Fracking and Tracking: The Effects of Oil and Natural Gas Well Locations on the Housing Market

Brent Norwood*, Ahmed El Fatmaoui, Samantha Johnson, and William Townsend

October 12, 2024

Abstract

Hydraulic fracturing or "fracking" has made it possible to produce vast new quantities of oil and natural gas causing states such as Colorado, Texas, and Oklahoma to dramatically increase their number of oil and natural gas wells. Using data from the U.S. Department of Homeland Security and Zillow's ZTRAX, we estimate the effect of hydraulically fractured oil and natural gas well sites on both urban and rural residential home prices between 2000 and 2018. Utilizing ArcGIS, we create varying buffer zone sizes around well sites to explore how average home prices change before and after a well opens and apply a ZIP code-level fixed effects model, a household-level fixed effects model, a repeat sales model, and a spatial differences-in-differences (SDID) approach. Our results show that homes within 0.5 miles of a well have a 4.5% increase in selling price and homes between 0.5-1 mile of a well see a 2.5% increase in selling price compared to properties located more than 2 miles away from a well.

Keywords: Hydraulic fracturing, Spatial dependence, Housing, Spatial Analysis
JEL Codes: Q35, Q33, R21

1 Introduction

Energy prices and their connection to other economic variables has been an area of interest for economists since the oil price shocks in the 1970s (Berndt and Wood, 1975). Over the past 25 years, there has been a significant increase in the construction and development of oil and natural gas well injection sites in the U.S. According to the US Energy Information Administration, hydraulically fractured and horizontally drilled wells are now the majority of all new wells drilled, and by 2016, 69% of all producing wells became hydraulically fractured and horizontally drilled (Fetzer, 2014). Hydraulic fracturing has led to a mining boom across the US as this new fracking technology, in combination with horizontal drilling, has made shale deposits that were previously unusable now profitable (Brown et al., 2014; Brown, 2015; Munasib and Rickman, 2015).

This progress is not without new challenges and risks. Over the past decade, news reports and politicians have negatively highlighted hydraulic fracturing [1] as it can cause earthquakes in heavily fracked areas (Frohlich, 2012). Leaking wells can also contaminate air and groundwater (Darrah et al., 2014; Muehlenbachs et al., 2015), and some have asserted the general presence of a well may be a dis-amenity for individuals (Popkin et al., 2013; Rivera, 2020). In 2013, The Wall Street Journal analyzed well location and census data for more than 700 counties in 11 major energy-producing states. They found that nearly 15.3 million Americans lived within one mile of a drilled well as of 2000 (Gold and McGinty, 2013).

Oil and gas development may also lead to serious local and global environmental concerns (Black et al., 2021; Cheung et al., 2018; Court et al., 2017; Garcia-Gonzales et al., 2019; Maguire and Winters, 2017; Stern, 2008; Torres et al., 2016), adverse health effects (Deziel et al., 2020; Hill, 2018; Black et al., 2021; Willis et al., 2022), and air pollution and water contamination (Jackson et al., 2014). This paper does not strive to ignore the environmental

[1] For example, both Bernie Sanders and Elizabeth Warren have outlined how they would end all fracking by 2025.

and health concerns regarding fracking nor make a definitive statement that fracking has an overall net positive effect on welfare, we are simply providing robust empirical evidence indicating that fracking is positively associated with housing prices.

Among the literature on hydraulic fracturing and home values, papers tend to focus on one or a few states within their empirical analysis (Muehlenbachs et al., 2015; Stephens and Weinstein, 2019) with mixed findings on the impact of fracturing and drilling on housing prices. Some research shows that fracking leads to lower home bids (Throupe et al., 2013), while other spatial econometric models provide no evidence of significant environmental impacts on housing values (He et al., 2018). Gopalakrishnan and Klaiber (2014) find that properties are adversely affected by proximate gas exploration sites/wells, and the home values were worse for properties with well water as a source of drinking water. Conversely, Muehlenbachs et al. (2012) showed that proximity to horizontal and vertical gas wells increases property values for homes with "piped water". These mixed results are due to several factors, including spatial heterogeneity, home amenities, varying government regulations, and regional practices.

Our paper contributes to the vast literature on fracking and property values by using a novel data set of wells and homes across 11 states over 18 years. This rich data allows us to test the impact of hydraulic fracturing and horizontal drilling on home values at a much larger scale than in previous work. With our novel dataset, we find a positive relationship between home value and proximity to wells using a panel data fixed effects model, a repeat sales model, and a spatial Difference-in-Differences model. We find that homes within 0.5 miles of a well have a 4.5% increase in selling price compared to those more than 2 miles away. In addition, we find homes between 0.5-1 mile from a well site experience a 2.5% increase in home value compared to homes that are more than 2 miles away.

Our findings also contribute to the literature covering the uncertainty of net benefits regarding oil and gas development on state and local economies(Cai et al., 2019; Marchand and Weber, 2018; Weinstein, 2014; White, 2012). Evidence indicates that increased oil and gas development could have positive effects for other industries and benefits from royalty recipients spending money on local goods and services (Brown, 2021; Whitacre et al., 2020). However, other research finds that increased oil and gas development could bid up local wages and other input prices. The higher local wages and input prices could make it significantly more difficult for other local export industries in the area to compete in global markets (Allcott and Keniston, 2018).

Among our sample, it is likely that the increase in housing values with proximity to wells stems from increased employment opportunities and wages in the short run. A large body of literature supports the argument that fracking leads to higher wages and employment prospects (Feyrer et al., 2020; Kearney and Wilson, 2018; Weber, 2012). There is a concern that in-migrants, temporary residents, and long-distance commuters may take most of the new oil and gas jobs (Gittings and Roach, 2020; Green et al., 2019; Jacobsen et al., 2023; Wilson, 2022; Wrenn et al., 2015); however, strong empirical evidence shows that fracking increases wage and job opportunities for long time local residents and in-migrant workers (Winters et al., 2021).

The remaining portions of the paper proceed as follows: Section 2 describes the background of hydraulic fracturing and some of the past literature, Section 3 elaborates on the data, Section 4 discusses the methodology, Section 5 presents the result, robustness checks, and various extensions, and Section 6 concludes.

2 Background and Related Literature

2.1 Hydraulic Fracturing and Horizontal Drilling

There have been two advances in well drilling technology that have made it economically meaningful to drill in shale deposit regions: hydraulic fracturing and horizontal drilling. Many areas across the U.S. have shale oil that is of low-pressure permeability, which means conventional drilling technology extraction would not be profitable. Hydraulic fracturing and horizontal drilling have alleviated these constraints. This new combination makes fracking extremely cost-effective as oil and gas companies can drill one well production pad that acts as multiple wells, making the above-ground footprint less invasive and eliminating costs of additional pads. Additionally, if the well connects directly to a pipeline, there is no need for above-ground storage tanks to hold the oil or gas.

According to Johnson et al. (2008), fracking starts by obtaining roughly half an acre of land to establish a pad foundation before using a 150-foot-tall drilling derrick to bore the well. They then drill multiple horizontal wells in any direction below the newly created vertical well. These horizontal wells then come in three varieties depending on the distance the well takes from ground zero to the depth of the targeted area. [2]

According to Halliburton, one of the largest fracking companies in the world, the process begins with a company drilling a 1,000 to 4,000 foot hole. Next, they place steel casings into the well and fill the space between the casing and the hole with cement. This process is repeated several times as the casings get smaller to prevent gas leaks and ensure groundwater protection. The wells eventually reach a total depth of 6,000 to 10,000 feet to access the gas.[3] After the depth and casings are complete, the driller inserts a perforated pipe gun into the horizontal part of the hole that produces explosions. These explosions create fractures in the shale. Next, they mix fracking fluids (consisting of 3 to 5 million gallons of water) with chemicals and sand and then pump the mixture at high pressure into fractures [4]. Once someone starts fracking, it takes about 3-10 days to start producing.

After fracking is complete and the drilling equipment is removed, the only visible structure above ground is 5-6 feet of surface valves left behind. On average, this well can go two years before it needs to be extensively serviced. It can be re-fracked many times over significantly

[2]Long radius wells, which require more than a thousand feet to reach 0-degree level (targeted area), and short and medium radius horizontal wells, that have sharper turns.

[3]The depth of 6,000 to 10,000 is below most aquifers

[4]The fracking fluid is 98% to 99.5% water and sand used to keep the cracks open. The remaining percent comprises chemicals that help reduce friction, kill microbes that might clog wells, and stop corrosion on and around the pipes. The exact chemicals used are usually unknown because the industry considers its fracking fluid formulas to be trade secrets, and it also depends on the particular conditions at the well (CHO (2014))

reducing the cost of drilling new wells and making it easier to place well sites in denser urban areas. Note, well location is not random. With the ease of horizontal drilling, companies can locate the wellhead closer to highways/roads for ease of access while still drilling horizontally in multiple directions to access the target areas. In residential areas, this can mean companies will be vying for the same land as home builders.

2.2 Previous Literature

Hydraulic fracturing has the potential to provide many benefits to local economies. Homeowners that have mineral rights can get royalty payments from the drilling and extraction of oil and gas.[5] Vissing (2015) evaluated factors driving tract-level heterogeneity in the caliber of leases executed by property owners who transfer their mineral rights to firms that drill and extract natural gas. Timmins and Vissing (2015) use a dual-gradient hedonic model to measure the capitalization of lease clauses into housing values. They analyze the effect of spatial proximity to hydraulically fractured wells and lease quality on houses' appraisal values.

Other broader benefits include increases in employment both within and outside of the mining sector. While some research argues that fracking does not increase employment or income (Paredes et al., 2015), a significant portion of the literature shows local communities around shale deposits can see an increase in employment in the mining community, as well as have significant spill-over effects into different sectors at the locations where resource extraction takes place (Winters et al., 2021; Fetzer, 2014; Weber, 2012, 2014; Weber et al., 2014). Maniloff et al. (2014) uncovered that counties with shale development had 6% higher wages. In other studies, findings suggest wages in the oil and gas sector increased by approximately 30%, wages in retail increased by 6%, and wages in the hotel sector increased by 17% (Marchand, 2012; Mason et al., 2015; Weber, 2012, 2014). These wage increases were not solely short-term gains. Feyrer et al. (2017) found that roughly two-thirds of the

[5] According to Reuters, almost all home builders now retain the mineral rights under homes they construct and the royalty payments would go to the builders and not homeowners

wage income increase persists for two years. Currently, employment in the mining sector has reached levels the public has not seen since the early 1990s (Fetzer, 2014). Fetzer (2014) found that each oil and gas sector job creates about 2.17 other jobs. Aside from creating thousands of jobs, fracking has also helped secure US energy security[6] and can make the US economy less carbon intensive, as it has relied less on coal and more on natural gas production (U.S. Energy Information Administration, 2018).

