and gas workers may move to the region for jobs. If these workers are exposed to higher levels of hydraulic fracturing, they may have worse health outcomes, leading to an upward bias in the results. Therefore, I conduct a series of tests to further investigate the mechanisms behind the increases. I begin by testing subpopulations that are unlikely to be impacted by drilling-related migration, including women and Medicare and Medicaid recipients. If the increases in healthcare utilization were due to migration, I would not expect to find increases in these subpopulations. However, I find similar, if not larger, increases in these groups.

Next, I use data on public groundwater use to examine whether increases in healthcare utilization are linked to groundwater contamination from hydraulic fracturing. Specifically, I compare changes in healthcare utilization in counties with unconventional drilling near a public water source to those without. I find that counties with unconventional drilling within one kilometer of a public groundwater source experience larger increases in inpatient healthcare utilization than those without drilling within one kilometer of a public groundwater source. If compositional changes were driving the increases in healthcare utilization, this result suggests that either healthier individuals are more likely to move to locations without drilling near a public water source or unhealthy individuals are more likely to move to locations with drilling near a public water source. Both of these scenarios are unlikely since households probably do not know the exact location of their water source, and therefore this results also supports the hypothesis that compositional changes are not driving the results.

Additionally, increases in healthcare utilization in counties with drilling near public groundwater sources suggests groundwater contamination serves as a mechanism for pollution exposure. I find that healthcare utilization increases in counties with drilling within one kilometer of a groundwater source, but not in counties with drilling within three to five kilometers of a groundwater source, which aligns with research that shows that groundwater contamination occurs near the drilling site (Rozell and Reaven 2012; Shonkoff, Hays, and Finkel 2014). Finally, if air pollution alone were driving the health effects of hydraulic fracturing, then I would not expect differences in healthcare utilization in counties with unconventional drilling near groundwater sources relative to counties with unconventional further away from groundwater sources since both groups are impacted by air pollution from drilling. Altogether, these results suggest that healthcare utilization increases due to hydraulic fracturing entry, and this increase can partially be attributed to groundwater contamination.

This paper adds to the growing literature on externalities related to unconventional drilling methods. One line of research studies the effect of natural gas development on property values which suggests that proximity to an unconventional drilling site is a disamenity. Muehlenbachs, Spiller, and Timmins (2015) find that the perceived risk of groundwater contamination negatively affects the property value of homes near gas wells, and particularly homes that depend on private groundwater wells for drinking water. The authors find that the property values of homes that use publicly supplied water are not affected or may even benefit from proximity to a producing well through royalty payments. Although the authors do not estimate groundwater contamination, their work suggests that individuals internalize the potential risk of contamination, driving down property values of affected homes. In a similar study, Balthrop and Hawley (2017) study home prices that largely use public water and still finds that houses closer (within 3,500 feet) to hydraulically fractured wells sell at lower prices than houses farther away. The authors suggest various reasons for the negative effect, such as pollution, risk of earthquakes, or visual disamenities, but they do not identify a mechanism.

Environmental analysis on air and water contamination near shale development through hydraulic fracturing and horizontal drilling suggests that pollution concerns may be valid (Hildenbrand et al. 2016; Hill and Ma 2022; Litovitz et al. 2013; Osborn et al. 2011; Xu 2023). The excess of toxins near shale gas development locations leads to concerns about the health of individuals living within close proximity to hydraulic fracturing sites. Much of the literature focuses on pregnancy and infant health outcomes, and finds a relationship between pregnancy complications, poor infant health, and hydraulic fracturing (Currie, Greenstone, and Meckel 2017; Hill and Ma 2022; Hill 2024). For example, Currie, Greenstone, and Meckel (2017) compare infant health outcomes before and after the entry of a hydraulically fractured well for infants born within one kilometer and within one to three kilometers of the drilling site. The authors find that mothers who live within three kilometers of a hydraulic fracturing site give birth to infants with relatively worse health outcomes after the entry of hydraulic fracturing, including a twenty-five percent increase in the probability of a low birth weight. Hill and Ma (2022) confirm these results and additionally establish that drilling near an infant's public water source leads to an increased risk of shorter gestation and a low birth weight.

Black et al. (2021) conduct a systematic review of recent literature on the costs and benefits of unconventional drilling, focusing the economic, environmental, and health effects of unconventional drilling. The authors describe several epidemiological and public health studies that document a positive relationship between increased unconventional drilling and poor health outcomes. For example, Denham et al. (2019) finds an association between increased drilling and both genitourinary and skin-related hospitalizations. Willis et al. (2018) finds a twenty-five percent increase in pediatric asthma-related hospitalizations in the same quarter that a new unconventional gas well is drilled. McKenzie et al. (2017) finds that children between 5 and 24 who were diagnosed with acute lymphocytic leukemia were 4.3 times more likely to live near a high concentration of oil and natural gas wells. However, Black et al. (2021) point out that most of the causal analysis on the health effects of hydraulic fracturing focuses on children an infants while causal research on adult health outcomes is a gap in the the literature. A limited number of papers do use quasi-experimental techniques to identify the causal impact of hydraulic fracturing on respiratory illnesses. Peng, Meyerhoefer, and Chou (2018) use difference-in-differences analysis and finds an increase in between pneumonia hospitalizations for elderly patients in counties with unconventional drilling. Similarly, Xu (2023) finds that Texas counties with higher oil reserves have more respiratory-related hospitalizations after the entry hydraulic fracturing. However, prior research ignores many conditions predicted to be related to hydraulic fracturing. This paper adds to the causal work on the relationship between hydraulic fracturing and adult health outcomes by analyzing a broad spectrum of health conditions.

I also extend prior research by looking at healthcare utilization in an understudied region: the Eagle Ford Shale formation. Much of the prior causal investigation of the health effects of unconventional drilling uses data from the Marcellus Shale geological formation in Pennsylvania and New York.[2] According to Speight (2019), the geology of each shale formation is unique, meaning that the techniques used in each region differ and minerals, metals, and hydrocarbons released during the fracking process also vary in each region. One major difference between the two regions is that the Marcellus Shale formation produces a considerable amount of natural gas while the Eagle Ford Shale formation and produces both oil and natural gas. Producing both resources releases methane and other pollutants, and thus the natural gas must be flared to reduce the most toxic compounds. The flaring process, however, contributes to greenhouse gas emissions.

Recovery of hydraulic fracturing fluid also differs between the Marcellus Shale region and Eagle Ford; most of the fluid in the Marcellus Shale region returns to the surface after use, while more of the fluid may remain underground in Eagle Ford because the region is more water-deprived (Speight 2019). Finally, drilling in the Marcellus Shale formation began much earlier than Eagle Ford, as portions of the natural gas reservoirs were accessible using conventional drilling techniques. Eagle Ford is inaccessible using conventional oil and gas drilling methods, and was only discovered in 2008. After discovery, the region experienced a sharp and dramatic increase in both oil and natural gas drilling in 2010. This sharp change in exposure to hydraulic fracturing and oil and gas drilling provides a unique opportunity to identify the health effects of hydraulic fracturing in the region.

In the next section, I describe the history of the Eagle Ford Shale formation, the process of hydraulic fracturing and horizontal drilling, and potential pathways for negative health outcomes. In section three, I outline the data and provide descriptive statistics. Section four reviews my methods and results. Section five concludes.

2. One exception is Willis et al. (2021), which looks at the effects all types of drilling on birth outcomes within the state of Texas, though their health outcomes data end in 2009.

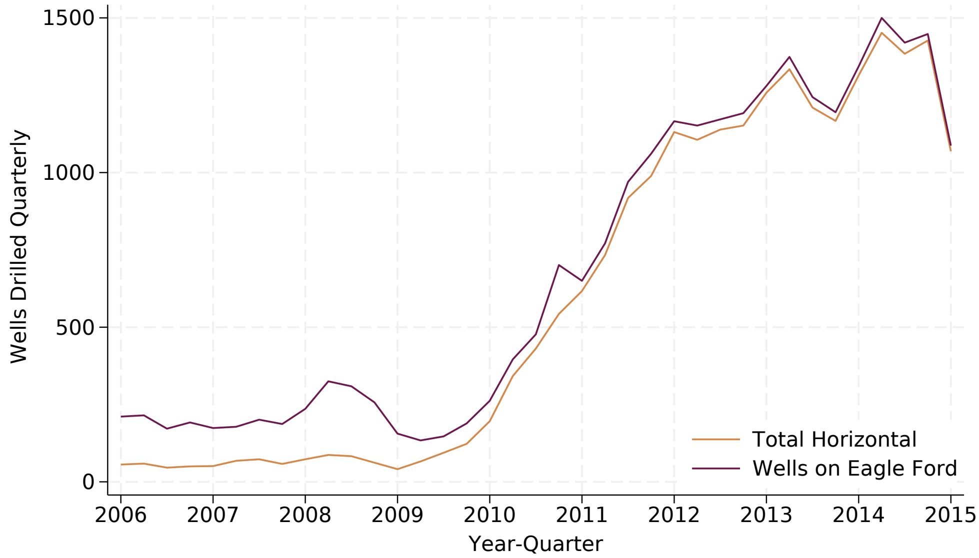
2 Background

2.1 UNCONVENTIONAL DRILLING

Unconventional drilling refers to newer methods of extracting oil, gas, and other substances from underground reservoirs that were inaccessible using previous technologies. While conventional drilling involves drilling vertically into oil or natural gas reservoirs, unconventional drilling uses alternative methods to release oil and natural gas trapped in tight geological formations. Two unconventional drilling methods, often used in combination, are hydraulic fracturing and horizontal drilling. The process of unconventional drilling involves several phases between preparation and production (Hill 2018). Once a drilling site is located and permitted, well pad preparation begins, which includes clearing and preparing the area for drilling. Next, wells are drilled on the well pad, which can take up to two months for horizontally drilled wells. After drilling, hydraulic fracturing takes place for around seven days to stimulate oil and natural gas production. The hydraulic fracturing process involves injecting water, chemicals, and/or sand into rock formations to break up the rock and release the oil and gas, and conversations with oil and natural gas engineers suggest that hydraulic fracturing takes place laterally, within a two mile radius of a well pad. After fracturing, there is a period of flowback. During this period, the injected fluid mixed with oil or gas, called "flowback fluid" or "produced water," is extracted and either reused or stored in tanks until it can be sent to injection disposal wells (Kondash, Albright, and Vengosh 2017).

Cost-effective hydraulic fracturing and horizontal drilling began in the late 1990s and early 2000s in the Barnett Shale region of Texas. By 2008, Texas produced more than seventy percent of shale gas in the United States (Gilmer and Kerr 2010). In addition to the Barnett Shale region, unconventional drilling technology opened new opportunities in several other Texas geological formations including the Permian Basin, the Haynesville Shale formation, and the Eagle Ford Shale formation. According to Gilmer, Hernandez, and Phillips (2012), drilling began in the Eagle Ford region in 2008 in La Salle county. Production subsided in 2009 as a result of declining oil and gas prices and the Great Recession, but picked up again in 2010 and grew rapidly, making it an ideal setting to better understand the health consequences of unconventional oil and gas drilling. As explained above, oil and natural gas in the Eagle Ford Shale formation was not accessible prior to hydraulic fracturing, and therefore the new technology allowed for a sharp increase in production after discovery (Speight 2019). By 2021, Eagle Ford was one of the most productive formations in Texas. Figure 1 shows the increase in oil and gas well permits on Eagle Ford after the entry of hydraulic fracturing and horizontal drilling, and captures the small increase in well permits after 2008 with a much larger increase after 2010.

Figure 1: Total wells drilled on Eagle Ford over time



Notes: Figure 1 shows the number of wells drilled each year-quarter in Eagle Ford during the sample period. 'Wells on Eagle Ford' includes all wells drilled during the period. 'Total Horizontal' represents all horizontally drilled wells, which are the wells most likely to use hydraulic fracturing. Horizontal wells are a subset of the 'Wells on Eagle Ford.' *Source: Authors Calculations using Texas Railroad Commission (2007–15) data.*

2.2 HYDRAULIC FRACTURING AND HEALTH

The process of extracting oil and natural gas through both conventional and unconventional methods gives rise to health concerns, primarily through air and water pollution (Shonkoff, Hays, and Finkel 2014; Speight 2019). The use of large volumes of fracking fluid, which contain potentially hazardous chemicals and could leach toxic substances from the shale, distinguishes unconventional methods from conventional methods. This water-intensive process contributes to concerns about drinking water contamination, and research finds high concentrations of hazardous chemicals in drinking water near hydraulic fracturing sites (Alawattegama et al. 2015; Hill 2018; Rozell and Reaven 2012; U.S. EPA 2016; Wollin et al. 2020). In a review of the relationship between drinking water contamination, hydraulic fracturing, and health, Shonkoff, Hays, and Finkel (2014) cite studies that find elevated concentrations of methane (Jackson et al. 2013; Osborn et al. 2011), heavy metals and total dissolved solids (Fontenot et al. 2013), and volatile organic compounds (VOCs) including benzene, toluene, ethylbenzene, and xylenes (BTEX) (DiGiulio et al. 2011; Gross et al. 2013; Hildenbrand et al. 2015) in drinking water near unconventional drilling sites.

Exposure to these chemicals can have significant health consequences. For example, Webb et al. (2018) draw attention to serious neurological health effects of chemicals used in the hydraulic fracturing process, such as heavy metals, particulate matter (PM), BTEX, and endocrine disrupting chemicals (EDCs), all of which are associated with poor neurodevelopment, cognitive, and mental health outcomes through drinking water exposure. BTEX chemicals are also known to cause damage to the nervous and endocrine systems, birth defects and low birth weight, respiratory conditions, abnormal heart rhythms, and blood diseases, including aplastic anemia and acute myelogenous leukemia (Epstein 2017). Shrestha et al. (2017) examine the health effects of chemicals found in groundwater near hydraulic fracturing sites, and finds chemicals related to gastrointestinal distress, kidney disease or kidney damage, skin sensitivity or sores, cancer, dehydration, and liver disease, in addition to the conditions mentioned above. Chemicals used in drinking water treatment can also interact with chemicals used in hydraulic fracturing, transforming them into more toxic compounds. For instance, Wollin et al. (2020) explain that chemicals used to disinfect drinking water can cause oxidation of bromide used in hydraulic fracturing, creating the chemical bromate, which is mutagenic and carcinogenic.

Although the health effects of some chemicals used in the hydraulic fracturing process are known, there remains a lot of uncertainty about both the chemicals used in the process and the health consequences of those chemicals. Colborn et al. (2011) compile a non-exhaustive list of products used in the fracturing process and find potential health effects from 353 chemicals identified by the Chemical Abstract Service members which could affect health through either air or water pollution. The chemicals with known health effects can impact the skin, eyes, kidneys, and brain as well as the nervous, immune, endocrine, respiratory, and cardiovascular systems. Twenty-five percent of the chemicals identified could cause cancer or mutations, and thirty-seven percent are candidate endocrine disrupting chemicals (EDCs), which can effect reproductive systems and fetal and early childhood development. In a 2016 study on the impact of hydraulic fracturing on drinking water, the EPA finds that hydraulic fracturing can impact drinking water and considers the health effects of over 1,600 chemicals used in the hydraulic fracturing process. The authors note that the health effects of many of the chemicals used in the process are unknown and many are not regulated or measured in drinking water (U.S. EPA 2016).

In a cautionary public health paper, Finkel and Law (2011) explain that the timing of health effects after exposure to the chemicals used in hydraulic fracturing is largely understudied. For example, a review of clinical, observational, and epidemiological studies on EDCs suggests that numerous factors, such as age at exposure, exposure dosage, and mixtures of contaminants can all impact the timing between exposure and disease, and some endocrine disrupting effects may appear decades after exposure or even trans-generationally. Nonetheless, researchers find increased incidence of diseases in areas with exposure to hydraulic fracturing, such as an association between babies born with congenital heart defects whose mothers lived near unconventional wells (McKenzie et al. 2014) or a higher rate of hematological cancer among children who live closer to oil and gas development (McKenzie et al. 2017).

Water contamination can affect either groundwater or surface water, however because groundwater contamination is local to the drilling site, I focus on groundwater contamination in this analysis. Rozell and Reaven (2012) suggest that water contamination can occur through spills during fracking fluid transportation, leaks in the well casing, leaks in fractured shale, discharge at the drilling site, or wastewater disposal. Surface water contamination, which includes contamination of streams, wetlands, lakes, and oceans, mainly occurs through spills and leaks. Exposure to surface water contamination can depend on environmental factors such as water flow and elevation (Bonetti, Leuz, and Michelon 2021; Hill and Ma 2022). Because of this, identifying the exact exposure area for surface water contamination is difficult. On the other hand, groundwater contamination, which includes contamination of water below the Earth's surface and soil water, most likely occurs at the water system's intake point, or wellhead. Some researchers suggest that groundwater contamination at the drilling site most likely occurs from surface spills in the pre-production period (Shanafield, Cook, and Simmons 2019), however additional studies also indicate that groundwater contamination may occur later in the well lifespan as a result of gas and fluid migration due to structural impairment or site discharge during the flowback period (Rozell and Reaven 2012; Shonkoff, Hays, and Finkel 2014).

Despite the focus on groundwater contamination, I do not rule out the possibility that air pollution also contributes to negative health effects. Air pollution can occur near the drilling site from both conventional and unconventional drilling methods, primarily from production activities, however transportation of fluid, equipment, and produced oil and natural gas to and from the fracturing site also leads to higher emissions throughout the entire region. Several studies evaluate air quality near drilling sites and find dangerous levels of air pollution related to both conventional and unconventional oil and natural gas production activities (McKenzie et al. 2012; Pétron et al. 2012; Roy, Adams, and Robinson 2014). In context of hydraulic fracturing, substantial air pollution comes from transporting the hydraulic fracturing fluid, leading to increased diesel particulate matter along the routes to the drilling sites (Shonkoff, Hays, and Finkel 2014, Speight 2019). A single well requires between two and five million gallons of water, which is transported in trucks with a capacity of only 3,000 gallons (Shonkoff, Hays, and Finkel 2014). Clay et al. (2017) also finds increased air pollution from transporting crude oil by rail. Therefore, while areas near oil and natural gas drilling are likely affected by air pollution, the boundaries of pollution exposure from increased traffic are less clear. Given that unconventional drilling-related air pollution affects counties on and off the shale formation, my results likely underestimate of the overall effect of hydraulic fracturing on healthcare utilization. I leave the exercise of identifying the impact of drilling-related air pollution from traffic to future research.

3 Data and Descriptive Statistics

3.1 HEALTHCARE DATA

Claim-level inpatient data from the Texas Health Care Information Collective (THCIC) provide the primary outcome data for this paper. All hospitals within the state of Texas are required to report claims, which are the invoices submitted to health insurance companies for payment, on discharged patients, and THCIC makes de-identified claim reports available for public use (Texas Health Care Information Collective 2007–15).[3] This rich data source includes information on each patient claim submitted by the hospital, the quarter and year the claim was submitted, details on the diagnosis and procedure codes associated with the claim, and patient demographic information.[4] Specifically, the data include hospital information, such as the hospital name, indicators for the facility type,[5] and categorical variables for the specialty unit[6] where the care occurred. I also observe patient demographic information for each claim including the patient's age range, gender, race, ethnicity, county, zip code, and payment information, such as insurance type. I use data between 2007 and 2015, starting one year prior to the entry of hydraulic fracturing in the Eagle Ford Shale formation. I also only including claims made from individuals with a county or zip code in the treatment or control groups. If workers temporarily move to Eagle Ford for labor opportunities, but do not establish residency in the area, they are excluded from the analysis.

I categorize the ICD-9[7] codes on each claim using the Healthcare Cost and Utilization Project's Clinical Classifications Software (HCUP CCS), which takes the over 14,000 ICD-9 diagnosis codes and 3,900 procedure codes and collapses them into clinically meaningful categories. Specifically, HCUP CCS categorizes diagnoses codes into a hierarchy of nested medical conditions. For example a medical claim may include a set of diagnosis codes that suggest the clinically meaningful diagnosis was "Essential Hypertension," which falls under the "Hypertension" category, which is nested under "Diseases of the Circulatory System." For most conditions, I use highest category in the hierarchy to test each medical category and I focus on health conditions that could be linked to hydraulic fracturing through pollution exposure according to prior literature described in Section 2.2. Specifically, I focus on mental health,[8] endocrine conditions, cancer, conditions of

3. Although data were available through 2016 at the time of this study, healthcare providers switched from ICD-9 to ICD-10 codes in October of 2015. Therefore, I use all claims made before this transition.

4. The public use files exclude claims made in hospitals located in a county with a population less than 35,000, or those in a larger county with less than 100 hospital beds. In addition, hospitals that do not seek insurance or government reimbursement are also exempt.

5. For example, this might include an indicator for a psychiatric facility or a rehabilitation facility.

6. Specialty units are the areas of the hospital dedicated to patients with specific needs. For example, the obstetric unit is the area of the hospital specifically for labor and delivery.

7. ICD-9 and ICD-10 codes are the World Health Organizations Ninth Revision of the International Classification of Diseases (ICD-9) diagnosis and procedure codes, which is the official coding system for diagnoses and procedures in the United States.

8. Mental health conditions include anxiety disorders, attention deficit disorders, depressive disorders, suicide and intentional self-harm, eating disorders, and substance abuse disorders.

the central nervous, respiratory, gastrointestinal, urological, circulatory, and integumentary systems, pregnancy and birth outcomes, and blood diseases (Colborn et al. 2011; Epstein 2017; Shonkoff, Hays, and Finkel 2014; Webb et al. 2018; Wollin et al. 2020; Garshick et al. 2008; Persico and Marcotte 2022; Shonkoff, Hays, and Finkel 2014). In addition, I test changes in 'General Illness,' which includes symptoms such as fever, nausea and vomiting, and general malaise that is not attributed to a specific condition. This list is not meant to be an exhaustive or exact list of conditions and exposure pathways, rather an analytical starting point to build upon. For more information on the conditions included in this paper, see Appendix A1.1.1.

3.2 UNCONVENTIONAL DRILLING TECHNOLOGY

I identify wells drilled within the time period using the Texas Railroad Commission (RRC) Drilling Permit Master data file combined with the RRC Geographic Information Systems (GIS) data files (Texas Railroad Commission 2007–15). The GIS data files include the latitude and longitude of both the surface and bottom of each well and the permit data include the spud date, which is the date drilling began on an individual well. I exclude well permit records that are missing location information or a spud date and I only use wells with a spud date between 2007 and 2015.

The data include several different types of wells and does not specifically identify hydraulic fracturing sites. Because my analysis focuses on unconventional oil and natural gas drilling, I include sites for oil, gas, and brine.[9] I also include injection sites, which are sites where oil, gas, and/or contaminated water is stored underground after it is produced or used. I exclude other types of drilling locations, such as water observation wells and geothermal wells. I then map the boundaries of the shale formation to identify which wells lie within a two mile radius of the formation. Ultimately, I find just over 20,000 wells drilled on the Eagle Ford Shale formation within the time period.

Treatment timing is determined by the spud date of the first well drilled in a geographic area after 2008, which makes two key assumptions. First, I assume that all wells drilled in the Eagle Ford region after 2008 used hydraulic fracturing technology. Speight (2019) states that oil and natural gas in the Eagle Ford Shale formation are inaccessible using conventional drilling techniques. However, I test this assumption by identifying all horizontally drilled wells as those that are either 1) identified as a horizontal well in the permit data or 2) have different surface and bottom locations, since hydraulic fracturing occurs along horizontally drilled wells. Figure 1 shows the increase in oil and gas wells on the Eagle Ford Shale formation and the increase in horizontally drilled wells on the Eagle Ford Shale formation. The increase in wells in the region is driven almost entirely by horizontally drilled wells. Second, I assume that hydraulic fracturing takes place shortly after the well is drilled. Conversations with engineers in the field suggest that this assumption is

9. Brine mining is used to extract chemicals and compounds, such as lithium, from brine deposits.

likely true, although hydraulic fracturing may be delayed after a well is drilled. Since hydraulic fracturing stimulates the oil production, producers may delay fracturing until oil prices are higher, for example.

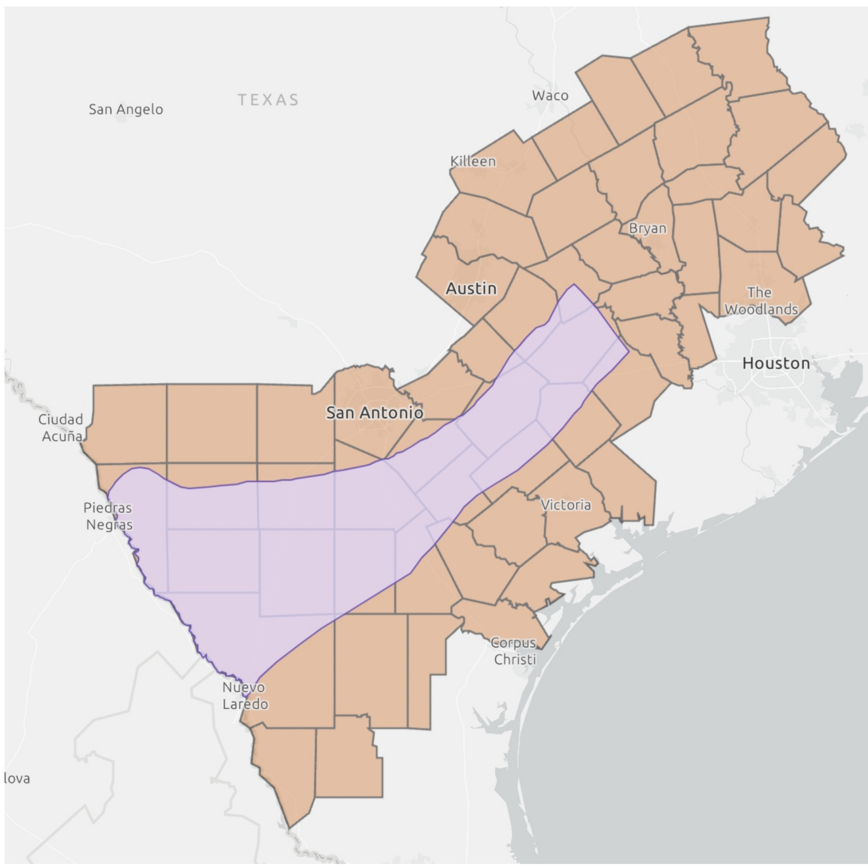
To ensure the robustness of my results, I conduct the analysis at both zip code and county levels, given the complementary strengths and weaknesses of each. My primary analysis is at the zip code level and I identify treated Zip Code Tabulation Areas (ZCTA)[10] using the boundaries of the shale formation. All zip codes that experienced drilling within the time period and that intersect with the shale formation are included in the treatment group. Treatment timing is staggered, with unconventional drilling entry dates between 2008 and 2012 at the zip code level. For the control group, I include all zip codes located in the 27 counties that the Texas RRC identifies as on the Eagle Ford Shale formation that do not experience drilling and/or do not intersect with the shale formation. I also include zip codes located in the 29 counties that border the Eagle Ford counties. Figure 2 shows the counties used in the estimation in orange and highlights the Eagle Ford Shale formation in purple.

The advantages of zip-code level analysis are twofold. First, proximity to the hydraulic fracturing site likely impacts exposure, and examining the effects at the zip code level allows for closer identification of individuals who live near drilling activity compared to county-level. In addition, zip-code-level exposure to hydraulic fracturing was staggered, with entry dates spanning from 2008 to 2012. Estimating the effects at the zip code level, therefore, leverages both the effect of exposure proximity and timing, likely producing more precise estimates.