Conversely, hydraulic fracturing has the potential to cost a lot. When someone drills and fractures a well, there can be an increase in noise and air pollution, traffic, potential spills, and other environmental hazards (Lipscomb et al., 2012). The greatest danger comes from wells leaking and potentially contaminating groundwater which decreases home values (Zabel and Guignet, 2012). Beyond spill concerns, the oil and gas industry equipment might bring other dis-amenities that may affect a home's sale price. While drilling newer well heads leaves a smaller above-ground footprint, the drilling process itself has a large above-ground footprint and may be unsightly (Lipscomb et al., 2012). Gopalakrishnan and Klaiber (2014) show that increased truck traffic and noise during construction and fracturing may annoy homeowners. Fear of pipeline explosions may drive some households to not want to live near well heads, but Boxall et al. (2005) found no evidence that pipelines negatively affect sales prices.

Previous studies on the effect of hydraulic fracturing on home values have been largely concentrated in the Marcellus Shale of Pennsylvania and New York and the Barnett Shale of Texas by the greater Dallas metro area. With the Marcellus Shale, the results are mixed. Two papers by Muehlenbachs et al. (2012, 2015) use a triple Difference-in-Differences approach to estimate the effect of fracking on home prices. The authors found that the adverse effects of hydraulic fracturing are felt mainly by rural homes dependent on groundwater, while the effect for homes with piped water is slightly positive. In another study focused on Washington County in Pennsylvania, Gopalakrishnan and Klaiber (2014) found that homes

[6]This boom in the oil and gas industry helped make the United States the top oil-producing country in the world, accounting for 18% of world oil production (U.S. Energy Information Administration, 2018).

feel the negative effects of fracking on groundwater during the well construction phase. Still, the effects vastly decrease once well construction has finished and when the distance from the well increases. Delgado et al. (2016) focused on two counties in northeastern Pennsylvania, failing to find strong evidence that fracking significantly reduces home values. Following that, Boslett et al. (2016) utilized a drilling suspension in New York and found the likelihood of shale oil and gas drilling increases the value of properties. Ooms and Tracewski (2011) discovered that counties with an increase in wells also saw an increase in housing demand.

Weber et al. (2014) focused on the Barnett shale of Texas and found that property values in shale-producing ZIP codes appreciated for a decade relative to non-shale ZIP codes. They show the increase in home prices is attributable to the effect shale gas development has had on public finances. The increased revenues have allowed for more public spending while maintaining lower property taxes. Balthrop and Hawley (2017) looked at the exploit variation in distance to nearby gas wells in home sale prices to estimate the effect. They looked at an urban area and found that wells within 3,500 feet reduce property values by 1.5%-3.0%.

Currently, the net effect on higher-density areas is still unclear. In these areas, homeowners face the full cost and spillovers of being in a spatial location close to hydraulic fracturing. It could be that higher density areas make homeowners less sensitive to nearby drilling, similar to findings by Raimi et al. (2020). Their paper argues that because someone can use numerous alternative roads during construction, homeowners around wells do not see any adverse effects from additional wells.

Our paper fits into the broader recent research that looks at the housing market and the amenities or dis-amenities that oil and gas wells have on home value ((Bartik et al., 2019; Loomis and Haefele, 2017; Gregory et al., 2011). Currently, there is no clear consensus on the short or long-term effects since many studies considering hydraulic fracturing and home prices have focused on small populations, rural areas, small sample sizes, or how groundwater treatment affects homes next to disposal sites (Boslett et al., 2016; Gopalakrishnan and

Klaiber, 2014; Muehlenbachs et al., 2015). Our paper instead focuses on the broader cross-state effects of oil and gas well locations and their effects on the housing industry. Since our study covers 11 states involving both rural and urban areas, we can estimate the net effect of wells on higher-density settlements similar to Balthrop and Hawley (2017). This will allow for a more aggregate analysis of the long-run impacts of the fracking boom on the housing market than what the literature currently includes.

3 Data

3.1 Housing Data

For this paper, we gathered data from two main sources. The first comes from Zillow and its centralized source of property transactions through its Zillow Transaction and Assessment Dataset (ZTRAX). [7] This unique dataset consists of detailed household-level data across all of the 11 oil-producing states we analyze. It tracks individual and repeat sales with household-level characteristics such as square footage, number of rooms, bedrooms, bathrooms, spatial location, and sale price.

The ZTRAX dataset uniquely sets our paper apart from past research by not only providing the main outcome of interest (house prices) but also enabling us to avoid omitted variable bias in our specification through the inclusion of various house-related characteristics. Furthermore, the data encompasses all the geographical levels of interest[8]. To the best of our knowledge at the time of writing, ZTRAX is the only dataset available to academics that provides extensive details on the housing market in the US at various geographical levels, such as ZIP code and county. We consider all homes in each of the 11 states used in the sample if the Zillow housing data represents a state's housing market. We excluded no individual observation within a state, as every state we analyzed had the needed variables

[7]More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author and do not reflect the position of Zillow Group.

[8]CA, CO, KS, KY, OK, PA, TN, UT, VA, WV, and WY

to run the regression analysis. This is also a key advantage that sets our research apart from past research: the analysis of multiple states followed by individual state analysis.

Since we focus on the locations of wells and their effects on the housing market, we only consider homes Zillow documents as residential properties. Some states do report business, government, and other non-residential properties, but we exclude these observations. We also filter the data for situations where it is likely a non-market transaction. We only include observations categorized as deed transfers with exchanging a property's title from one party to another. To ensure that incorrect or improbable observations do not drive our results, we exclude transactions with sale prices below $10,000 or above $10,000,000, similar to Cheng et al. (2018). It is unlikely that transactions with prices falling below $10,000 occurred on the market[9], and transactions with prices above $10,000,000 are extraordinary and are not representative of the state's housing market. [10] Additionally, we filter house characteristics to exclude observations in the top thousandth or top ten-thousandth percentile. This filtering, for example, eliminates any observation with over 1,000 bedrooms. This also removed many observations in states that do not require counties to report the home characteristics, such as Utah and Wyoming. In the Appendix, we provide a more comprehensive examination of our data cleaning process for the Zillow data (see table 15).

3.2 Oil and Natural Gas Well Data

We acquire the second data set containing the well information from the U.S. Department of Homeland Security. It provides a mostly complete listing of the oil and natural gas well locations across the U.S.[11] This data set contains just over 1 million well locations and has detailed information such as the wells operators, location by longitude and latitude, approval

[9]Such transactions may have slipped through the DataClassStndCode filter

[10]These could also be due to an error in the data entry

[11]Geographic coverage includes the United States (Alabama, Alaska, Arizona, Arkansas, California, Colorado, Florida, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maryland, Michigan, Mississippi, Missouri, Montana, North Dakota, Nebraska, Nevada, New Mexico, New York, Ohio, Oklahoma, Oregon, Pennsylvania, South Dakota, Tennessee, Utah, Virginia, Washington, West Virginia, Wyoming) as well as oil and natural gas wells in the Canadian provinces of British Columbia and Manitoba that are within 100 miles of the country's border with the United States.

date, status, and type of well. Note that companies are not required to give public notice about their intent to drill, but permit records are public information. While this information is accessible to individuals, it is located in a state-level database on the agency's website of whomever is in charge of oil and natural gas production. [12] This means that unless an individual is explicitly looking for this information, they would not know if/when someone could place a well near their home suggesting that home prices should not be affected before a well is built since there is little foreknowledge of the drilling.

Figure 1 shows the number of wells the U.S. drilled between 2000 and 2018, along with that year's average crude oil closing price. Between 2000 and 2008, the U.S. gradually increased their yearly wells drilled, spiking between 2006 and 2008, which is expected as oil prices hit record highs. The U.S. then slowed down its drilling after oil prices went from a yearly closing average of \$99.67 in 2008 to \$61.95 in 2009. The biggest one-year jump in well drilling came in 2014 when the U.S. drilled 33,285 wells, corresponding to a spike in oil prices when the yearly closing average went back up to \$93.17. In the following years, the U.S. drastically slowed down its well drilling, partially due to OPEC and Russia flooding the oil markets near the end of 2014. This resulted in the average yearly closing price of oil in 2015 dropping to \$48.72. Figure 1 also shows how quickly the oil and gas sector can respond and quickly ramp up production or just as easily slow it down.

[12] The agency in charge of oil and gas law and production varies from state to state. In Oklahoma, it is the Oklahoma Corporation Commission, but in Louisiana, it is the Department of Natural Resources.