The drawback of conducting the analysis at the zip-code level is that THCIC masks zip codes identifiers for patients living in zip codes with a small number of claims, making some utilization rates artificially low. Therefore, I also estimate the effects at the county level. Of the 27 counties that the Texas RRC identifies as on the Eagle Ford Shale formation, 19 counties experienced drilling within the time period, which comprise the treatment group. The other eight counties and the 29 border counties make up the control group. At the county-level, all treated counties were treated in the first quarter of 2008. However, according to Gilmer, Hernandez, and Phillips (2012), while oil and gas permitting increased in 2008, drilling expenditure spiked in 2010, suggesting that producers delayed production during 2009 when oil and gas prices were low, and expanded production in 2010 as oil prices increased. Therefore, I choose 2010 as the effective entry date due to the rapid expansion of hydraulic fracturing after 2010.

10. Zip codes are groups of street networks and buildings used by the United States Postal Service for mail delivery, while ZCTA's are designed by the United States Census Bureau for geographic analysis.

Figure 2: Eagle Ford Shale formation



Notes: This figure shows the Eagle Ford Shale formation in southern Texas and highlights counties included in both the treatment and control. The purple shape overlaying the counties is the Eagle Ford Shale formation; all counties that intersect with the shale formation are included in the treatment group. *Source: Authors Calculations. Generated using ArcOnline.*

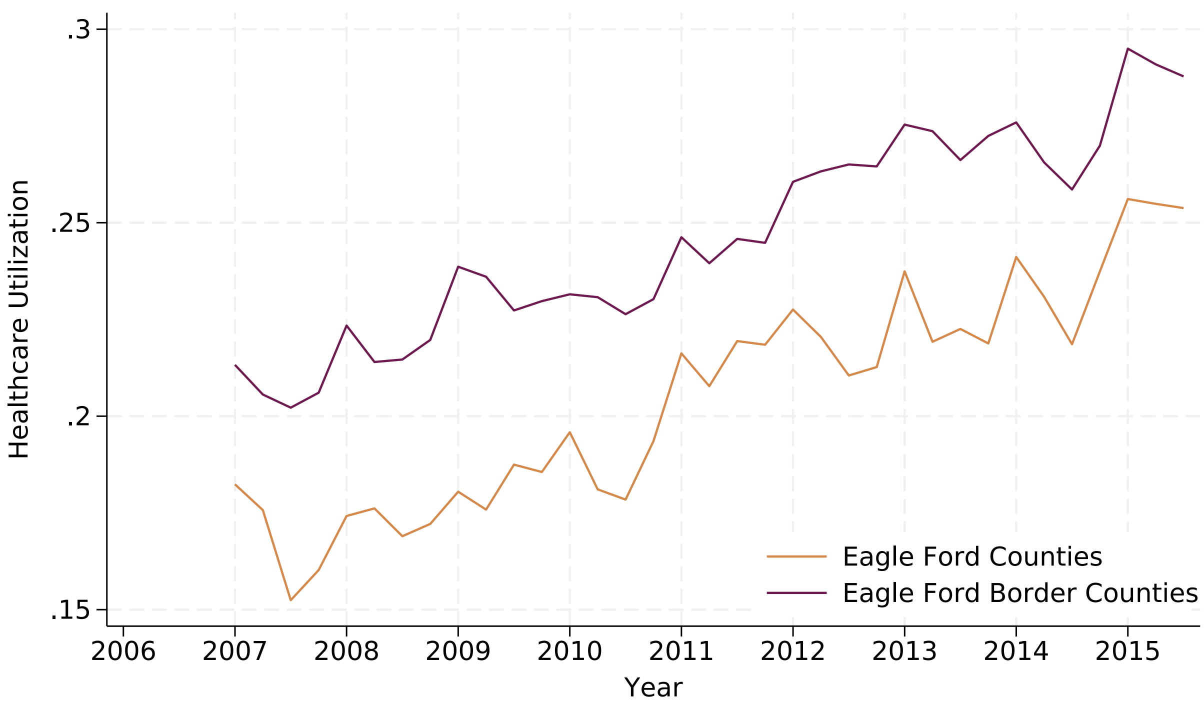
3.3 DESCRIPTIVE STATISTICS

My outcome variable of interest is the geographic-level healthcare utilization rate for each condition considered in this paper. I calculate the healthcare utilization rate as follows:

$$Y_{ct} \equiv \frac{\sum_{claim \in HC} claim_{ct}}{Pop_{cy}} \tag{1}$$

Where utilization, Y_{ct}, is defined in zip-code (county) c in year-quarter t as the sum of total claims, $claim_{ct}$, within healthcare category type HC in zip-code (county) c in year-quarter t divided by the population of zip-code (county) c in year y. Population is calculated annually using the U.S. Census Bureau Decennial Census total at the zip-code-level or County Population Totals at the county level.

Figure 3: Healthcare utilization in the Eagle Ford and border counties



Notes: Figure 3 plots the average quarterly inpatient healthcare utilization for counties located on the Eagle Ford Shale formation and counties just outside of Eagle Ford using the THCIC data and the U.S. Census Bureau County Population totals. The utilization rate for each healthcare category is calculated by taking the sum of all fracking-related claims considered in this study made by residents of a county each quarter divided by the annual population of that county. *Source: Authors Calculations using data from Texas Health Care Information Collective (2007–15).*

Table 1 shows the average healthcare utilization for each condition before and after the treatment date in 2010 for the treated Eagle Ford counties and the counties in the control group. 'All conditions' is the sum of all claims made for conditions potentially related to hydraulic fracturing through pollution as described in

Section 2.2, including circulatory conditions, general illness, mental health and endocrine conditions, cancer, conditions of the central nervous system, respiratory conditions, gastrointestinal conditions, urological conditions, pregnancy complications, dermatological conditions, perinatal conditions, congenital abnormalities, and blood disease. Utilization increased from the pre-hydraulic fracturing to the post-hydraulic fracturing periods in all categories in the Eagle Ford counties, and similarly increased in nearly all categories in the control counties. However, the utilization rate for both pregnancy complications and perinatal utilization decreased in the control counties after 2010, while utilization increased for these categories in the treatment counties.

Table 1: Average, county-level healthcare utilization pre- and post-hydraulic fracturing entry

Health Condition	Eagle Ford; Pre-2010 (1)	Eagle Ford; Post-2010 (2)	Border Pre-2010 (3)	Border Post-2010 (4)
All Conditions	0.174	0.221	0.219	0.260
	(0.052)	(0.062)	(0.061)	(0.072)
Circulatory	0.049	0.060	0.061	0.071
	(0.017)	(0.021)	(0.019)	(0.022)
General Illness	0.005	0.007	0.006	0.008
	(0.003)	(0.004)	(0.003)	(0.003)
Mental Health	0.016	0.020	0.022	0.026
	(0.005)	(0.006)	(0.006)	(0.007)
Endocrine	0.034	0.046	0.042	0.053
	(0.011)	(0.013)	(0.012)	(0.016)
Cancer	0.008	0.009	0.009	0.010
	(0.003)	(0.003)	(0.003)	(0.003)
Central Nervous	0.013	0.016	0.016	0.019
	(0.004)	(0.005)	(0.005)	(0.005)
Respiratory	0.018	0.023	0.024	0.029
	(0.007)	(0.008)	(0.009)	(0.009)
Gastrointestinal	0.012	0.015	0.016	0.018
	(0.004)	(0.005)	(0.005)	(0.006)
Urological	0.013	0.016	0.014	0.017
	(0.004)	(0.004)	(0.004)	(0.005)
Pregnancy Complications	0.003	0.005	0.006	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Dermatological	0.003	0.004	0.004	0.004
	(0.001)	(0.001)	(0.002)	(0.002)
Perinatal	0.003	0.003	0.003	0.003
	(0.002)	(0.001)	(0.002)	(0.001)
Blood Disease	0.008	0.010	0.009	0.011
	(0.003)	(0.003)	(0.003)	(0.003)
Congenital Abnormalities	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
Observations	228	437	444	851

Notes: Standard deviations in parenthesis.
Each column summarizes the average healthcare utilization for the conditions considered in this paper, calculated as the number of claims for the condition divided by the total population. Columns (1) and (2) show utilization for the Eagle Ford counties with drilling pre- and post-2010, respectively, while Columns (3) and (4) show utilization for the border counties and Eagle Ford counties without drilling. 'All Conditions' includes the sum of healthcare claims for each condition studied in this paper divided by the geographic-area population.

Figure 3 plots the average, quarterly utilization rate for all conditions related to hydraulic fracturing in the Eagle Ford counties and in the border counties. In 2007, prior to the entry of hydraulic fracturing, the Eagle Ford counties and the border counties both experienced a decrease in utilization in the early part of the year, followed by a slight increase leading up to 2008. In the early years, the healthcare utilization rate in the Eagle Ford counties is lower than utilization in the border counties with the gap between the two groups narrowing in 2011.[11]

4 Methods and Results

4.1 HEALTHCARE UTILIZATION

I begin by estimating the general effect of hydraulic fracturing entry on healthcare utilization using a difference-in-differences estimation strategy. At the county-level, the treatment timing is the same for all counties, and therefore I used a standard, two-way fixed-effects (TWFE) estimator. My specification corresponds to the equation:

$$Y_{ct} = \beta_0 + \beta_1 1[Post]_t \times 1[E\text{-}Ford]_c + \eta_c + \tau_t + \epsilon_{ct} \tag{2}$$

Where Y_{ct} is the healthcare utilization rate in county c in year-quarter t. $1[Post]$ is an indicator equal to one if year-quarter t is after the 2010 treatment date and $1[E\text{-}Ford]$ is an indicator equal to one if the county lies on the Eagle Ford Shale formation and experienced drilling between 2010 and 2015. The parameter of interest, β_1, captures the differential impact of hydraulic fracturing entry on healthcare utilization in treated counties compared to untreated counties. This estimation method assumes that healthcare utilization in counties near Eagle Ford without hydraulic fracturing represent changes in healthcare utilization that would have happened if hydraulic fracturing never entered the Eagle Ford region. With time and county fixed effects, β_1 is identified by the expansion of hydraulic fracturing technology in 2010 in the Eagle Ford region.[12]

At the zip-code-level, I take advantage of the variation in treatment timing by using the estimator described in Callaway and Sant'Anna (2021). Zip codes are considered treated if a well was drilled in a zip code within two miles of the shale formation. Similar to the county level, treatment is binary and an

11. A few counties on the border of Eagle Ford include highly populated, urban areas, which likely have better access to healthcare and other resources. This leads to potential concerns about including the highly populated counties in the control group, particularly if individuals in these counties have improved access to healthcare resources around the time of hydraulic fracturing entry. Given the passage of the Affordable Care Act (ACA) in 2010, including these counties could confound the results. However, if I exclude these counties, the results are similar.

12. To test that the treated and control areas would have evolved in parallel without treatment, I create an event study plot that corresponds to the zip-code-level analysis. The event studies plot year-quarter level estimates, capturing the difference in healthcare utilization in the Eagle Ford counties relative to the border counties. The event plots, displayed in Appendix **??** and A2, are somewhat noisy, but generally show little evidence of differential trends between the treated counties on the Eagle Ford Shale formation and the control counties outside of Eagle Ford prior to treatment, suggesting that these assumptions likely hold in this situation.

absorbing state. This estimator calculates the average treatment effect on the treated (ATT) in zip-code-year-quarter cohorts, or the average treatment effect on the cohort of Eagle Ford zip-codes first treated in year-quarter yq. I include the not-yet treated zip-codes in the control group.

Column (1) of Table 2 reports the β_1 coefficients of the county-level estimation, capturing the average change in the county-level healthcare utilization rate after hydraulic fracturing. Column (2) of Table 2 reports the average healthcare utilization rate for each condition in the treatment group before the entry of hydraulic fracturing. While most of the changes in healthcare utilization at the county-level are not statistically distinguishable from zero, I find an increase in healthcare utilization related to adverse birth outcomes and pregnancy complications, consistent with Currie, Greenstone, and Meckel (2017) and Hill and Ma (2022). I estimate a 27 percent increase in utilization for pregnancy complications, a 15 percent increase in utilization for perinatal conditions,[13] and an 14 percent increase in utilization for congenital abnormalities in the Eagle Ford counties relative to the border counties. However, this result could be due to an overall increase in births related to the oil and natural gas production boom (Kearney and Wilson 2018), which would also lead to an increase in negative birth outcomes. I test Equation 2 using claims related to births as the outcome variable and I find a statistically significant increase of 0.0014 (SD 0.0003), which translates to a 20 percent increase in birth-related claims relative to the pre-fracking mean. Thus, increases in pregnancy related conditions may be driven by an increase in pregnancy and births in the region.

Column (3) of Table 2 displays the coefficients from the zip-code level estimation, revealing larger increases in the healthcare utilization rate for treated zip codes.[14] While most changes at the county-level are not statistically distinguishable from zero, estimates at the zip code level reveal significant increases in the utilization rate for circulatory conditions (12 percent), cancer (14 percent), conditions of the central nervous system (20 percent), dermatological conditions (50 percent), gastrointestinal conditions (18 percent), blood diseases (29 percent), and urological conditions (36 percent). Some of these estimates are large, but represent a small increase in actual claims because the conditions are rare. For instance, only two inpatient, dermatological claims were submitted per 1000 residents prior to hydraulic fracturing, and therefore a 50 percent increase in dermatological utilization represents one additional claim per 1000 residents.

I do not find increases in pregnancy complications, perinatal conditions, or congenital abnormalities at the zip-code-level. The THCIC data do not provide zip-code-level identifiers for conditions with infrequent claims within a particular zip code to protect patient privacy. This means that some zip codes appear to have zero claims, when in reality they may have had very few claims. This may prevent identification of an

13. Perinatal conditions include respiratory conditions, birth trauma, and low birth weight.

14. I drop one outlier zip code with late treatment timing since the lack of post-treatment observations for these group leads to noisy estimates and a violation of the parallel trends assumption. However, leaving these zip codes in the estimation does not change the estimates.

Table 2: Change in healthcare utilization

	County		Zip Code	
Health Condition	β (1)	Sample Mean (2)	β (3)	Sample Mean (4)
All Conditions	0.0055	0.174	0.022**	0.150
	(0.011)		(0.010)	
Circulatory	0.0009	0.049	0.005*	0.041
	(0.003)		(0.003)	
General Illness	0.0004	0.005	0.001	0.004
	(0.001)		(0.001)	
Mental Health	0.0008	0.016	0.002	0.014
	(0.001)		(0.001)	
Endocrine	0.0006	0.034	0.003	0.029
	(0.002)		(0.002)	
Cancer	0.0003	0.008	0.001**	0.007
	(0.001)		(0.001)	
Central Nervous	0.0003	0.013	0.002**	0.010
	(0.001)		(0.001)	
Respiratory	0.0002	0.018	0.001	0.016
	(0.001)		(0.001)	
Gastrointestinal	0.0008	0.012	0.002**	0.011
	(0.001)		(0.001)	
Urological	0.0001	0.013	0.004**	0.011
	0.001)		(0.002)	
Pregnancy Complications	0.0008**	0.003	0.000	0.004
	(0.000)		(0.000)	
Dermatological	0.0002	0.003	0.001***	0.002
	(0.000)		(0.000)	
Perinatal	0.0005***	0.003	0.000	0.003
	(0.000)		(0.000)	
Blood Disease	0.0004	0.008	0.002***	0.007
	(0.000)		(0.000)	
Congenital Abnormalities	0.0001**	0.001	0.000	0.001
	(0.000)		(0.000)	
Observations	1,960	228	13,720	167

*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.

Notes: Table 2 reports the county-level β_1 coefficients from Equation 2 in Column (1). Column (3) reports the analogous estimates from the staggered difference-in-differences specification at the zip-code-level. Columns (2) and (4) give the pre-fracking entry sample mean at the county- and zip-code- levels respectively. 'All Conditions' includes all healthcare claims for each condition studied.

effect at the zip-code-level due to an abundance of observations with zero claims.

One potential concern about comparing Eagle Ford areas to the border areas is that the composition of the population of Eagle Ford may change in response to new opportunities related to oil and natural gas production in the region. Therefore, the health effects estimated by this specification may be capturing changes in the composition of people living in the region, rather than the negative effects of pollution exposure to bystanders. Wilson (2022) estimates that individuals who move for fracking-related employment opportunities account for an increase of less than two percent of the total population in fracking counties in the South, which suggests that these concerns affect a small, but non-zero, percentage of the population.

The demographic groups that move to fracking counties tend to be unmarried men under the age of 65 (Wilson 2022). Therefore, to address concerns about the results capturing compositional changes, I repeat Equation 2 for women and individuals with insurance through Medicaid or Medicare, which are subpopulations less likely to be impacted by compositional changes related to hydraulic fracturing. If the results in Table 2 were driven by an influx of young men drawn to the region by the hydraulic fracturing boom, I would not expect to see similar effects among these subpopulations. However, among conditions that respond to hydraulic fracturing, the estimated increases in utilization for these groups are similar to the results in Table 2, and many are larger, particularly for Medicare and Medicaid recipients. Utilization for all fracking-related healthcare conditions increased by over 20 percent for Medicare and Medicaid recipients.[15] The results are included in Appendix Tables A4, A5, A6.

4.2 WATER CONTAMINATION

To more rigorously evaluate whether compositional changes account for the effects shown in Section 4.1 and to investigate groundwater contamination as a pollution exposure pathway, I examine how drilling near a public groundwater source affects healthcare utilization. In Texas, drinking water is distributed from a water source location to a set of households by a community water system. The geographic boundaries that contain the households on system are known as the public water service area, or PWSA. The water source location is typically located near, but not always within the boundaries of, the PWSA. Water can also be purchased from one source location and served to another PWSA. Therefore, households are served drinking water from water sources that may or may not be near their home location. This lends it self to a natural experiment to test the impact of hydraulic fracturing near source water locations on healthcare utilization. In this experiment, geographic areas with households served water sourced from a location with

15. One particularly striking result is an 82 percent increase in urological utilization for Medicaid recipients. However, this is likely due to compositional changes in Medicaid coverage for kidney disease related treatments due to increased kidney disease in the full population. Table 2 shows that urological conditions respond to the entry of hydraulic fracturing, and individuals with kidney disease are eligible for Medicaid coverage for treatment of kidney disease.

unconventional drilling nearby are considered treated, and households served water sourced from a location without unconventional drilling nearby act as a control.

Although water can be drawn from a surface water source, such as rivers or water reservoirs, or a groundwater source, like an aquifer, I focus on groundwater source locations since groundwater contamination is local to the drilling site, as explained in Section 2.1. I use data provided by the Texas Water Development Board (TWDB), which includes annual information on the public water service area and the name and the location of their the water source. The data also include and the county and number of individuals served by each PWSA. Using the latitude and longitude of each oil and natural gas well and each public water source, I identify all oil and natural gas wells drilled within one, three, and five kilometers of a public groundwater source location. I assume that if oil or natural gas drilling occurred within the radius, individuals served water from the water source were potentially exposed to water contaminated by hydraulic fracturing.

These data allow me to exploit variation in potential exposure to drinking water contamination to test if the health effects observed in Section 4.1 are 1) partially linked to drinking water contamination and 2) not likely a result of compositional changes in the population. Ideally, I would like to use the ZCTA as the geographic unit to precisely capture the effect of drilling near a public water source. However, PWSA's do not follow ZCTA geographies and therefore I aggregate this analysis at the county level. A county is treated if a PWSA within the county sourced water from a public groundwater source with unconventional drilling nearby.

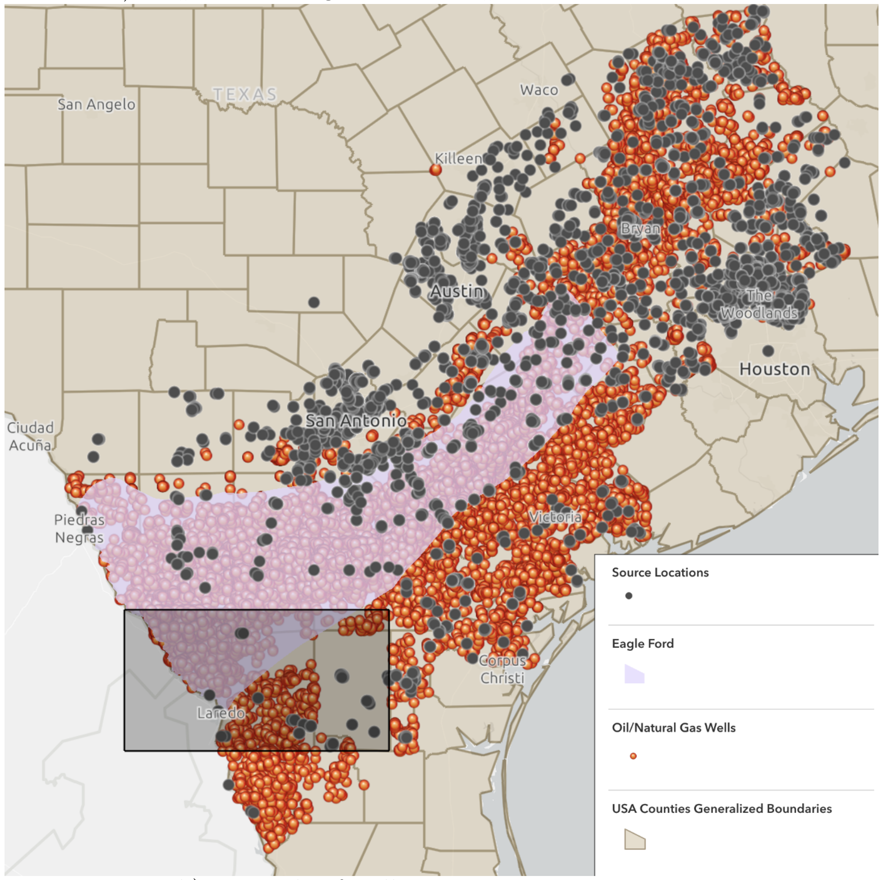
Figure 4 offers a visual representation of the oil and natural gas well locations and the water source locations to describe the identification strategy. Panel (a) of Figure 4 maps the Eagle Ford Shale region and marks the shale formation in purple. Each oil or natural gas well in the region is represented by an orange point and unconventional wells are the orange points that lie within the shale formation. The gray points are water source locations. Overall, there are 439 public water sources in the sample and 83 experience drilling within one kilometer of the water source. 18 of these locations are on the shale formation, for a total of ten counties treated at the one kilometer level. At the three kilometer level, 16 counties are considered treated with a total of 49 wells drilled on the shale formation. The same 16 counties are treated at the five kilometer level, however a total of 70 wells are drilled on the shale formation. Treatment timing is also often earlier at the five kilometer level.[16]

Panel (b) focuses on Webb County in the section of the formation highlighted by the gray box in Panel (a). The northern part of Webb County is located on the Shale formation and the southern part of the county is outside of the formation. This map highlights the boundaries of the PWSA's with a blue polygon.

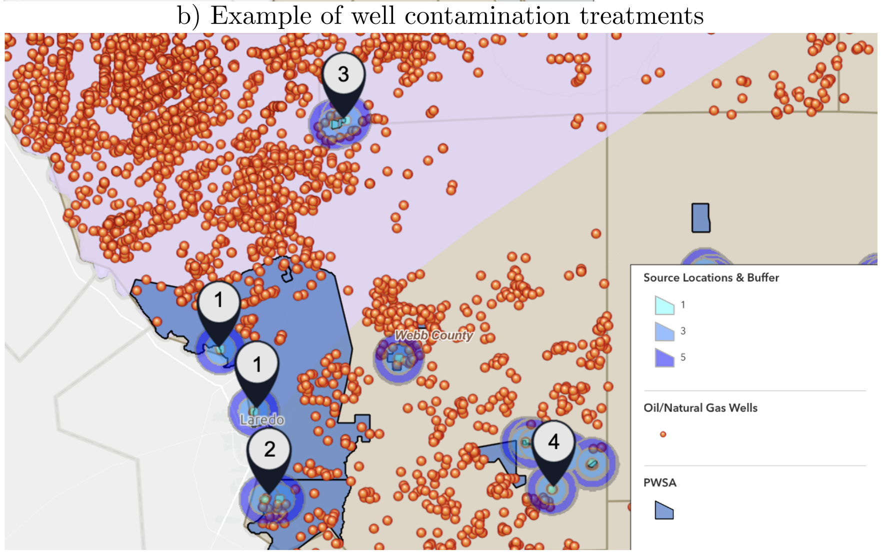
16. A total of 190 water sources experience drilling within three kilometers and a total of 272 water sources experience drilling at the five kilometer level.

Figure 4: Oil and gas wells near Public Water Sources

a) Map of oil and gas wells and public water source wells



b) Example of well contamination treatments



Notes: Panel (a) plots (in orange) all oil and natural gas wells on and off the shale formation drilled within the treatment period. The gray points represent water sources. The purple area is the Eagle Ford Shale formation; wells within the formation are considered to be unconventional wells. The bottom panel focuses on the gray box in Panel (a) to show the one, three, and five kilometer buffers drawn around each PWS as an example of how treated water sources are identified. *Source: Authors calculations using data from Texas Railroad Commission (2007–15) and Texas Water Development Board (2023). Generated using ArcOnline.*

Each water source and the one, three, and five kilometer buffers drawn around them are identified by the bullseye shapes.

I mark a few water sources to illustrate the identifying variation used in this analysis. The markers labeled 1 both provide water for the adjacent PWSA and are the only two water sources in Webb County located on the shale formation. There are no oil or natural gas wells within the one-, three-, or five-kilometer buffers drawn around the source locations. Notably, several oil and natural gas drilling sites exist within the boundaries of the PWSA, and therefore households within the PWSA may be exposed to air and noise pollution. Yet, their water source is not classified as contaminated under the treatment definition used in this analysis.

Marker 2 highlights two water sources serving a separate PWSA. Although there is drilling activity within one kilometer of this water source—indicating potential exposure to contamination from conventional drilling—the oil and natural gas wells nearby are not on the shale formation. As a result, this water source is also not classified as contaminated by unconventional drilling.

Finally, marker 3 identifies a cluster of water sources in La Salle County that serve a small PWSA straddling the border between Webb and La Salle Counties. This location has unconventional drilling sites within three and five kilometers of the water source. Consequently, Webb County is considered treated at the three- and five-kilometer levels, but not at the one-kilometer level. Note that this is a blunt measure of treatment since fewer than 90 households in Webb County are served by this PWSA. For context, just over one percent of a county's population was exposed to contaminated drinking water from drilling within one kilometer of a water source on average, with a maximum of eleven percent.

One potential issue with this analysis is that households may know where their public water source location is, observe that there is drilling nearby, and change their water consumption behavior. However, there are several reasons to believe that households may not be aware of their water source location. First, marker 4 depicts an example of a water source located over five kilometers from the PWSA boundary, demonstrating that water source locations may be far away from the PWSA where the water is distributed. Next, many PWSAs have multiple water sources, like those identified by the markers labeled 1, and it may be difficult for households to know how much of their water supply comes form each location. Finally, some PWSA's purchase water from water sources much further away. Thus, it is unlikely that households know the location of their water source.