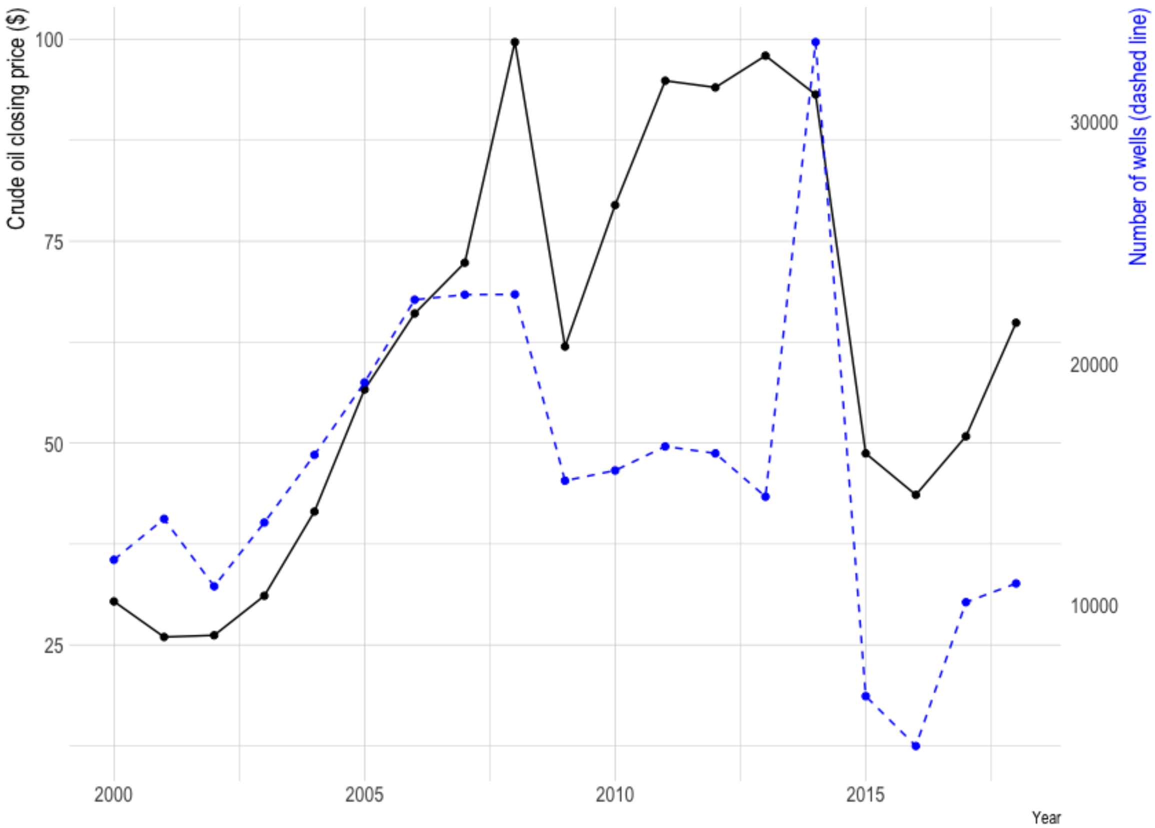


Figure 1: Well and Crude Oil Price Trends

The crude oil closing price on the left side vertical axis refers to the average crude oil closing price (the solid line). The right vertical axis shows the number of wells drilled per year (the dashed line). These are the number of new wells drilled per year during the sample period regardless of the method used, fracking or otherwise. The significant drop-off after 2014 occurred when the oil and gas market bottomed out from a decrease in demand and OPEC deciding not to cut production. Overall the oil prices and the number of wells seem to have a common trend.

3.3 ArcGIS

Using each house and well longitude and latitude, we utilized ArcGIS to spatially plot the wells on a map of the U.S. and create various-sized buffers around each well. By selecting wells used from that year and all previous years, we were able to identify established wells and correctly isolate the number of new wells and their location each year. Once we created layers for each year, we imported the housing transactions. Again, we isolated the homes sold each year to create a layer that just contained each year's housing transactions. After creating each yearly layer, we matched corresponding well and housing years together to separate the homes that fell inside each well's half mile buffer. We used the same process to identify homes that fell between 0.5-1 mile and again for homes that fell between 1-2 miles of a well. Figure 2 shows a map of all wells used in the sample.

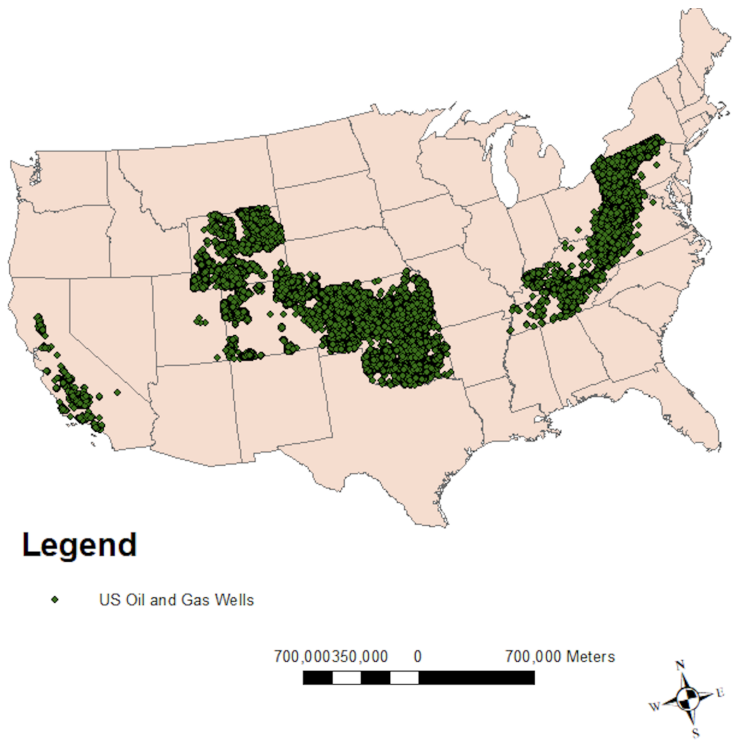


Figure 2: Oil and natural gas well locations

We created this map using ArcGIS and show the location of all wells in our data's sample.

Once we merged these geographical levels and periods with all the house variables from ZTRAX with oil and natural gas well data, we created summary statistics in Table 1 below.

Table 1: State Housing Summary Statistics

	Count	Home Price	Price/ SqFt	Rooms	Bedrooms	Bathrooms	SqFt	Year Built	Wells
CA	8,503,009	434,726.30	251.03	2.97	3.12	2.21	1,747.96	1977.32	12,201
		(445,818.60)	(189.58)	(3.544)	(0.95)	(0.84)	(5,307.71)	(24)	
CO	1,771,029	290,086.40	171.45	3.70	2.96	2.41	1,882.01	1981.	47,374
		(289,518.60)	(398.42)	(3.17)	(0.98)	(1.07)	(1,113.15)	(26)	
KS	45,683	114,765.80	77.42	6.04	3.02	1.90	1,427.10	1960	86,790
		(149,634.10)	(100.15)	(1.53)	(0.85)	(0.91)	(628.88)	(28)	
KY	359,049	174,939.90	93.41	2.39	2.93	2.11	1,840.95	1980	12,297
		(20,0437.80)	(107.12)	(3.14)	(0.90)	(0.85)	(1,160.86)	31	
OK	555,914	133,791.80	71.94	5.71	3.07	1.86	1,809.81	1979	25,605
		(168,148.50)	(95.29)	(2.30)	(0.69)	(0.64)	(824.50)	26	
PA	2,183,286	196,639.60	105.46	6.36	3.13	1.79	1,804.77	1960	10,6073
		(200,441.00)	(88.88)	(1.96)	(.80)	(0.85)	(941.01)	(36)	
TN	874,251	170,397.20	74.99	6.33	3.08	2.10	2,256.33	1978	13,072
		(194,195.00)	(72.81)	(2.05)	(0.80)	(0.89)	(1,322.68)	(26)	
UT	39,461	168,552.10	118.36	8.75	3.69	2.02	1,421.36	1974	11,949
		(145,809.10)	(65.03)	(2.78)	(1.18)	(0.87)	(711.17)	(27)	
VA	1,967,558	295,113.90	151.22	5.02	3.33	2.55	2,024.50	1984	10,132
		(233,333.70)	(96.39)	(3.33)	(0.89)	(0.97)	(1,035.81)	(23)	
WV	127,345	142,562	77.53	6.28	3.02	1.94	1,829.23	1969	113,463
		(166,406.80)	(98.04)	(1.62)	(0.76)	(0.84)	(911.45)	(33)	
WY	3,529	171,798.60	117.86	4.98	3.09	1.99	1,478.70	1970	51,325
		(262,969.70)	(144.82)	(3.10)	(0.94)	(0.80)	(1,591.91)	(29)	
Total	16,430,114	337,004.40	190.11	4.05	3.12	2.20	1,833.14	1976	29,720
		(374,030.30)	(208.45)	(3.47)	(0.92)	(0.92)	(3,890.13)	(27)	(33,155)

Variable means are presented with standard deviations in parenthesis below. Observations are not aggregated so house price is the i_{st} level where i is household, s is the state, and t is the exact date of house contracts. Years 2000-2018 are included in the sample. Column 9 lists the number of oil and gas wells that are located in the state according to the data provided by the U.S. Department of Homeland Security.

Table 2 looks at the summary statistics for houses in ZIP codes with wells present relative to houses in ZIP codes that do not. Ideally, the means of these two samples should be similar, and we note that this is true for every variable measured in Table 2. Further details about the summary statistics by buffer zone are provided in the appendix Table 13.

Table 2: State Base Treatment Summary Statistics

	Home Price	Price/SqFt	Rooms	Bedrooms	Bathrooms	SqFt	Year Built
Homes in Control ZIP's							
Mean	339,859.60	189.95	4.65	3.14	2.19	1,856.55	1977
(SD)	(360,985.20)	(167.61)	(3.43)	(0.93)	(0.90)	(5,168.10)	(28.45)
Homes in Treatment ZIP's							
Mean	333,727.20	190.39	3.31	3.09	2.21	1,804.63	1975
(SD)	(389,580.90)	(249.65)	(3.37)	(0.90)	(0.93)	(957.32)	(26.32)
Total Sample							
Mean	337,108.40	190.15	4.05	3.12	2.20	1,833.25	1976
(SD)	(374,097)	(208.45)	(3.47)	(0.92)	(0.92)	(3,890.81)	(27.53)
Observations	16,424,190						

Table 2 looks at the summary statistics between houses in ZIP codes that have wells present to houses in ZIP codes that do not. The means of both groups should be similar; otherwise, if the treatment and control samples differed in the observed variables, the samples might also vary in correlated unobserved variables, which could bias the estimates. As the table shows, they are quite similar.