To investigate how increased water contamination affects healthcare utilization, I use the the estimator in Callaway and Sant'Anna (2021) to causally identify the effect of exposure to public groundwater contamination on healthcare utilization at the county level. Treated counties are those with unconventional drilling that occurred with in one, three, or five kilometers of a groundwater source location. Treatment timing aligns with the the year-quarter after the spud date for the first oil or natural gas well drilled on the shale formation near a groundwater source. The identifying assumption is that healthcare utilization in counties with residents who were not served contaminated drinking water, both on and outside of the shale formation, would have trended similar to counties with residents served contaminated drinking water had they not experienced drilling near their water source. Treatment is binary and an absorbing state. I calculate the ATT in county-year-quarter cohorts.

Table 3 shows that drilling within one kilometer of a water source leads to a 2.35 percentage point increase in county-level healthcare utilization relative to the comparison counties. This effect corresponds to a 13 percent increase over the pre-hydraulic fracturing mean, or approximately 4 additional inpatient claims per 1,000 residents in a treated county. Increasing the distance from one kilometer to three or five kilometers results in a sharp decrease in the effect size with small increases in healthcare utilization of around 4 to 5 percent over the pre-hydraulic fracturing mean, and these changes are not statistically different from zero. Consistent with the zip code level estimates shown in Table 2, exposure to water contamination is positively correlated with healthcare utilization for circulatory conditions, cancer, gastrointestinal conditions, urological conditions, conditions of the central nervous system, dermatological conditions, and blood diseases. Additionally, this analysis reveals a relationship between endocrine conditions and water contamination, with a 16 percent increase in endocrine conditions after drilling near a water source relative to the pre-hydraulic fracturing mean.

This analysis relies on a blunt instrument for treatment definition, and therefore I test the robustness of the results to higher exposure levels. I consider a subsample of counties with an affected population share larger than the median affected population share among all counties treated at the one kilometer level. This sample includes six counties with more than three percent of their total population exposed to contaminated drinking water. I also use the subsample of counties with more than 9 percent of their total population exposed to contaminated drinking water, which is the 75th percentile of exposure populations within one kilometer, for a total of three treated counties. These results are noisier, but show a significant increase in healthcare utilization, particularly for conditions of the central nervous system, gastrointestinal conditions, urological conditions, and blood diseases. However, the increase in endocrine conditions is not robust this analysis. I include the full table of results in Appendix Table A7.

If the results in this analysis were driven by exposure to hydraulic fracturing alone, I would not expect to find an effect at the one kilometer level when only looking at the shale counties alone. To see if the results are robust to excluding the border counties from the analysis, I repeat the estimation strategy using only counties that intersect with the Eagle Ford Shale formation. However, the effect is nearly identical at the one kilometer level when focusing on the Eagle Ford counties alone. Further, at the three and five kilometer

Table 3: The effects of public groundwater contamination

Health Condition	1km (1)	3km (2)	5km (3)
All Conditions	0.0235***	0.0073	0.0080
	(0.0081)	(0.0092)	(0.0097)
Circulatory	0.0064***	0.0029	0.0032
	(0.0024)	(0.0029)	(0.0030)
General Illness	0.0009**	0.0008	0.0006
	(0.0005)	(0.0006)	(0.0005)
Mental Health	0.0011	0.0006	0.0006
	(0.0007)	(0.0007)	(0.0008)
Endocrine	0.0056***	0.0013	0.0012
	(0.0020)	(0.0020)	(0.0020)
Cancer	0.0008*	-0.0005	-0.0003
	(0.0004)	(0.0009)	(0.0008)
Central Nervous	0.0019***	0.0008	0.0007
	(0.0004)	(0.0007)	(0.0008)
Respiratory	0.0020	0.0011	0.0011
	(0.0014)	(0.0011)	(0.0011)
Gastrointestinal	0.0026***	0.0007	0.0008
	(0.0008)	(0.0009)	(0.0009)
Urological	0.0017***	-0.0001	-0.0001
	(0.0004)	(0.0009)	(0.0009)
Pregnancy Complications	0.0002	0.0005*	0.0007**
	(0.0005)	(0.0003)	(0.0003)
Dermatological	0.0004**	0.0003	0.0004
	(0.0002)	(0.0003)	(0.0003)
Perinatal	0.0003	0.0003	0.0002
	(0.0003)	(0.0003)	(0.0002)
Blood Disease	0.0013***	0.0000	0.0001
	(0.0004)	(0.0006)	(0.0005)
Congenital Abnormalities	0.0001***	0.0002***	0.0001
	(0.0000)	(0.0001)	(0.0001)
Observations	1,960	1,960	1,960

Standard errors in parenthesis, clustered at the county-level.
*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.
This table shows the results of a staggered, two-way fixed effects estimation strategy where Eagle Ford counties with that oil or natural gas well drilling within 1km, 3km, and 5km of a public water source are the treatment counties, and all other counties in the sample are the control counties.

levels the increases in healthcare utilization are once again much smaller and not statistically different from zero. These results are included in Appendix Table A8.

Finally, I match the identification number associated with each water source to data on groundwater contamination from the Texas Commission Environmental Quality (2024) to understand the first-stage effect of hydraulic fracturing near a groundwater source location. I find evidence of increases in the concentration of water contaminants related to hydraulic fracturing after drilling within one kilometer of a water source location. However, the monitoring schedule for different drinking water contaminants varies considerably depending on the contaminant in question. Some contaminants are tested frequently, while others are tested annually or every three years. If groundwater contamination is due to surface spills during the pre-production period, the lack of testing precisely before and after the entry of hydraulic fracturing near a water source means that fracking-related groundwater contamination may be missed in this analysis. Therefore, I take these results as suggestive evidence of water contamination from drilling near public water sources, particularly in combination with prior research that finds that unconventional drilling within one kilometer of a water source increases the concentration of contaminants related to hydraulic fracturing (Hill and Ma 2022). I also include this analysis in Appendix Section 1.2.

5 Discussion and Conclusion

The past two decades saw dramatic increases in hydraulic fracturing and horizontal drilling technologies used in the production of oil and natural gas, leading to changes in environmental quality in areas with shale gas and oil. This paper estimates how hydraulic fracturing affects healthcare utilization on the Eagle Ford Shale formation in southern Texas.

I begin with a focus on general increases in healthcare utilization after the entry of hydraulic fracturing on the Eagle Ford Shale formation. I find that compared to counties outside of Eagle Ford, healthcare utilization for conditions that are plausibly related to hydraulic fracturing increase by 15 percent in zip codes on the Eagle Ford Shale formation relative to the pre-hydraulic fracturing mean. Although these estimates could be due to population-level compositional changes, I repeat the specification for subpopulations that are less likely to be impacted by compositional changes and estimate larger increases in utilization for Medicaid and Medicare recipients. Finally, I find a causal relationship between increases in healthcare utilization and drilling near a public water source, which is not likely explained by compositional changes or air pollution alone, suggesting that groundwater contamination may serve as a primary pathway for pollution exposure.

At the zip code level, which offers a more precise estimate of the effects of hydraulic fracturing due to more a accurate representation of both timing and proximity to fracking, I find positive and significant increases in utilization for circulatory conditions, cancer, neurological conditions, gastrointestinal conditions, urological conditions, dermatological conditions, and blood diseases. These conditions align with conditions highlighted in public health and medical research that identifies chemicals used in hydraulic fracturing fluid and their potential effect on human health through groundwater and drinking water exposure (Colborn et al. 2011; Epstein 2017; Shonkoff, Hays, and Finkel 2014; Shrestha et al. 2017; Webb et al. 2018; Wollin et al. 2020). Although this study represents a longer exposure time than other papers, Eagle Ford residents were only exposed to hydraulic fracturing for a maximum of just over seven years in my sample. Additionally, for some of these conditions, the effects were larger among older adults and low-income individuals. Therefore, the increase in utilization appeared quite quickly and particularly affected vulnerable populations. This result may indicate worse outcomes over longer periods of time.

I also provide evidence that groundwater pollution acts as one exposure mechanism to hydraulic fracturing pollution. I find that counties that experienced hydraulic fracturing within one kilometer of a public groundwater source experienced a 13 percent increase in healthcare utilization over the sample mean for conditions related to unconventional drilling compared to counties that did not experience hydraulic fracturing within one kilometer of a public groundwater source. The effects are smaller and not statistically significant for counties with hydraulic fracturing within a three or five kilometer radius. This aligns with research on groundwater contamination, which suggests that contamination occurs at the groundwater intake point as a result of surface spills, structural impairment, or site discharge (Shanafield, Cook, and Simmons 2019; Rozell and Reaven 2012; Shonkoff, Hays, and Finkel 2014). Additional research at the household level, particularly for households that rely on private groundwater sources, would help to pinpoint the magnitude of the effects. Policymakers should consider additional protections for individuals living near shale development, such as expanding drinking water tests to include more hydraulic fracturing chemicals, offering free or low cost water tests for individuals on private wells, and restricting drilling locations to less populated areas, far away from drinking water sources.

Researchers should continue to monitor health outcomes in areas near shale development and work to diagnose the exact chemicals driving the negative health outcomes to more efficiently support affected families. In both the primary analysis and robustness checks on the relationship between hydraulic fracturing and water contamination, I identify a particular set of conditions among adults that appear to respond to exposure to contaminated water: conditions of the central nervous system, gastrointestinal conditions, urological conditions, and blood diseases. Public health researchers should consider focusing their attention on identifying hydraulic fracturing chemicals that could be causing these results to inform public policy on which chemicals to regulate in hydraulic fracturing fluid.

Although this paper finds a relationship between hydraulic fracturing and healthcare utilization, there are several limitations. First, I am unable to estimate actual exposure to hydraulic fracturing. Instead, I use methods to estimate the potential for exposure. However, these strategies do not account for potential spillover effects resulting from hydraulic fracturing near the borders of counties or zip codes, nor does it take mobility between geographic areas into consideration, both of which may change the amount of exposure at the individual-level. I also only directly test water pollution as a mechanism for pollution exposure, however air and noise pollution also likely impact health as well. Research identifying the causal impact of these alternative pathways on adult health conditions would provide a more complete picture of the effects of hydraulic fracturing on human health.

Additionally, for econometric convenience, this paper makes the assumption that certain healthcare conditions are likely related to hydraulic fracturing; in reality, healthcare is nuanced, and understanding the interaction between the chemicals used in hydraulic fracturing and human health is still developing. I use current research by public health and medical researchers to categorize the conditions, but future work on understanding how hydraulic fracturing effects health would allow for more precision.

This paper points to widespread health effects of hydraulic fracturing across the lifespan and links the effects to groundwater contamination. These increases in healthcare utilization imply negative externalities from hydraulic fracturing borne by those who live on shale formations. Although these new drilling technologies offer substantial consumer benefits, as shown by Hausman and Kellogg (2015), these externalities should be considered in an overall welfare analysis. Future work should continue to track health outcomes for individuals living on and near shale formations to better understand the long term effects of hydraulic fracturing.

6 References

References

Alawattegama, Shyama K, Tetiana Kondratyuk, Renee Krynock, Matthew Bricker, Jennifer K Rutter, Daniel J Bain, and John F Stolz. 2015. "Well water contamination in a rural community in southwestern Pennsylvania near unconventional shale gas extraction." *Journal of Environmental Science and Health, Part A* 50 (5): 516–528.

Balthrop, Andrew T, and Zackary Hawley. 2017. ""I can hear my neighbors' fracking"': The effect of natural gas production on housing values in Tarrant County, TX." *Energy Economics* 61:351–362.

Black, Katie Jo, Andrew J Boslett, Elaine L Hill, Lala Ma, and Shawn J McCoy. 2021. "Economic, environmental, and health impacts of the fracking boom." *Annual Review of Resource Economics* 13 (1): 311–334.

Bonetti, Pietro, Christian Leuz, and Giovanna Michelon. 2021. "Large-sample evidence on the impact of unconventional oil and gas development on surface waters." *Science* 373 (6557): 896–902.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics* 225 (2): 200–230.

Clay, Karen, Akshaya Jha, Nicholas Muller, and Randall Walsh. 2017. *The external costs of transporting petroleum products by pipelines and rail: evidence from shipments of crude oil from North Dakota.* Technical report. National Bureau of Economic Research.

Colborn, Theo, Carol Kwiatkowski, Kim Schultz, and Mary Bachran. 2011. "Natural gas operations from a public health perspective." *Human and Ecological Risk Assessment: An International Journal* 17 (5): 1039–1056.

Currie, Janet, Michael Greenstone, and Katherine Meckel. 2017. "Hydraulic fracturing and infant health: New evidence from Pennsylvania." *Science Advances* 3 (12): e1603021.

Denham, Alina, Mary Willis, Alexis Zavez, and Elaine Hill. 2019. "Unconventional natural gas development and hospitalizations: evidence from Pennsylvania, United States, 2003–2014." *Public health* 168:17–25.

DiGiulio, Dominic C, Richard T Wilkin, Carlyle Miller, and Gregory Oberley. 2011. *Investigation of ground water contamination near Pavillion, Wyoming.* Office of Research / Development, National Risk Management Research Laboratory.

Environmental Quality, Texas Commission on. 2024. *Drinking Water Watch.* Accessed: 2024-11-11. https://dww2.tceq.texas.gov/DWW/.

Epstein, Anne C. 2017. "The human health implications of oil and natural gas development." In *Advances in chemical pollution, environmental management and protection,* 1:113–145. Elsevier.

Finkel, Madelon L, and Adam Law. 2011. "The rush to drill for natural gas: a public health cautionary tale." *American Journal of Public Health* 101 (5): 784–785.

Fontenot, Brian E, Laura R Hunt, Zacariah L Hildenbrand, Doug D Carlton Jr, Hyppolite Oka, Jayme L Walton, Dan Hopkins, Alexandra Osorio, Bryan Bjorndal, Qinhong H Hu, et al. 2013. "An evaluation of water quality in private drinking water wells near natural gas extraction sites in the Barnett Shale Formation." *Environmental science & technology* 47 (17): 10032–10040.

Garshick, Eric, Francine Laden, Jaime E Hart, Bernard Rosner, Mary E Davis, Ellen A Eisen, and Thomas J Smith. 2008. "Lung cancer and vehicle exhaust in trucking industry workers." *Environmental Health Perspectives* 116 (10): 1327–1332.

Gilmer, Robert W., Raúl Hernandez, and Keith R. Phillips. 2012. "Oil Boom in Eagle Ford Shale Brings New Wealth to South Texas." *Southwest Economy.*

Gilmer, Robert W., and Emily Kerr. 2010. "Natural Gas from Shale: Texas Revolution Goes Global." *Southwest Economy.*

Gross, Sherilyn A, Heather J Avens, Amber M Banducci, Jennifer Sahmel, Julie M Panko, and Brooke E Tvermoes. 2013. "Analysis of BTEX groundwater concentrations from surface spills associated with hydraulic fracturing operations." *Journal of the Air & Waste Management Association* 63 (4): 424–432.

Hausman, Catherine, and Ryan Kellogg. 2015. *Welfare and distributional implications of shale gas.* Technical report. National Bureau of Economic Research.

Hildenbrand, Zacariah L, Doug D Carlton Jr, Brian E Fontenot, Jesse M Meik, Jayme L Walton, Josh T Taylor, Jonathan B Thacker, Stephanie Korlie, C Phillip Shelor, Drew Henderson, et al. 2015. "A comprehensive analysis of groundwater quality in the Barnett Shale region." *Environmental science & technology* 49 (13): 8254–8262.

Hildenbrand, Zacariah L, Phillip M Mach, Ethan M McBride, M Navid Dorreyatim, Josh T Taylor, Doug D Carlton Jr, Jesse M Meik, Brian E Fontenot, Kenneth C Wright, Kevin A Schug, et al. 2016. "Point source attribution of ambient contamination events near unconventional oil and gas development." *Science of the Total Environment* 573:382–388.

Hill, Elaine L. 2018. "Shale gas development and infant health: evidence from Pennsylvania." *Journal of Health Economics* 61:134–150.

———. 2024. "The Impact of Oil and Gas Extraction on Infant Health." *American Journal of Health Economics* 10 (1): 68–96.

Hill, Elaine L, and Lala Ma. 2022. "Drinking water, fracking, and infant health." *Journal of Health Economics* 82:102595.

Jackson, Robert B, Avner Vengosh, Thomas H Darrah, Nathaniel R Warner, Adrian Down, Robert J Poreda, Stephen G Osborn, Kaiguang Zhao, and Jonathan D Karr. 2013. "Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction." *Proceedings of the National Academy of Sciences* 110 (28): 11250–11255.

Kearney, Melissa S, and Riley Wilson. 2018. "Male earnings, marriageable men, and nonmarital fertility: Evidence from the fracking boom." *Review of Economics and Statistics* 100 (4): 678–690.

Kondash, Andrew J, Elizabeth Albright, and Avner Vengosh. 2017. "Quantity of flowback and produced waters from unconventional oil and gas exploration." *Science of the Total Environment* 574:314–321.

Litovitz, Aviva, Aimee Curtright, Shmuel Abramzon, Nicholas Burger, and Constantine Samaras. 2013. "Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania." *Environmental Research Letters* 8 (1): 014017.

McKenzie, Lisa M, William B Allshouse, Tim E Byers, Edward J Bedrick, Berrin Serdar, and John L Adgate. 2017. "Childhood hematologic cancer and residential proximity to oil and gas development." *PloS one* 12 (2): e0170423.

McKenzie, Lisa M, Ruixin Guo, Roxana Z Witter, David A Savitz, Lee S Newman, and John L Adgate. 2014. "Birth outcomes and maternal residential proximity to natural gas development in rural Colorado." *Environmental health perspectives* 122 (4): 412–417.

McKenzie, Lisa M, Roxana Z Witter, Lee S Newman, and John L Adgate. 2012. "Human health risk assessment of air emissions from development of unconventional natural gas resources." *Science of the Total Environment* 424:79–87.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The housing market impacts of shale gas development." *American Economic Review* 105 (12): 3633–3659.

Osborn, Stephen G, Avner Vengosh, Nathaniel R Warner, and Robert B Jackson. 2011. "Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing." *Proceedings of the National Academy of Sciences* 108 (20): 8172–8176.

Peng, Lizhong, Chad Meyerhoefer, and Shin-Yi Chou. 2018. "The health implications of unconventional natural gas development in Pennsylvania." *Health economics* 27 (6): 956–983.

Persico, Claudia, and Dave E Marcotte. 2022. *Air quality and suicide.* Technical report. National Bureau of Economic Research.

Pétron, Gabrielle, Gregory Frost, Benjamin R Miller, Adam I Hirsch, Stephen A Montzka, Anna Karion, Michael Trainer, Colm Sweeney, Arlyn E Andrews, Lloyd Miller, et al. 2012. "Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study." *Journal of Geophysical Research: Atmospheres* 117 (D4).

Roy, Anirban A, Peter J Adams, and Allen L Robinson. 2014. "Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas." *Journal of the Air & Waste Management Association* 64 (1): 19–37.

Rozell, Daniel J, and Sheldon J Reaven. 2012. "Water pollution risk associated with natural gas extraction from the Marcellus Shale." *Risk Analysis: An International Journal* 32 (8): 1382–1393.

Shanafield, Margaret, Peter G Cook, and Craig T Simmons. 2019. "Towards quantifying the likelihood of water resource impacts from unconventional gas development." *Groundwater* 57 (4): 547–561.

Shonkoff, Seth BC, Jake Hays, and Madelon L Finkel. 2014. "Environmental public health dimensions of shale and tight gas development." *Environmental Health Perspectives* 122 (8): 787–795.

Shrestha, Namita, Govinda Chilkoor, Joseph Wilder, Venkataramana Gadhamshetty, and James J Stone. 2017. "Potential water resource impacts of hydraulic fracturing from unconventional oil production in the Bakken shale." *Water Research* 108:1–24.

Speight, James G. 2019. *Shale oil and gas production processes.* Gulf Professional Publishing.

Texas Health Care Information Collective. 2007–15. "Inpatient Public Use Data Files." https://www.dshs.texas.gov/THCIC/Hospitals/Download.shtm.

Texas Railroad Commission. 2007–15. "Well Distribution by County Well Counts." https://www.rrc.texas.gov/oil-and-gas/research-and-statistics/well-information/well-distribution-by-county/.

Texas Water Development Board. 2023. "Water Use Survey Data."

U.S. Census Bureau. 2024. *Decennial Census (2020, 2010, 2000).* Accessed: 2024-09-29. https://www.census.gov/data/developers/data-sets/decennial-census.html.

U.S. EPA. 2016. *Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report).* Technical report EPA/600/R-16/236F. Washington, DC: U.S. Environmental Protection Agency.

Webb, Ellen, Julie Moon, Larysa Dyrszka, Brian Rodriguez, Caroline Cox, Heather Patisaul, Sheila Bushkin, and Eric London. 2018. "Neurodevelopmental and neurological effects of chemicals associated with unconventional oil and natural gas operations and their potential effects on infants and children." *Reviews on environmental health* 33 (1): 3–29.

Willis, Mary D, Todd A Jusko, Jill S Halterman, and Elaine L Hill. 2018. "Unconventional natural gas development and pediatric asthma hospitalizations in Pennsylvania." *Environmental Research* 166:402–408.

Wilson, Riley. 2022. "Moving to Economic Opportunity The Migration Response to the Fracking Boom." *Journal of Human Resources* 57 (3): 918–955.

Wollin, Klaus-Michael, Georg Damm, Heidi Foth, Alexius Freyberger, Thomas Gebel, Aswin Mangerich, Ursula Gundert-Remy, Falko Partosch, Claudia Röhl, Thomas Schupp, et al. 2020. "Critical evaluation of human health risks due to hydraulic fracturing in natural gas and petroleum production." *Archives of Toxicology* 94:967–1016.

Xu, Ziyue. 2023. "Fracking Boom and Respiratory Health: Evidence from Texas." Job Market Paper, Department of Economics, University of Texas at Austin, August. https://minio.la.utexas.edu/colaweb-prod/event_images/5/52088/ziyue_xu_jmp_957a3e24-10cc-4363-99a4-5fa09eb8e5c4.pdf.

Appendix

1.1 Data Description

This paper combines several sources of data, and in this section I provide links to publicly available data sources and describe how I merge the sources to create the dataset used in the analysis.

1.1.1 HEALTH OUTCOME DATA

Publicly available, inpatient healthcare discharge data from the Texas Health Care Information Collective (2007–15) provides quarterly details about inpatient claims, along with the patient county and zip code of residence, which I use create both the county- and zip-code-level panels. Additionally, the data includes the health insurer, and patient gender, which I use for the heterogeneity analysis. I use data between 2007 and the third quarter of 2015. The data is available for download on the THCIC webpage.

I classify the healthcare conditions using the HCUP Multi-Level CCS software, which takes the diagnosis and procedural codes and categorizes them in to clinically-meaningful categories. The table below breaks down the CCS codes that I include for each condition.

Table A1: HCUP CCS Codes

Condition Name	L1 CSS Codes	L2 CSS Codes
Cancer	2	All codes
Endocrine	3	All codes
Blood Disease	4	All codes
Mental Illness	5	All codes
Central Nervous	6	All codes
Circulatory	7	All codes
Respiratory	8	All codes
Gastrointestinal	9	All codes
Urinary	10	All codes
Pregnancy Complications	11	11.2, 11.3, 11.4, 11.5, 11.6
Dermatological	12	All codes
Congenital	14	All codes
Perinatal	15	15.2, 15.3, 15.4, 15.5, 15.6, 15.7
General Illness	17	17.1

1.1.2 OIL AND NATURAL GAS PERMIT DATA

To identify the location of oil and gas wells, I use public data from the RRC Drilling Permit Master File and the RRC GIS data files. I exclude permits that are missing location or spud date data, and I keep all permits with a spud date between 2007 and 2015. I also only use oil, gas, and brine mining sites, as well as injection sites.

Using the GIS file and ArcGIS, I match well locations to the the county and zip code where each well was drilled using county and ZCTA boundary shapefiles from the U.S. Census Bureau using the Spatial Join tool. I identify which wells were drilled within two miles of the boundary of the Eagle Ford Shale formation using shapefiles from the U.S. Energy Information Administration, again using a Spatial Join. Figure 2 shows the counties included in the analysis and the boundaries of the Eagle Ford Shale formation, and Figure 4a shows all wells drilled in both the Eagle Ford and non-Eagle Ford counties in the time period, plotted in orange.

Finally, I also use the GIS file and the ArcGIS Near Table tool to find oil and natural gas wells drilled within one, three, and five kilometers of a public water source. I then use the spud dates from the Permit Master File to determine the treatment timing. Figure 4a shows the water sources with oil and gas wells drilled nearby in gray, and Figure 4b shows an example of how I identify oil and gas wells drilled near water sources.

1.1.3 WATER SOURCE DATA

Groundwater data, including the latitude and longitude of the water source, volume of surface and groundwater used, and number of individuals served was privately requested and complied for this project by the Texas Water Development Board. As described above, I use the latitude and longitude of the PWS location from these data to match nearby oil and natural gas wells using the Near Table tool in ArcGIS.

1.1.4 ADDITIONAL DATA

I also use data from the U.S. Census Bureau Population Totals and the Decennial Census totals (U.S. Census Bureau 2024) for the total population of the counties and ZCTA's used in the analysis, as well as descriptive statitistics comparing demographics of the counties.

1.2 Groundwater Contaminant Analysis

Using data purchased from the Texas Commission on Environmental Quality's Texas Drinking Water Watch, I estimate the first-stage effect of hydraulic fracturing near a groundwater source location on groundwater contamination. These data include the date and PWS where the sample was collected, the sample point (distribution or entry point), the analysis method, contaminant names, and the test results. This allows me to analyze the change in contaminant concentrations before and after the entry of hydraulic fracturing on the Eagle Ford Shale formation, even at concentrations below the maximum threshold set by the Safe Drinking Water Act.

An accurate measure of the relationship between hydraulic fracturing near a public water source and drinking water quality would require testing the source water at the affected and non-affected sites regularly for the exact contaminants used in the hydraulic fracturing process, including testing before and after hydraulic fracturing. However, the monitoring schedule for different drinking water contaminants varies considerably depending on the contaminant in question. The variation in testing frequency makes it challenging to precisely identify changes in contaminant concentrations around the hydraulic fracturing entry date. Therefore, I estimate the average change in concentrations of hydraulic fracturing contaminants for PWS's with unconventional drilling within one kilometer compared to PWS's without unconventional drilling nearby. I use the following, two-way fixed effects specification:

$$Y_{pt} = \beta_0 + \beta_1[Treat]_{pt} + \eta_p + \gamma_j\tau_t + \epsilon_{pjt} \tag{3}$$

Where Y_{pt} is the test result for each contaminant at PWS p measured at time t. $Treat_{pt}$ is an indicator equal to one for all dates after the first spud date of an oil or gas well drilled within one kilometer of PWS p on the Eagle Ford Shale formation. The underlying assumption is that exposure to and detection of contamination occurs in the periods after the spud date. I leave out water sources with oil or natural gas drilling within three or five kilometers since these also may experience an increase in hydraulic fracturing chemicals during this time period, which would negatively bias the results. To account for seasonality that may impact water contamination, I include year-quarter fixed effects, τ_t. I also include a fixed effect for each PWS, η_p. Standard errors are clustered at the PWS level. β_1 estimates the average change in a contaminant level within PWS-contaminant groups, measured in micrograms per liter ($\mu g/L$), after a unconventional oil or gas well is drilled within one kilometer of a PWS.