4 Methodology

The primary goal of this paper is to find the average effect of oil and natural gas well locations on the housing market as well as assess if oil/gas sites drive away consumers, the consumers are indifferent to the well sites, or if the oil/gas sites increase local economic activity, thus increasing housing demand. To test this hypothesis, we used three approaches. First, we start with a panel data fixed effects model, followed by a repeat sales model, and finally a spatial Difference-in-Differences. First, consider the following fixed effects model:

$$\log(Price_{ijt}) = \alpha_1 Treat_{it} + \alpha_2 Treat1_{it} + \alpha_3 Treat2_{it} + \alpha_4 Treat3_{it} + \gamma_1 Z_{ij} + \gamma_2 X_{ijt} + \delta_j + \rho_t + \epsilon_{ijt} \quad (1)$$

where the primary dependent variable $\log(Price_{ijt})$, is the logged house price of home i, in ZIP code j, at time t. Independent variables of interest are $Treat_{it}$, $Treat1_{it}$, $Treat2_{it}$, and $Treat3_{it}$. $Treat = 1$ if the home is located within 2 miles of a well, $Treat1 = 1$ if the home is located with 0.5 miles of a well, $Treat2 = 1$ if the home is located within 0.5-1 mile of a well, and $Treat3 = 1$ if the home is located within 1-2 miles of a well. Additionally, Z_{ij} is a vector containing traditional housing characteristic controls such as the number of rooms, bedrooms, square feet, and property age. X_{ijt} refers to local characteristics, including state GDP, state population, state land area, county poverty counts, and county median household income to control for state and county-level time-varying demographic factors. Finally, δ_j is ZIP code level fixed effects, and ρ_t is a month-year fixed effect. We cluster the standard errors by ZIP code to allow for arbitrary heteroskedasticity and serial correlation within the same ZIP code over time.

In our analysis, we use these various ring boundaries to examine the sensitivity of the threshold definition for the treatment group to the measured treatment effect. These boundaries also align with the thresholds used by Balthrop and Hawley (2017) and Muehlenbachs et al. (2015), who used 3,500, 5,000, or 6,500 ft (in miles, 0.5, 0.94, and 1.23) and 1 km, 1.5 km, and 2 km. (in miles, 0.62, 0.93, and 1.24). Gopalakrishnan and Klaiber (2014) used boundaries between 0.75 miles and 2 miles, while Delgado et al. (2016) used between 1 and 4 miles. Since Gopalakrishnan and Klaiber found significant fading of the treatment effect by 2 miles and Delgado et al. (2016) estimated that the treatment effect reduces to zero around 4 miles, we stop our boundaries at 2 miles.

We divided our data into sub-samples to check for potential sources of bias. Table 3 shows the total number of treatment homes for each state and the total number of treatment homes in one of the three buffer zones per state. States such as Utah and Virginia have low treatment-to-control ratios, but we included them in our initial analysis. For a robustness check, we will drop states with a low treatment to control ratio and re-estimate later in our analysis.

Table 3: State Treatment Summary Statistics

	Treat	Treat1	Treat2	Treat3	Non-Treat
CA	64,100	1,678	11,585	50,837	8,438,766
CO	80,430	14,144	240,10	42,276	1,690,358
KS	1,932	44	187	1,701	45,651
KY	676	49	138	489	356,959
OK	12,057	1,218	2,194	8,645	541,462
PA	60,674	5,474	19,067	36,133	2,122,528
TN	367	0	21	346	870,738
UT	11	0	0	11	39,802
VA	921	84	244	593	1,966,052
WV	15,226	1,635	3,500	10,091	111,244
WY	262	251	1	10	3974
Total	236,656	24,577	60,947	151,132	16,187,534

Treat 1 is any home in the sample within 0-0.5 of a mile of a well. Treat 2 is any homes between 0.5-1 mile of a well. Treat 3 is homes that are 1-2 miles from a well. Non-Treat are homes that are farther than 2 miles from a well.

In equation 1, we use both ZIP code or household-level fixed effects to address the possible existence of unobservable characteristics of the properties that might be correlated with observed variables including the three boundary zones. After that, we restricted the sample to only repeat sales. We use this approach as there are two distinct advantages of a repeat sales model. First, we can run the model with ZIP code and household-level fixed effects.

Second, we can differentiate out the unobservables effect, assuming they are constant at both sale points. Balthrop and Hawley (2017) also used a repeat sales model, but when they restrict their observations to only repeat sales, their sample is reduced by 64%. Since our panel covers many years, our sample is only reduced by 26.6% with just over 9.5 million transactions remaining. We follow a repeat sales model similar to Balthrop and Hawley (2017) with the estimating equation:

$$\log(Price_{ijt+\theta}/Price_{ijt}) = (\alpha_1 Treat_{it+\theta} - \alpha_1 Treat_{it}) + \gamma_1 Z_{ij} + \gamma_2 X_{ijt} + \lambda_i + \rho_t + \epsilon_{ijt} \quad (2)$$

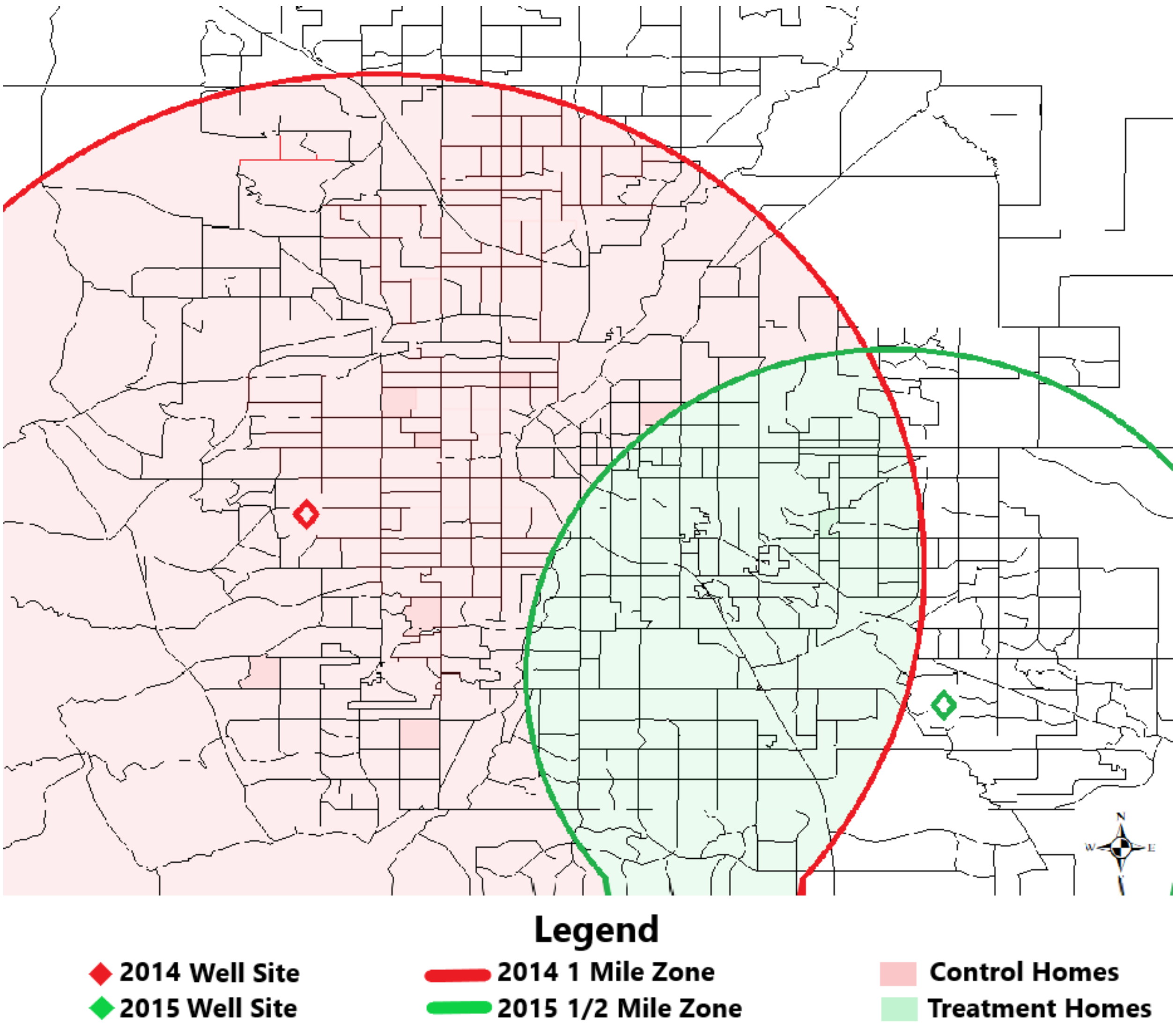
Each property we include in the repeat sales model has at least two transactions: the original transaction price at time t given by $Price_{ijt}$, and the second price at time $t + \theta$ denoted $Price_{ijt+\theta}$. With this method, the change in proximity to a well can explain the percentage change in price. In this model, we use λ_i first with ZIP code-level fixed effects and then again with household-level fixed effects.[13] We also cluster the standard errors in this model as we did in Equation 1.

A source of possible endogeneity in equations 1 and 2 is the non-random location of an oil/gas well site. Home builders/buyers and oil/gas companies are looking for places to build/buy/drill that have easy access to roads, interstates, and other access points, so the location of wellheads might affect the selling price of homes. To account for this possible endogeneity concern, we use a spatial Difference-in-Differences (SDiD) approach such as the one used in Dronyk-Trosper (2017). This study examined the local government's construction of public service facilities, such as fire departments and police stations, to see if they impacted the local housing market. In their application, control homes were homes that maintained their distance from the closest facility throughout the sample period. Treatment homes are ones at period t_0 that have the same distance as the control group, but at some future

[13] With this model, we assume that the home is not significantly altered during the sample period. The Zillow data does have data tracking if the house underwent any modifications, but due to reporting standards, not all states report it. For the states that did report, none showed any homes during the sample period with significant modifications. This is the same assumption that Balthrop and Hawley (2017) used in their model.

period t_s (where $s > 0$), a new public service facility reduces the distance to the nearest option. We follow the same intuitive logic but focus on the placement of new wells, leading to quasi-experimental differences in proximity to wells. Figure 3 illustrates these two groups.

Figure 3: Spatial Difference in Difference Illustration



The red diamond marks a well constructed in 2014, and the red ring is the 1-mile buffer zone that captures all homes. The green diamond is the location of a well built the following year and has a 0.5 mile buffer zone. The intersection of the 2, the green shaded area, are the new treatment homes, and the red shaded area are the control homes.