Panel A of Table A2 shows the change in the level of total dissolved solids (TDS) and the level of hydraulic fracturing related contaminants. The results show small increases in TDS for the water sources near hydraulic fracturing sites when compared to the control water sources, however these estimates are not statistically different from zero. Additionally, is a small increase in the concentration hydraulic fracturing contaminants in the treated water sources when compared to the control sources. The results consider the full set of contaminants possibly used in hydraulic fracturing, however whether each of these contaminants are actually used is unknown. I test each contaminant related to hydraulic fracturing either through fracturing fluid or produced water and include the results in Table A3. I find a statistically significant increase in barium, which is used in fracking fluid and is found in produced water. In addition, I find increases in bromoform, radium-226, calcium, and potassium, which can be found in produced water. The overall levels of iron, which could be used in both fracturing fluid and is found in produced water, decreased during this time period.

The magnitude of these effects is large, however many contaminants are not detectable below a certain threshold and are recorded as zero in the data, biasing the effect size away from zero. In Panel B of Table A2, I show the results of the same specification, but take the log of the test results measured in $\mu g/L$. This gives the percent change in the contaminant concentration conditional on the contaminant exceeding detectable levels. The results for total dissolved solids and the full set of hydraulic fracturing chemicals are not statistically different than zero, but the individual specifications reveal statistically significant increases in xylenes, selenium, and zinc, in addition to increases in barium, calcium and potassium. This specification confirms a decrease in iron and also shows a statistically significant decrease in sodium.

Table A2: The effects of hydraulic fracturing near a public groundwater source on water contamination

Panel A

	Concentration (1)	Observations (2)	Mean (3)
Total Dissolved Solids	4.27	1,478	464.9
	(17.5)		(375.7)
All Fracking Chemicals	238.9	188,353	3,613.3
	(506.5)		(33,114.6)

Panel B

	Log(Concentration) (1)	Observations (2)	Mean
Log(Total Dissolved Solids)	-0.009	1,478	
	(0.030)		
Log(All Fracking Chemicals)	0.071	85,666	
	(0.112)		

Standard errors in parenthesis, clustered at the PWS level.
*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.

Notes: This table shows the change in concentration of drinking water contaminants after oil and gas drilling occurred within 1km of a PWS location, relative to PWS locations without drilling within 5km both on and off the Eagle Ford Shale formation. Panel A shows the level-change in contaminants and Panel B shows the percent change in contaminants that exceed threshold for detection.

Table A3: Groundwater Contaminant: Individual contaminants

Chemical	Log(Concentration) (1)	Observations (2)	Raw Concentration (3)	Observations (4)	Mean (5)	Fracking Fluid (6)	Produced Water (7)
Lead (1030)	-0.122	5,825	0.475	9,712	1.647	X	X
	(0.146)		(0.449)		(7.619)		
Bromodichloromethane (2943)	0.0413	5,316	0.441	7,135	8.142		X
(0.124)		(1.173)		(9.016)			
Copper (1022)	-0.0231	8,734	-13.73	9,608	127.096	X	X
	(0.237)		(24.480)		(233.954)		
Xylenes, total (2955)	0.799**	124	0.266	1,496	0.17	X	X
	(0.309)		(0.210)		(0.938)		
Nitrate (1040)	-0.033	3,283	46.06	4,278	389.085		X
	(0.088)		(28.960)		(891.398)		
Arsenic (1005)	-0.272	631	-2.045	1,296	10.611	X	X
	(0.237)		(3.353)		(24.725)		
Chloroform (2941)	0.21	4,767	1.168	7,135	7.27		X
	(0.166)		(1.215)		(16.074)		
Bromoform (2942)	0.0656	5107	3.793***	7,135	6.597		X
	(0.087)		(1.460)		(10.952)		
Fluoride (1025)	0.0493	1,435	44.73	1,530	497.336		X
	(0.092)		(38.890)		(577.606)		
Barium (1010)	0.311***	1,005	30.04*	1,006	201.703	X	X
.	(0.101)		(17.510)		(377.455)		
Nickel (1036)	0.0464	498	0.24	960	1.047		X
	(0.268)		(0.250)		(1.593)		
Selenium (1045)	0.516**	265	1.652	960	1.268		X
	(0.244)		(1.921)		(3.271)		

Standard errors in parenthesis, clustered at the county-level.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Notes: This table shows the β_1 coefficients from Equation (3), and estimates the change in contaminant levels for individual contaminants related to hydraulic fracturing. The list of contaminants is generated from Hill and Ma (2022) based on the contaminants measured regularly in the Environmental Quality (2024) data. Columns (1) and (3) give the change in the log of the concentration for contaminants above the detection threshold and the raw change in concentration, respectively. Column (5) gives the average concentration for each chemical. Columns (6) and (7) identifies which contaminants are used in fracking fluid and which come from produced water, based on Hill and Ma (2022).

Chemical	Log(Concentration)	Observations	Raw Concentration	Observations	Mean	Fracking Fluid	Produced Water
Radium-228 (4030)	-0.164	305	0.416	1,061	0.659		X
	(0.128)		(0.326)		(1.283)		
Radium-226 (4020)	0.509	401	3.281*	441	4.892		X
	(0.318)		(1.868)		(5.484)		
Manganese (1032)	-0.397	865	1.171	998	17.479		X
	(0.348)		(9.244)		(35.702)		
Iron (1028)	-0.663***	733	-172.3***	1,018	160.77	X	X
	(0.243)		(61.950)		(346.042)		
Sodium (1052)	-0.211**	1023	-14,936	1,023	105,132.70		X
	(0.096)		(13,199)		(99,450.21)		
Calcium (1016)	0.225*	941	8,857*	959	37,775.20		X
	(0.125)		(5,029)		(31,031.99)		
Chloride (1017)	0.0159	1478	6,283	1,479	71581.87	X	X
	(0.061)		(8,230)		(84,448.74)		
Sulfate (1055)	-0.0196	1373	4,634	1,479	44,778.75	X	X
	(0.090)		(6,162)		(64,162.48)		
Aluminium (1002)	0.0105	266	11.82	960	24.479	X	X
	(0.215)		(21.29)		(298.740)		
Zinc (1095)	0.992***	768	4.254	960	56.134	X	X
	(0.285)		(25.600)		(256.283)		
Magnesium (1031)	-0.0465	812	401.3	959	6,725.82		X
	(0.139)		(1,094)		(6,885.575)		
Potassium (1042)	0.344***	252	1,589***	258	4,836.30		X
	(0.074)		(483.4)		(3,710.869)		

Standard errors in parenthesis, clustered at the county-level.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Notes: This table shows the β_1 coefficients from Equation (3), and estimates the change in contaminant levels for individual contaminants related to hydraulic fracturing. The list of contaminants is generated from Hill and Ma (2022) based on the contaminants measured regularly in the Environmental Quality (2024) data. Columns (1) and (3) give the change in the log of the concentration for contaminants above the detection threshold and the raw change in concentration, respectively. Column (5) gives the average concentration for each chemical. Columns (6) and (7) identifies which contaminants are used in fracking fluid and which come from produced water, based on Hill and Ma (2022).

1.3 Additional Tables and Figures

Table A4: Change in healthcare utilization: Medicare recipients

	County		Zip Code	
Health Condition	β (1)	Sample Mean (2)	β (3)	Sample Mean (4)
All Conditions	0.0054	0.103	(0.0180***	0.178
	(0.007)		(0.0064)	
Circulatory	0.0020	0.034	0.0053***	0.050
	(0.002)		(0.0021)	
General Illness	0.0003	0.003	0.0004	0.006
	(0.000)		(0.0003)	
Mental Health	0.0001	0.005	0.0009**	0.015
	(0.001)		(0.0005)	
Endocrine	0.0012	0.022	0.0028*	0.037
	(0.002)		(0.0015)	
Cancer	0.0003	0.005	0.0008	0.008
	(0.000)		(0.0005)	
Central Nervous	0.0003	0.007	0.0010*	0.013
	(0.001)		(0.0005)	
Respiratory	0.0003	0.011	0.0013*	0.012
	(0.001)		(0.0008)	
Gastrointestinal	0.0008	0.007	0.0016***	0.012
	(0.001)		(0.0005)	
Urological	0.0002	0.008	0.0030**	0.013
	(0.001)		(0.0013)	
Dermatological	0.0002	0.002	0.0003	0.003
	(0.000)		(0.0002)	
Blood Disease	0.0003	0.005	0.0012***	0.008
	(0.000)		(0.0003)	
Observations	1,960	437	13,720	1,828

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Notes: Table A4 focuses on the change in healthcare utilization for the subsample of claims submitted by Medicare recipients. Column (1) reports the county-level β_1 coefficients from Equation 2. Column (3) reports the analogous estimates from the staggered difference-in-differences specification at the zip-code-level. Columns (2) and (4) give the pre-fracking entry mean utilization rates for Medicare recipients at the county- and zip-code levels respectively. Because most Medicare recipients are over the age of 65, I exclude pregnancy-related outcomes.

Table A5: Change in healthcare utilization: Medicaid recipients

	County		Zip Code	
Health Condition	β (1)	Sample Mean (2)	β (3)	Sample Mean (4)
All Conditions	0.0052	0.053	0.0102	0.0419
	(0.0038)		(0.0063)	
Circulatory	0.0006	0.013	0.0025*	0.0101
	(0.0009)		(0.0014)	
General Illness	0.0001	0.002	0.0006	0.0011
	(0.0002)		(0.0005)	
Mental Health	0.0008	0.007	0.0005	0.0061
	(0.0005)		(0.0005)	
Endocrine	0.0010	0.010	0.0013	0.0070
	(0.0008)		(0.0011)	
Cancer	0.0003	0.002	0.0003	0.0014
	(0.0002)		(0.0002)	
Central Nervous	0.0006	0.004	0.0008*	0.0030
	(0.0003)		(0.0004)	
Respiratory	0.0005	0.005	0.0007	0.0041
	(0.0005)		(0.0006)	
Gastrointestinal	0.0003	0.004	0.0012	0.0026
	(0.0003)		(0.0007)	
Urological	0.0004	0.004	0.0025	0.0031
	(0.0003)		(0.0018)	
Pregnancy Complications	0.0004	0.003	-0.0001	0.0024
	(0.0004)		(0.0003)	
Dermatological	0.0002*	0.001	0.0003**	0.0005
	(0.0001)		(0.0002)	
Perinatal	0.0002	0.003	-0.0001	0.0019
	(0.0002)		(0.0003)	
Blood Disease	0.0004*	0.003	0.0005*	0.0021
	(0.0002)		(0.0003)	
Congenital Abnormalities	0.0001	0.0004	-0.0001	0.0003
	(0.0000)		(0.0001)	
Observations	1,960	228	13,720	167

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Notes Table A5 reports focuses on the subsample of claims submitted by Medicaid recipients and reports the county-level β_1 coefficients from Equation 2 in Column (1). Column (3) reports the analogous estimates from the staggered difference-in-differences specification at the zip-code-level. Columns (2) and (4) give the pre-fracking entry mean utilization rates for Medicaid recipients at the county- and zip-code levels respectively. This reveals a strikingly large increase in urological conditions, which is likely related to kidney disease, as patients with kidney disease are eligible for Medicaid coverage.

Table A6: Change in healthcare utilization: Women

	County		Zip Code	
Health Condition	β (1)	Sample Mean (2)	β (3)	Sample Mean (4)
All Conditions	0.0061	0.09	0.0058	0.0820
	(0.0061)		(0.0062)	
Circulatory	0.0011	0.024	0.0003	0.0210
	(0.0018)		(0.0020)	
General Illness	0.0004	0.003	0.0003	0.0030
	(0.0004)		(0.0005)	
Mental Health	0.0007	0.007	0.0006	0.0070
	(0.0005)		(0.0006)	
Endocrine	0.0011	0.019	0.0001	0.0160
	(0.0014)		(0.0018)	
Cancer	0.0003	0.004	0.0007***	0.0040
	(0.0002)		(0.0003)	
Central Nervous	0.0005	0.006	0.0008*	0.0060
	(0.0004)		(0.0005)	
Respiratory	0.0003	0.009	-0.0003	0.0080
	(0.0007)		(0.0009)	
Gastrointestinal	0.0007	0.006	0.0014**	0.0060
	(0.0005)		(0.0006)	
Urological	0.0003	0.007	0.0015**	0.0060
	(0.0004)		(0.0006)	
Pregnancy Complications	0.0008	0.004	0.0002	0.0040
	(0.0004)		(0.0005)	
Dermatological	0.0001	0.001	0.0005**	0.0010
	(0.0001)		(0.0002)	
Perinatal	0.0002	0.001	0.0004	0.0010
	(0.0001)		(0.0004)	
Blood Disease	0.0005	0.004	0.0004	0.0040
	(0.0003)		(0.0003)	
Congenital Abnormalities	0.0001	0.0004	0.000	0.0003
	(0.0000)		(0.0001)	
Observations	1,960	228	13,720	167

*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.

Notes: Table A6 reports focuses on the subsample of claims submitted by female claimants and reports the county-level β_1 coefficients from Equation 2 in Column (1). Column (3) reports the analogous estimates from the staggered difference-in-differences specification at the zip-code-level. Columns (2) and (4) give the pre-fracking entry mean utilization rates for females at the county- and zip-code levels respectively. Note that because some observations are missing patient sex, the pregnancy related estimates do not match the estimates in Table 2 exactly.