Homes in the control group are in the red 1 mile zone with no change in their distance to the construction of a new oil/gas well. We compare these to the treatment homes in the original 1 mile zone but move closer to the 0.5 mile zone when the new oil/gas well is constructed. This method provides additional clarity as to the probable pathway of housing capitalization effects. Significant results here would help strengthen the idea that other significant findings are not likely from a spurious effect relating to oil/gas well location choices. Equation 3 represents our SDiD model:

$$\log(Price_i) = \beta_1 Treatment_i + \beta_2 BaseTreat_i + \beta_3 (Treatment_i * BaseTreat_i) + \gamma X_i + \epsilon_i \quad (3)$$

In this model, $Treatment_i$ is an indicator variable that reflects whether a home is in one of the treatment groups, $BaseTreat_i$ is a dummy interaction term for whether a home sale occurred before or after the construction of a new oil/gas well, and X_i is a vector of home characteristic controls. β_3 is our variable of interest, which represents the change in home values for treated units following the opening of a new well site. A significant value here would demonstrate that constructing a new well site altered local housing values. Again, the standard errors are clustered at the ZIP code level. Lastly, to round out the analysis further, we run Equations 1 and 3 individually for each state to analyze results at a more micro/state level.

To run the models at the state level, each state needed to have enough observations, and we could not skew the treatment-to-control ratio too far to the control side. This restriction did not allow us to analyze some states in the main models. States that we could analyze include California, Colorado, Oklahoma, Pennsylvania, and West Virginia. Running the model at the state level is essential because regulations, finances, local zoning, and noise regulation impacting the amenities/dis-amenities from hydraulic fracturing may vary from state to state [14]. Given all of these differences, conducting a state-by-state analysis will help give more insight into how regulatory differences may affect the oil and gas industry which

[14]It is not clear which state has more overall stringent regulations. Pennsylvania, for example, has longer setback restrictions on wells from buildings and water sources, whereas California has no rule regulating setbacks. Oklahoma has more restrictive venting and flaring regulations, whereas Colorado requires more stringent pre-drilling water testing requirements (Richardson et al., 2013).

could translate into housing market impacts.

5 Results

5.1 ZIP Code Level Fixed Effects and SDiD

Table 4 exhibits the estimated results for Equations 1 and 3. Columns 1 and 2 show the results for the ZIP code level fixed effect model when we include all states with clustering at the ZIP code level. Columns 3 and 4 show the SDiD results. The fixed effects model has positive results, but they are not significant. However, turning to columns 3 and 4, the SDID models show significantly positive results. When comparing homes that started off 2 miles from a well site to those that gained a new well, decreasing the distance to between 0.5-1 mile, there is an increase in selling price of 1.9%. Additionally, homes now within 0.5 miles from a well see an increase of 6%. Column 4 shows homes that were once between 0.5-1 mile of a well to now within 0.5 miles of a well increase their selling price by 4.5%.

Table 5 shows the results when we remove states with low treatment counts and balance the overall sample. Now, both the ZIP code level fixed effects model and the SDiD show positive and significant results. Like before, columns 1 and 2 show the ZIP code level fixed effect model results with clustering at the ZIP code, and columns 3 and 4 show the SDiD results. In this case, columns 1 and 2 still show increased point estimates, but they have gained significance. Column 1 shows that homes within two miles of a well get an increase in selling price of 1.6%, or roughly a $5,393.73 increase based on the mean transaction price. Homes that are within 0.5 miles of a well experience a 4.5% increase, homes that are 0.5-1 mile from a well see a 2.5% increase, both significant at the 0.05 level, and homes that are 1-2 miles from a well increase their selling price by 0.9%. The point estimates decreasing as the distance from the well increases follows the assumption that both well sites and home sites or owners compete for the same ease of access to roads, highways, and interstates.

Table 4: Ln HP Results

	Outcome variable: LnHP			
Independent Variables	(1)	(2)	(3)	(4)
Within 2 Miles	0.006			
	(0.009)			
.5 Mile Zone		0.029	0.060***	0.045***
		(0.018)	(0.016)	(0.014)
.5-1 Miles Zone		0.012	0.019*	
		(0.010)	(0.010)	
1-2 Mile Zone		0.000		
		(0.009)		
Rooms	0.005***	0.005***	0.015***	0.015***
	(0.001)	(0.001)	(0.003)	(0.004)
Bedrooms	0.131***	0.131***	0.177***	0.154***
	(0.006)	(0.006)	(0.022)	(0.030)
Ln SqFeet	0.716***	0.716***	0.622***	0.554***
	(0.008)	(0.008)	(0.036)	(0.052)
Year Built	0.002***	0.002***	0.004***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)
Ln State GDP	1.883***	1.883***	1.851***	1.766***
	(0.079)	(0.079)	(0.156)	(0.214)
Ln State Population	0.360***	0.356***	-0.240	-0.097
	(0.115)	(0.115)	(0.271)	(0.280)
Ln State Area	-8.522***	-8.512***		
	(0.366)	(0.367)		
Ln County Poverty Count	0.017***	0.017***	-0.001	-0.012**
	(0.002)	(0.002)	(0.005)	(0.006)
County Median House Income	0.000***	0.000***	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Sample	Full	Full	Base Treat 3	Base Treat 2
Observations	15,875,600	15,875,600	986,540	478,250
R-squared	0.734	0.734	0.704	0.663
ZIP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered S.E.	YES	YES	YES	YES

Results are in the Logged value of the sales price (LnHP), and the standard errors are clustered at the ZIP code level. Columns 3 and 4 are the spatial DID results. Column 3 shows when we restrict the sample to just Base Treat 3 homes, and it displays the effect of when a home is initially in the 2-mile zone and then moves to the 0.5-1 or 0-0.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 4 is the same but represents homes that were once in the 0.5-1 mile zone and are now in the 0-0.5 mile zone.

Table 5: Ln HP Results dropping low Treatment States

	Outcome variable: LnHP			
Independent Variables	(1)	(2)	(3)	(4)
Within 2 Miles	0.016*			
	(0.009)			
.5 Mile Zone		0.045**	0.029*	0.028**
		(0.018)	(0.015)	(0.014)
.5-1 Mile Zone		0.025**	0.007	
		(0.011)	(0.009)	
1-2 Mile Zone		0.009		
		(0.009)		
Rooms	0.008***	0.008***	0.016***	0.017***
	(0.001)	(0.001)	(0.003)	(0.003)
Bedrooms	0.132***	0.132***	0.097***	0.091***
	(0.007)	(0.007)	(0.012)	(0.013)
Ln SqFeet	0.725***	0.725***	0.669***	0.662***
	(0.009)	(0.009)	(0.016)	(0.017)
Year Built	0.002***	0.002***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Ln State GDP	2.387***	2.388***	2.156***	2.050***
	(0.074)	(0.074)	(0.130)	(0.147)
Ln State Population	-0.481***	-0.487***	-0.295**	-0.218*
	(0.112)	(0.112)	(0.115)	(0.118)
Ln State Area	-8.013***	-8.000***	-7.492***	-7.303***
	(0.437)	(0.437)	(0.416)	(0.434)
Ln County Poverty Count	0.024***	0.024***	-0.006**	-0.007***
	(0.003)	(0.003)	(0.003)	(0.003)
County Median House Income	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Sample	Restricted	Restricted	Base Treat 3	Base Treat 2
Observations	13,082,632	13,082,632	5,000,861	4,712,508
R-squared	0.743	0.743	0.663	0.651
ZIP FE	YES	YES	YES	YES
Year-Month FE	YES	YES	YES	YES
Clustered S.E.	YES	YES	YES	YES

Results are in the Logged value of the sales price (LnHP), and the standard errors are clustered at the ZIP code level. Homes with low treatment counts have been restricted from the sample. These include KS, KY, TN, UT, VA, and WY. Columns 1 and 2 are the fixed effects model results. Columns 3 and 4 are the spatial DID results. In column 3, we restrict the sample to just Base Treat 3 homes, and it looks at the effect when a home is initially in the 2 mile zone and then moves to the 0.5-1 or 0-0.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 4 is the same, but looking at homes that were once in the 0.5-1 mile zone and now are in the 0-0.5 mile zone.

The SDiD results in columns 3 and 4 using the restricted sample show similar results to the full sample results in Table 4. While the point estimates are slightly lower, most of the significance continues. Using the restricted sample, homes that started two miles from a well and then experienced a new well drilling, decreasing the distance to be less than 0.5 miles, show a 2.9% increase in column 3 and significance at the 0.10 p-value. In column 4, results for homes once 1 mile from a well and transitioned to within 0.5 miles of a well experienced a 2.8% increase in selling price with significance at the 0.05 p-value. These results align directly with what Fetzer (2014), Weber (2012), Weber et al. (2014), and Muehlenbachs et al. (2012) show and help support the hypothesis that the boom in hydraulic fracturing has increased the local economy through increases in wages, employment, and local tax revenue among our sample which helps increase the demand for housing. There is some concern that hydraulic fracturing effects on home values differ during the life-cycle of the well, but Balthrop and Hawley (2017) looked at the impact of drilling on sales within six months of well completion and sales made shortly after well drilling. They find that well construction does not seem to be the primary driver of any amenity or dis-amenity on a home's selling price. The data also supports the argument that new well sites and homes are attracted to the same plots of land with easy access. This also hints at the possibility that potential home buyers are not "scared" away by oil/gas well sites as fracking continues and grows, becoming a part of our normal reality.

5.2 Household Fixed Effects and Repeat Sales

Next, we focus on Equation 1 using household-level fixed effects and Equation 2 analyzing repeat sales with ZIP and household-level fixed effects. These results are in Table 6. These approaches allow us to better control for constant and unobserved neighborhood quality than the previous estimates. In addition, they should produce the same, or nearly the same, point estimates as analyzing all transactions using a household-level fixed effect should only look at homes that have been sold more than once. For these reasons, it is the preferred specification.

Columns 1 and 2 show the results of applying household-level fixed effects. Looking at the entire sample, homes within 2 miles of a well still experience a positive and significant increase in their selling price by 3.9%, shown in column 1. Balancing the sample and taking out low treatment states, the point estimate increases to 4.9% in column 2 and becomes more significant, corroborating our earlier findings.

In columns 3 and 4, we focus on repeat sales. We identify these estimates as a home that goes from being greater than 2 miles from a well to being within 2 miles of a well. This is different than Equation 1 because it does not rely on spatial differences in well exposure, which requires comparing homes that may be in different neighborhoods. When we restrict our sample to repeat sales, excluding the states with a low treatment count, we have a sample of $9,594,780$ observations. The richness of our data allows for a more in-depth examination than previous repeat sales models, allowing a significant contribution to the literature [15]. In column 3, we use ZIP code-level fixed effects, and in column 4, we use household-level fixed effects. We cluster the standard errors at the ZIP code level in both models. The point estimates for both repeat sales models are essentially the same compared to the entire sample in columns 1 and 2. This was expected since using household-level fixed effects in Equation 1 will automatically consider only repeat sales in the full sample. In addition, restricting the sample to just repeat sales allowed for the use of two different layers of fixed effects. The outputs in Table 6 confirm the results from our outer estimations.

[15] As noted earlier, Balthrop and Hawley (2017) conducted a repeat sales model but lost 64% of their total sample when they restricted it to just repeat sales for a total of roughly $15,000$ observations

Table 6: Household ID and Repeat Sales Model Results

	Outcome variable: LnHP			
Independent Variables	(1)	(2)	(3)	(4)
Within 2 Miles	0.039*	0.049**	0.016*	0.049**
	(0.020)	(0.022)	(0.010)	(0.022)
Rooms	0.006	0.006	0.008***	0.006
	(0.005)	(0.005)	(0.001)	(0.005)
Bedrooms	0.131	0.135	0.142***	0.135
	(0.142)	(0.145)	(0.007)	(0.145)
Ln SqFeet	0.084	0.082	0.718***	0.082
	(0.151)	(0.155)	(0.010)	(0.155)
Year Built	0.000	0.000	0.000	0.000
	(0.002)	(0.003)	(0.000)	(0.003)
Ln State GDP	3.213***	3.278***	2.486***	3.278***
	(0.069)	(0.084)	(0.089)	(0.084)
Ln State Population	-2.852***	-2.866***	-0.513***	-2.866***
	(0.160)	(0.201)	(0.128)	(0.201)
Ln County Poverty Count	-0.163***	-0.196***	0.032***	-0.196***
	(0.007)	(0.009)	(0.003)	(0.009)
Ln County Median House Income	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Sample	Full	Restricted	Repeat Sales	Repeat Sales
Observations	15,875,600	13,082,632	9,594,780	9,594,780
R-squared	0.893	0.894	0.725	0.845
Household FE	YES	YES	-	YES
ZIP Code FE	-	-	YES	-
Clustered S.E.	YES	YES	YES	YES

Results are the Logged value of the sales price and the standard errors are clustered at the ZIP code level. These results represent when we use the household level fixed effects to capture how repeat sales of homes are affected when a well comes within 2 miles or closer. Column 1 is the full sample, whereas in column 2 we drop low treatment to control states. These states include KS, KY, TN, UT, VA, and WY. The results here help support the hypothesis that hydraulically fractured wells positively stimulate housing demand, and the boost to the local economy helps boost the demand for housing.

5.3 Individual States

The results in Tables 4 are across 11 states, and results in Tables 5 and 6 are across 7 states. Although state and local policies regarding hydraulic fracturing vary from state to state (Krupnick et al., 2015; Warner and Shapiro, 2013; Kaplan, 2014; Weber et al., 2016), it is still unclear which states have more stringent regulations and which have less. Some states might be strict on certain aspects but simultaneously more relaxed on others. As a robustness check, we estimate Equations 1 and 2 for each state, intending to provide a deeper view to decipher if there are any heterogeneous effects from state to state. Because some states have low treatment counts, this was impossible for all states in the sample.[16] Tables 7 through 11 (see appendix) show the results for California, Colorado, Oklahoma, Pennsylvania, and West Virginia.

Richardson et al. (2013) performed a state-by-state analysis that analyzed numerous different regulations by state and created a ranking for individual regulations but not a clear overall state-wide regulation ranking. Some regulations included well spacing, venting regulations, building setback requirements, flaring regulations, pre-drilling water well testing, severance taxes, and accident reporting requirements. For some of the regulations, they could determine how stringently states regulate. This includes setback restrictions, casing/cementing depth requirements, and command-and-control regulations that are quantitatively reported. Their analysis shows how widely the states vary in regulations that are beyond federal regulations.

The results from the state-by-state analysis reveal positive and significant results except for West Virginia, which showed negative and significant impacts. Starting with California and Equation 1, the results show positive and significant point estimates for all three buffer zones. The 0.5 mile zone has a 5.1% increase, the 0.5-1 mile zone has a 2.2% increase, and the 1-2 mile zone has a 3.8% increase. When we analyze the spatial model, there is no

[16] The same states we excluded from the total sample were the same that could not do an individual state analysis. These include KY, TN, VA, and WY.

significance, whereas, in Oklahoma and Colorado, which have a higher treatment-to-control ratio than California, there are positive and significant point estimates for both the ZIP code fixed effect model and the spatial difference in differences.

Pennsylvania, a state that past literature has focused heavily on, shows positive point estimates only in the SDID model. It is difficult to draw a clear reason why some states show significant point estimates in just one or both of the models. We hypothesize that it involves the state's regulation and its history with hydraulic fracturing. Pennsylvania encountered numerous groundwater contamination reports in the early 2000s, which might contribute to a negative overall impression of fracking. Since they have not had as many reports in recent years, homeowners around wells or homes that do not rely on groundwater might not see them as a dis-amenity anymore. This could explain why the SDiD shows significant point estimates but not with the overall fixed effect model. According to the US Energy Information Agency, Oklahoma accounted for 9% and 4.5% of the US natural gas and oil, respectively. Colorado accounted for 5.6% and 4.2% of the US natural gas and oil, respectively. These two states had a balance of both oil and natural gas, whereas states like California and Pennsylvania did not. In California, they accounted for 3.3% of the total US oil but only 0.8% towards natural gas. Pennsylvania was the country's second-largest gas-producing state, accounting for 16%, but only 0.1% went towards oil.

A similar result was found for West Virginia which accounts for 12.6% of total US production with 5.5% and 0.4% towards gas and oil, respectively. Looking at the total state's regulations towards shale production, West Virginia is the only state that regulates all 20 elements, whereas Oklahoma regulates 16. States such as Colorado and Oklahoma have more quantitatively regulated elements than states such as California and West Virginia. West Virginia is the state with the most non-quantitatively regulated elements. Looking at specific regulations, the situation becomes more varied. West Virginia has cement-type regulations, but Oklahoma and Colorado do not, and California addresses the regulations in the drilling permit. All states require at least a permit to withdraw water over the threshold of 1,000 gal/day, but Pennsylvania require permits for any water withdrawals for hydraulic fracturing and operate ecosystem models that provide the basis for rejecting applications for water withdrawals that would put stress on ecosystems. Meanwhile, West Virginia requires a similar water management plan for withdrawals of more than 210,000 gallons per month. To withdraw more than 210,000 gallons per month, they need to document the source of the water withdrawal and show that its impact will be minimal. Lastly, we see differences in places such as Oklahoma and Colorado regarding how states can store fracturing fluids. Oklahoma requires sealed tanks for some fluids, but Colorado allows pits and regulates all fluids.

The differences in how heavily the state relies on the oil and gas industry for employment and income and in regulatory standards help contribute to the reason we see mixed results when analyzing individual states and the effects hydraulic fracturing has on the housing market. Colorado and Oklahoma seem to have the most in common of the five states as they both share a similar balance when it comes to the extraction of oil and natural gas and many state regulations. Both states show positive and significant point estimates for both the ZIP code fixed effect model and the spatial difference-in-differences. On average, four of the five states we examined showed positive and significant point estimates for either the ZIP code fixed effect model, the spatial difference-in-differences, or both. West Virginia had negative and significant point estimates for both the ZIP code fixed effect model and the spatial Difference-in-Differences. These estimates might be driven by the states' heavy reliance on coal which, from an industry perspective, has been declining over the past decade and is being replaced by cheaper natural gas.

6 Conclusion and Policy Implications

This paper uses data from over 15 million house transactions across 11 different states to demonstrate that houses within 2 miles of a hydraulically fractured natural gas well sell at higher prices compared to homes that are not. The price increase is largest in states with a more balanced treatment-to-control ratio and for homes within 0.5 miles of a well pad [17]. After controlling for both household-level fixed effects and repeat sales, we find homes within 2 miles of a well experience an increase of 4.9% in selling price. These estimates extend previous studies because they indicate how the broader housing market is affected, and they also depict the benefit to property net value of any lease payments and reflect the overall costs and benefits to homeowners.

These findings show that in relation to property value and fracking, while potentially controversial, the short-run negative externalities related to global pollution and any local disamenities are offset by the gains in local economic opportunities. This does not mean potentially adverse long-run impacts are not important, but it does help clarify the better socially relevant costs and benefits associated with fracking. Tax revenue generated an increase in economic activity, both directly through well permits and indirectly through higher property tax assessments. These tax revenues will also affect local public finances.

Finally, this paper looked across five different states to see how differences in state regulation on the oil and gas industries impacted the housing market. Four of the five states showed positive results. However, due to the complicated nature and the wide variance in state regulations, it is difficult to pinpoint precisely what regulations significantly hinder the oil and gas sector and what impacts those may have on the housing market. While we have done our best to uncover how the fracking boom has affected the housing market across states, the results of the fracking literature on property values are highly diverse, so these and other questions remain. Future research should continue to explore the impact of fracking on home values across different states and regions. In addition, given diverse state regulations and practices, there is also a need for further research to focus on the impact of fracking on environmental factors, such as local water use and contamination, among varying regional samples.

[17] To illustrate this increase, a house priced at $\$337,004.00$ would experience a 4.5% increase, adding $\$15,165.20$ to its value.

References

Allcott, H., and Keniston, D. (2018). "Dutch disease or agglomeration? the local economic effects of natural resource booms in modern america." *The Review of Economic Studies*, *85*(2), 695–731.

Apergis, N. (2019). "The impact of fracking activities on oklahoma's housing prices: A panel cointegration analysis." *Energy Policy*, *128*, 94–101.

Balthrop, A. T., and Hawley, Z. (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.

Bartik, A. W., Currie, J., Greenstone, M., and Knittel, C. R. (2019). "The local economic and welfare consequences of hydraulic fracturing." *American Economic Journal: Applied Economics*, *11*(4), 105–155.

Berndt, E. R., and Wood, D. O. (1975). "Technology, Prices, and the Derived Demand for Energy." *The Review of Economics and Statistics*, *57*(3), 259.