Table A7: One kilometer groundwater contamination: Exposure population greater than 3 and 9 percent

Health Condition	Treated Pop >3%	Treated Pop >9%
All Conditions	0.0176*	0.0134
	(0.0105)	(0.0111)
Circulatory	0.0036	0.0030
	(0.0028)	(0.0031)
General Illness	0.0005	0.0007
	(0.0006)	(0.0004)
Mental Health	0.0010	0.0000
	(0.0009)	(0.0006)
Endocrine	0.0040	0.0016
	(0.0027)	(0.0027)
Cancer	0.0007*	0.0007
	(0.0003)	(0.0005)
Central Nervous	0.0017***	0.0016***
	(0.0005)	(0.0000)
Respiratory	0.0004	0.0006
	(0.0018)	(0.0024)
Gastrointestinal	0.0026***	0.0026***
	(0.0010)	(0.0011)
Urological	0.0015***	0.0016***
	(0.0005)	(0.0004)
Pregnancy Complications	0.0010***	0.0009**
	(0.0003)	(0.0004)
Dermatological	0.0004	0.0002
	(0.0003)	(0.0003)
Perinatal	0.0007***	0.0004
	(0.0003)	(0.0003)
Blood Disease	0.0011**	0.0008**
	(0.0005)	(0.0004)
Congenital Abnormalities	0.0002***	0.0001**
	(0.0001)	(0.0000)
Observations	1,820	1,680

Standard errors in parenthesis, clustered at the county-level.
*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.

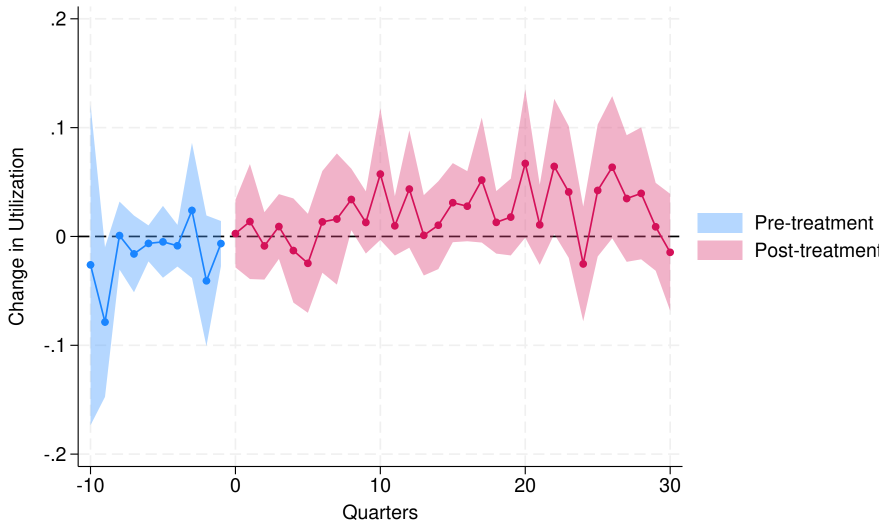
Notes: Table A7 shows the results of a staggered differences-in-differences estimation that focuses on the subset of counties with drilling within one kilometer of a water source that served more than 3 percent and more than 9 percent of the population. These results can be compared to Column (1) of Table 3.

Table A8: The effects of public groundwater contamination: Eagle Ford Counties only

Health Condition	1km (1)	3km (2)	5km (3)
All Conditions	0.0267**	0.0137	0.008
	(0.0123)	(0.0125)	(0.0010)
Circulatory	0.0079**	0.0071**	0.0032
	(0.0037)	(0.0030)	(0.0030)
General Illness	0.0007	0.0011	0.0006
	(0.0005)	(0.0007)	(0.0005)
Mental Health	0.0009	0.0005	0.0006
	(0.0012)	(0.0013)	(0.0008)
Endocrine	0.0063***	0.0017	0.0012
	(0.0028)	(0.0039)	(0.0020)
Cancer	0.0012	-0.0002	-0.0003
	(0.00090	(0.0009)	(0.0008)
Central Nervous	0.0021***	0.0005	0.0007
	(0.0008)	(0.0015)	(0.0008)
Respiratory	0.0027	0.0031**	0.0011
	(0.0019)	(0.0012)	(0.0011)
Gastrointestinal	0.0026**	0.0004	0.0008
	(0.0010)	(0.0004)	(0.0009)
Urological	0.0020**	0.0006	-0.0001
	(0.0008)	(0.0013)	(0.0008)
Pregnancy Complications	-0.0003	0.0004	0.0007**
	(0.0006)	(0.0006)	(0.0003)
Dermatological	0.0003	0.0003	0.0004
	(0.0003)	(0.0003)	(0.0003)
Perinatal	0.0001	0.0001	0.0002
	(0.0003)	(0.0004)	(0.0002)
Blood Disease	0.0014*	-0.0002	0.0001
	(0.0008)	(0.0007)	(0.0005)
Congenital Abnormalities	0.0000	0.0001	0.0001
.	(0.0001)	(0.0001)	(0.0001)
Observations	665	665	665

Standard errors in parenthesis, clustered at the county-level.
*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.
Notes: Table A8 is identical to Table 3, but only uses counties that intersect with the shale formation in both the treatment and control groups.

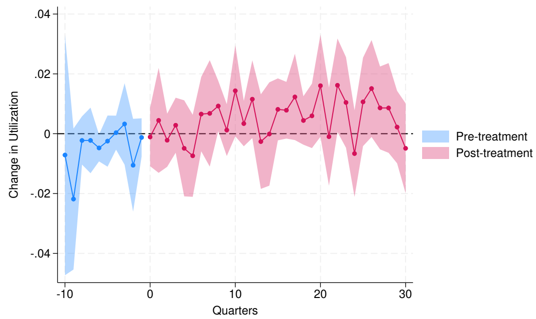
Fracking-related healthcare utilization event study: zip-code-level



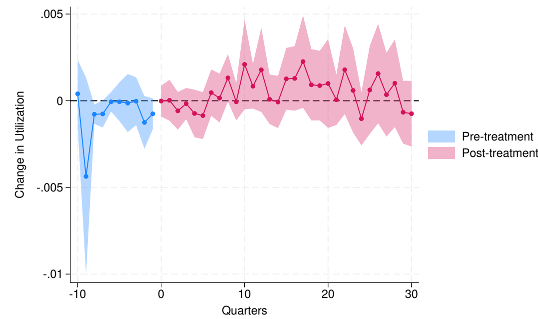
Notes: The figure above plots the coefficients and 95 percent confidence intervals for the event study that corresponds to the zip-code level analysis.

Figure A2: Condition-level healthcare utilization event studies: zip-code-level

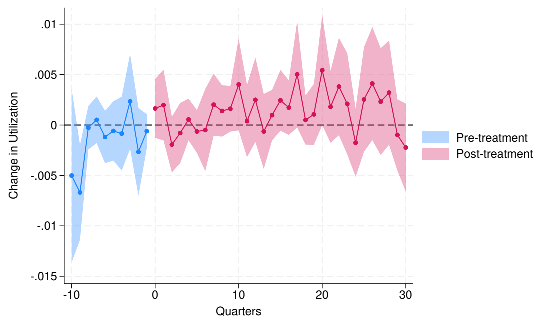
(a) Circulatory



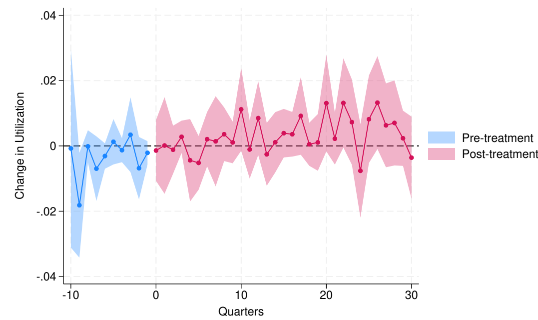
(b) General Illness



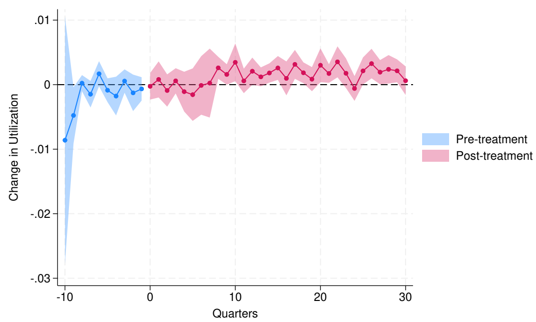
(c) Mental Health



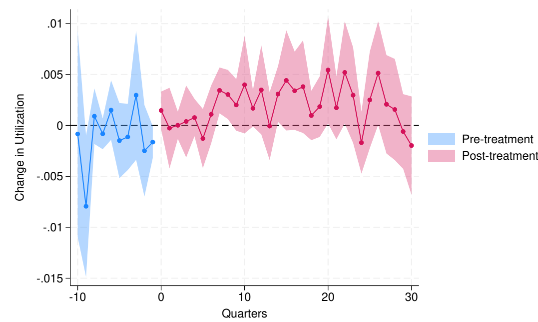
(d) Endocrine



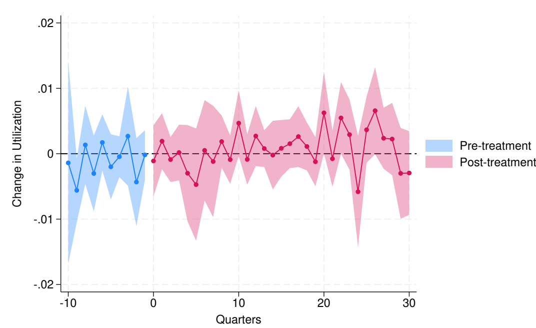
(e) Cancer



(f) Central Nervous



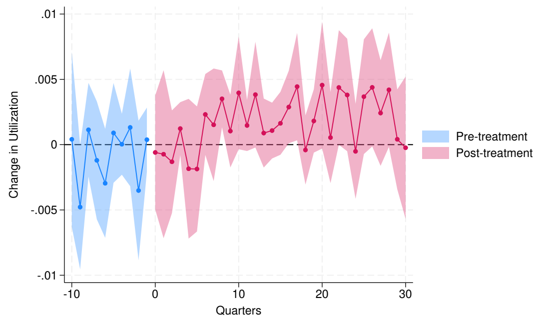
(g) Respiratory



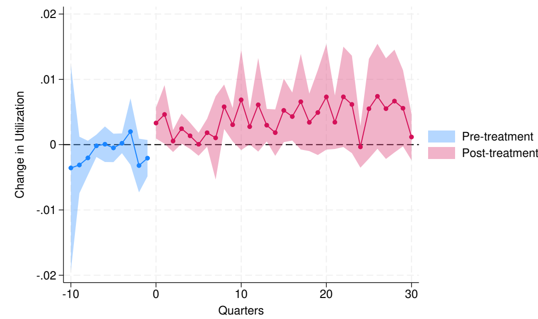
Notes: The figure above plots the coefficients and 95 percent confidence intervals for the event study that corresponds to the zip-code level analysis for each individual condition.

Condition-level healthcare utilization event studies; continued

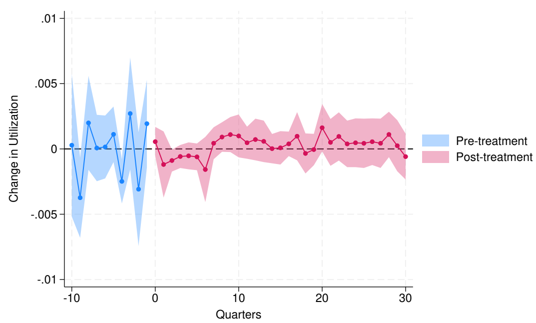
(a) Gastrointestinal



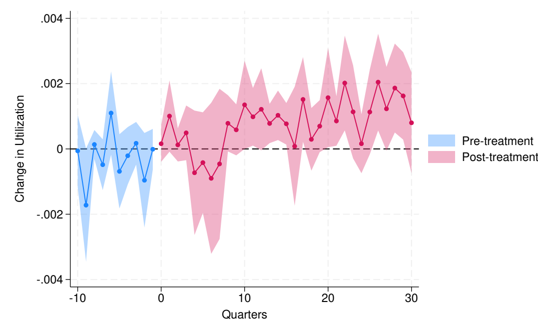
(b) Urological



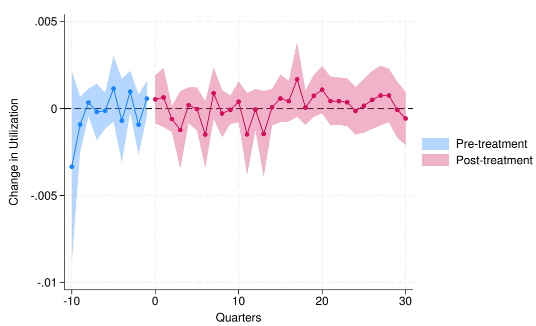
(c) Pregnancy Complications



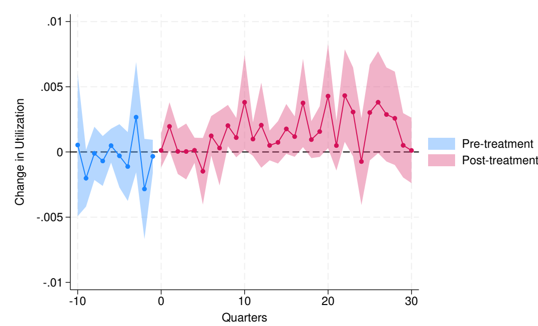
(d) Dermatological



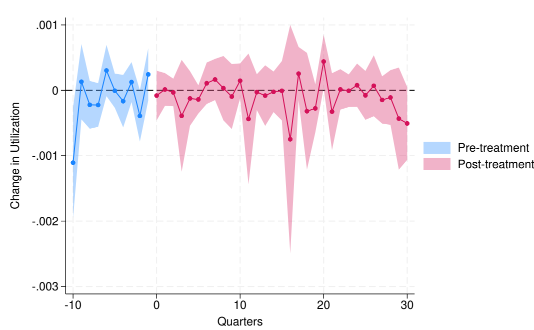
(e) Perinatal



(f) Blood Disease



(g) Congenital Abnormalities



Notes: The figure above plots the coefficients and 95 percent confidence intervals for the event study that corresponds to the zip-code level analysis for each individual condition.