Black, K. J., Boslett, A. J., Hill, E. L., Ma, L., and McCoy, S. J. (2021). "Economic, environmental, and health impacts of the fracking boom." *Annual Review of Resource Economics*, *13*(1), 311–334.

Boslett, A., Guilfoos, T., and Lang, C. (2016). "Valuation of expectations: A hedonic study of shale gas development and New York's moratorium." *Journal of Environmental Economics and Management*, *77*, 14–30.

Boxall, P. C., Chan, W. H., and McMillan, M. L. (2005). "The impact of oil and natural gas facilities on rural residential property values: A spatial hedonic analysis." *Resource and Energy Economics*, *27*(3), 248–269.

Brown, J. P. (2015). "The response of employment to changes in oil and gas exploration and drilling." *Economic Review (01612387)*, *100*(2).

Brown, J. P. (2021). "Response of consumer debt to income shocks: The case of energy booms and busts." *Journal of Money, Credit and Banking*, *53*(7), 1629–1675.

Brown, J. P., et al. (2014). "Production of natural gas from shale in local economies: a resource blessing or curse." *Economic Review*, *99*(1), 119–147.

Cai, Z., Maguire, K., and Winters, J. V. (2019). "Who benefits from local oil and gas employment? labor market composition in the oil and gas industry in texas and the rest of the united states." *Energy Economics*, *84*, 104515.

Cheng, C., Mayer, W. J., and Mayer, Y. (2018). "The effect of legalizing retail marijuana on housing values: Evidence from Colorado." *Economic Inquiry*, *56*(3), 1585–1601.

Cheung, R., Wetherell, D., and Whitaker, S. (2018). "Induced earthquakes and housing markets: Evidence from oklahoma." *Regional Science and Urban Economics*, *69*, 153–166.

CHO, R. (2014). "The Fracking Facts."

Court, C. D., Hodges, A. W., Clouser, R. L., and Larkin, S. L. (2017). "Economic impacts of cancelled recreational trips to northwest florida after the deepwater horizon oil spill." *Regional Science Policy & Practice*, *9*(3), 143–164.

Darrah, T. H., Vengosh, A., Jackson, R. B., Warner, N. R., and Poreda, R. J. (2014). "Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales." *Proceedings of the National Academy of Sciences of the United States of America*, *111*(39), 14076–14081.

Delgado, M. S., Guilfoos, T., and Boslett, A. (2016). "The cost of unconventional gas extraction: A hedonic analysis." *Resource and Energy Economics*, *46*, 1–22.

Deziel, N. C., Brokovich, E., Grotto, I., Clark, C. J., Barnett-Itzhaki, Z., Broday, D., and Agay-Shay, K. (2020). "Unconventional oil and gas development and health outcomes: A scoping review of the epidemiological research." *Environmental research*, *182*, 109124.

Dronyk-Trosper, T. (2017). "Searching for Goldilocks: The Distance-Based Capitalization Effects of Local Public Services." *Real Estate Economics*, *45*(3), 650–678.

Fetzer, T. (2014). "Fracking Growth." Tech. rep.

Feyrer, J., Mansur, E., and Sacerdote, B. (2020). "Geographic dispersion of economic shocks: Evidence from the fracking revolution: Reply." *American Economic Review*, *110*(6), 1914–1920.

Feyrer, J., Mansur, E. T., and Sacerdote, B. (2017). "Geographic dispersion of economic shocks: Evidence from the fracking revolution." *American Economic Review*, *107*(4), 1313–1334.

Frohlich, C. (2012). "Two-year survey comparing earthquake activity and injection-well locations in the Barnett Shale, Texas." *Proceedings of the National Academy of Sciences of the United States of America*, *109*(35), 13934–13938.

Garcia-Gonzales, D. A., Shonkoff, S. B., Hays, J., and Jerrett, M. (2019). "Hazardous air pollutants associated with upstream oil and natural gas development: a critical synthesis of current peer-reviewed literature." *Annual review of public health*, *40*(1), 283–304.

Gittings, R. K., and Roach, T. (2020). "Who benefits from a resource boom? evidence from the marcellus and utica shale plays." *Energy Economics*, *87*, 104489.

Gold, R., and McGinty, T. (2013). "Energy boom puts wells in america's backyards." *Wall Street Journal*, *25*.

Gopalakrishnan, S., and Klaiber, H. A. (2014). "Is the shale energy boom a bust for nearby residents? evidence from housing values in pennsylvania." *American Journal of Agricultural Economics*, *96*(1), 43–66.

Green, D. A., Morissette, R., Sand, B. M., and Snoddy, I. (2019). "Economy-wide spillovers from booms: Long-distance commuting and the spread of wage effects." *Journal of Labor Economics*, *37*(S2), S643–S687.

Gregory, K. B., Vidic, R. D., and Dzombak, D. A. (2011). "Water management challenges associated with the production of shale gas by hydraulic fracturing." *Elements*, *7*(3), 181–186.

He, X., Lu, N., and Berrens, R. P. (2018). "The case of the missing negative externality? housing market effects of fracking in the niobrara shale play, colorado." *Journal of Environmental Economics and Policy*, *7*(3), 223–243.

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

Jackson, R. B., Vengosh, A., Carey, J. W., Davies, R. J., Darrah, T. H., O'sullivan, F., and Pétron, G. (2014). "The environmental costs and benefits of fracking." *Annual review of Environment and Resources*, *39*(1), 327–362.

Jacobsen, G. D., Parker, D. P., and Winikoff, J. B. (2023). "Are resource booms a blessing or a curse?: Evidence from people (not places)." *Journal of Human Resources*, *58*(2), 393–420.

Johnson, J. R., Johnson, G. R., McDonough, M. H., Burban, L. L., and Monear, J. K. (2008).

"Tree owner's manual for the northeastern and midwestern united states." Accessed: 2024-09-17.

Kaplan, T. (2014). "Citing health risks, cuomo bans fracking in new york state." *New York Times*, *17*.

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

Krupnick, A., Richardson, N., and Gottlieb, M. (2015). "Heterogeneity of state shale gas regulations." *Economics of Energy & Environmental Policy*, *4*(1), 53–68.

Lipscomb, C. A., Wang, Y., and Kilpatrick, S. J. (2012). "Unconventional Shale Gas Development and Real Estate Valuation Issues." *The Review of Regional Studies*, *42*(2).

Loomis, J., and Haefele, M. (2017). "Quantifying market and non-market benefits and costs of hydraulic fracturing in the united states: a summary of the literature." *Ecological Economics*, *138*, 160–167.

Maguire, K., and Winters, J. V. (2017). "Energy boom and gloom? local effects of oil and natural gas drilling on subjective well-being." *Growth and Change*, *48*(4), 590–610.

Maniloff, P., Mastromonaco, R., et al. (2014). "The local economic impacts of hydraulic fracturing and determinants of dutch disease." *Division of Economic and Business Working Paper Series, Colorado School of Mines*.

Marchand, J. (2012). "Local labor market impacts of energy boom-bust-boom in Western Canada." *Journal of Urban Economics*, *71*(1), 165–174.

Marchand, J., and Weber, J. (2018). "Local labor markets and natural resources: A synthesis of the literature." *Journal of Economic Surveys*, *32*(2), 469–490.

Mason, C. F., Muehlenbachs, L. A., and Olmstead, S. M. (2015). "The Economics of Shale Gas Development." *Annual Review of Resource Economics*, *7*(1), 269–289.

Mothorpe, C., and Wyman, D. (2021). "What the frack? the impact of seismic activity on residential property values." *Journal of Housing Research*, *30*(1), 34–58.

Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). "The Housing Market Impacts of Shale Gas Development." *American Economic Review*, *105*(12), 3633–3659.

Muehlenbachs, L., Spiller, E., Timmins, C., Thank, W., Bishop, K., Chu, J., Kousky, C., Krupnick, A., Lang, C., Linn, J., Ma, L., Mares, J., Mastromonaco, R., Staubli, S., Walsh, R., and We, J. W. (2012). "Shale Gas Development And Property Values: Differences Across Drinking Water Sources." Tech. rep.

Munasib, A., and Rickman, D. S. (2015). "Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis." *Regional Science and Urban Economics*, *50*, 1–17.

Ooms, T., and Tracewski, S. (2011). "Issue Brief The Impact on Housing in Appalchian Pennsylvania as a Result of Marcellus Shale-Social Services." *The Institute for Public Policy and Economic Development at Wilkes University*.

Paredes, D., Komarek, T., and Loveridge, S. (2015). "Income and employment effects of shale gas extraction windfalls: Evidence from the marcellus region." *Energy Economics*, *47*, 112–120.

Popkin, J. H., Duke, J. M., Borchers, A. M., and Ilvento, T. (2013). "Social costs from proximity to hydraulic fracturing in new york state." *Energy Policy*, *62*, 62–69.

Raimi, D., Bazilian, M., and Krupnick, A. (2020). "Who Wants Fracking in Their Backyard?"

Richardson, N., Gottlieb, M., Krupnick, A., and Wiseman, H. (2013). "The State of State Shale Gas Regulation." Tech. rep.

Rivera, N. M. (2020). "Is mining an environmental disamenity? evidence from resource extraction site openings." *Environmental and Resource Economics*, *75*(3), 485–528.

Stephens, H. M., and Weinstein, A. L. (2019). "Household valuation of energy development in amenity-rich regions." *Growth and Change*, *50*(4), 1375–1410.

Stern, N. (2008). "The economics of climate change." *American Economic Review*, *98*(2), 1–37.

Throupe, R., Simons, R., and Mao, X. (2013). "A review of hydro âĂIJfrackingâĂİ and its potential effects on real estate." *Journal of Real Estate Literature*, *21*(2), 205–232.

Timmins, C., and Vissing, A. (2015). "Valuing Leases for Shale Gas Development." *Working Paper*.

Torres, L., Yadav, O. P., and Khan, E. (2016). "A review on risk assessment techniques for hydraulic fracturing water and produced water management implemented in onshore unconventional oil and gas production." *Science of the Total Environment*, *539*, 478–493.

U.S. Energy Information Administration (2018). "Hydraulically fractured horizontal wells account for most new oil and natural gas wells - Today in Energy - U.S. Energy Information Administration (EIA)."

Vissing, A. (2015). "Private contracts as regulation: A study of private lease negotiations using the Texas natural gas industry." In *Agricultural and Resource Economics Review*, vol. 44, 120–137, Northeastern Agricultural and Resource Economics Association (NAREA).

Warner, B., and Shapiro, J. (2013). "Fractured, fragmented federalism: A study in fracking regulatory policy." *Publius: The Journal of Federalism*, *43*(3), 474–496.

Weber, J. G. (2012). "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming." *Energy Economics*, *34*(5), 1580–1588.

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

Weber, J. G., Burnett, J. W., and Xiarchos, I. M. (2014). "Shale Gas Development and Housing Values over a Decade: Evidence from the Barnett Shale." *SSRN Electronic Journal*.

Weber, J. G., Burnett, J. W., and Xiarchos, I. M. (2016). "Broadening benefits from natural resource extraction: housing values and taxation of natural gas wells as property." *Journal of Policy Analysis and Management*, *35*(3), 587–614.

Weinstein, A. (2014). "Local labor market restructuring in the shale boom." *Journal of Regional Analysis & Policy*, *44*(1), 71–92.

Whitacre, B. E., Johnston, D. L., Shideler, D. W., and Lansford, N. H. (2020). "The influence of oil and natural gas employment on local retail spending: evidence from oklahoma panel data." *The Annals of Regional Science*, *64*, 133–157.

White, N. E. (2012). "A tale of two shale plays." *Review of Regional Studies*, *42*(2), 107–119.

Willis, M. D., Hill, E. L., Kile, M. L., Carozza, S., and Hystad, P. (2022). "Associations between residential proximity to oil and gas extraction and hypertensive conditions during pregnancy: a difference-in-differences analysis in texas, 1996–2009." *International Journal of Epidemiology*, *51*(2), 525–536.

Wilson, R. (2022). "Moving to economic opportunity: the migration response to the fracking

boom." *Journal of Human Resources*, *57*(3), 918–955.
Winters, J. V., Cai, Z., Maguire, K., and Sengupta, S. (2021). "Causal effects of the fracking boom on long-term resident workers." *Journal of Regional Science*, *61*(2), 387–406.
Wrenn, D. H., Kelsey, T. W., and Jaenicke, E. C. (2015). "Resident vs. nonresident employment associated with marcellus shale development." *Agricultural and Resource Economics Review*, *44*(2), 1–19.
Zabel, J. E., and Guignet, D. (2012). "A hedonic analysis of the impact of LUST sites on house prices." *Resource and Energy Economics*, *34*(4), 549–564.

Appendices

Table 7: California Results

	Outcome variable: LnHP		
Independent Variables	(1)	(2)	(3)
Within 2 Miles	0.035***		
	(0.002)		
.5 Mile Zone	0.051***	0.013	0.010
	(0.010)	(0.010)	(0.011)
.5-1 Mile Zone	0.022***	-0.017***	
	(0.004)	(0.004)	
1-2 Mile Zone	0.038***		
	(0.002)		
Rooms	0.002***	0.003***	0.022***
	(0.000)	(0.001)	(0.003)
Bedrooms	0.156***	0.218***	0.246***
	(0.002)	(0.006)	(0.008)
Bathrooms	0.005***	0.010***	0.006***
	(0.000)	(0.001)	(0.001)
Ln Sqfeet	0.784***	0.855***	0.865***
	(0.001)	(0.003)	(0.007)
Year Built	-0.001***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)
Ln County Poverty Count	-0.002	0.002	0.029
	(0.002)	(0.008)	(0.019)
County Median House Income	0.000*	0.000**	0.000
	(0.000)	(0.000)	(0.000)
Observations	8,502,866	421,095	132,742
Sample	CA	CA Base Treat 3	CA Base Treat 2
R-squared	0.750	0.765	0.758
ZIP FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Robust S.E.	YES	YES	YES

Results are in the Logged value of the sales price (LnHP). Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2-mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 8: Colorado Results

	Outcome variable: LnHP		
Independent Variables	(1)	(2)	(3)
Within 2 Miles	0.013***		
	(0.002)		
.5 Mile Zone	0.066***	0.035***	0.011**
	(0.004)	(0.005)	(0.005)
.5-1 Mile Zone	0.040***	0.026***	
	(0.003)	(0.004)	
1-2 Miles Zone	0.018***		
	(0.003)		
Rooms	0.016***	-0.030***	-0.024***
	(0.000)	(0.002)	(0.003)
Bedrooms	0.091***	0.118***	0.092***
	(0.003)	(0.007)	(0.006)
Bathrooms	0.092***	0.072***	0.061***
	(0.002)	(0.005)	(0.007)
Ln Sqfeet	0.534***	0.278***	0.275***
	(0.002)	(0.005)	(0.005)
Year Built	0.000***	0.001***	0.002***
	(0.000)	(0.000)	(0.000)
Ln County Poverty Count	-0.005**	-0.011	-0.008
	(0.002)	(0.009)	(0.010)
County Median House Income	0.000*	-0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	1,668,992	122,901	94,305
Sample	CO	CO Base Treat 3	CO Base Treat 2
R-squared	0.621	0.456	0.427
ZIP FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Robust S.E.	YES	YES	YES

Results are in the Logged value of the sales price (LnHP). Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2 mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 9: Oklahoma Results

	Outcome variable: LnHP		
Independent Variables	(1)	(2)	(3)
Within 2 Miles	0.031***		
	(0.006)		
.5 Mile Zone	0.060***	0.040**	0.024
	(0.016)	(0.017)	(0.019)
.5-1 Mile Zone	0.070***	0.057***	
	(0.012)	(0.012)	
1-2 Mile Zone	0.018**		
	(0.007)		
Rooms	0.020***	0.021***	0.017***
	(0.001)	(0.002)	(0.003)
Bedrooms	0.160***	0.141***	0.132***
	(0.014)	(0.021)	(0.027)
Ln Sqfeet	0.826***	0.719***	0.698***
	(0.004)	(0.008)	(0.012)
Year Built	0.003***	0.004***	0.005***
	(0.000)	(0.000)	(0.000)
Ln County Poverty Count	-0.005	0.021***	0.025**
	(0.004)	(0.008)	(0.012)
County Median House Income	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Observations	553,519	133,522	52,323
Sample	OK	OK Base Treat 3	OK Base Treat 2
R-squared	0.401	0.482	0.508
ZIP FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Robust S.E.	YES	YES	YES

Results are in the Logged value of the sales price (LnHP). Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2-mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 10: Pennsylvania Results

	Outcome variable: LnHP		
Independent Variables	(1)	(2)	(3)
Within 2 Miles	0.015***		
	(0.003)		
.5 Mile Zone	0.011	0.028***	0.019**
	(0.008)	(0.009)	(0.009)
.5-1 Mile Zone	0.043***	0.016***	
	(0.005)	(0.005)	
1-2 Mile Zone	0.003		
	(0.003)		
Rooms	-0.002*	0.101***	0.114***
	(0.001)	(0.013)	(0.018)
Bedrooms	0.071***	0.148***	0.095***
	(0.003)	(0.028)	(0.030)
Bathrooms	0.100***	0.268***	0.240***
	(0.002)	(0.016)	(0.019)
Ln Sqfeet	0.613***	0.444***	0.440***
	(0.002)	(0.007)	(0.008)
Year Built	0.004***	0.006***	0.007***
	(0.000)	(0.000)	(0.000)
Ln County Poverty Count	-0.006*	-0.011	-0.006
	(0.004)	(0.013)	(0.016)
County Median House Income	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	2,183,202	171,198	109,586
Sample	PA	PA Base Treat 3	PA Base Treat 2
R-squared	0.645	0.453	0.427
ZIP FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Robust S.E.	YES	YES	YES

Results are in the Logged value of the sales price (LnHP). Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2 mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 11: West Virginia Results

	Outcome variable: LnHP		
Independent Variables	(1)	(2)	(3)
Within 2 Miles	-0.016**		
	(0.007)		
.5 Mile Zone	-0.049***	-0.041**	-0.028
	(0.018)	(0.018)	(0.018)
.5-1 Mile Zone	-0.033***	-0.026**	
	(0.012)	(0.012)	
1-2 Mile Zone	-0.006		
	(0.007)		
Rooms	0.012**	0.016***	0.018***
	(0.005)	(0.005)	(0.005)
Bedrooms	0.086***	0.096***	0.115***
	(0.013)	(0.015)	(0.017)
Bathrooms	0.290***	0.314***	0.308***
	(0.016)	(0.014)	(0.016)
Ln Sqfeet	0.558***	0.560***	0.565***
	(0.007)	(0.009)	(0.010)
Year Built	0.003***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Ln County Poverty Count	-0.001	0.001	-0.001
	(0.006)	(0.007)	(0.008)
County Median House Income	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	126,470	89,475	67,611
Sample	WV	WV Base Treat 3	WV Base Treat 2
R-squared	0.485	0.472	0.467
ZIP FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Robust S.E.	YES	YES	YES

Results are in the Logged value of the sales price (LnHP). Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2 mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 12: Effect on Unemployment Rate

	(1)	(2)	(3)	(4)
Within 2 miles	-0.055***			
	(0.018)			
.5 Mile Zone		-0.132**	-0.092**	-0.080**
		(0.052)	(0.038)	(0.032)
.5-1 Mile Zone		-0.115***	-0.076***	
		(0.027)	(0.025)	
1-2 Mile Zone		-0.033**		
		(0.015)		
Rooms	0.048	0.049	0.044	0.037
	(0.035)	(0.035)	(0.039)	(0.041)
Bedrooms	-0.091**	-0.094**	-0.115**	-0.117**
	(0.045)	(0.045)	(0.052)	(0.055)
bedrooms2	0.008	0.008	0.012	0.014
	(0.007)	(0.007)	(0.008)	(0.009)
Ln SqFeet	0.058	0.059	0.094*	0.111*
	(0.048)	(0.048)	(0.057)	(0.060)
Year Built	0.000	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Ln State Population	-6.455***	-6.456***	-4.915***	-4.588***
	(0.027)	(0.027)	(0.085)	(0.069)
Ln population	12.355***	12.361***	12.588***	13.680***
	(0.065)	(0.065)	(0.192)	(0.354)
Ln County Poverty Count	0.355***	0.355***	0.293***	0.276***
	(0.009)	(0.009)	(0.012)	(0.011)
County Median House Income	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
N	8308248	8308248	2069312	1847273
R-squared	0.915	0.915	0.839	0.758
ZIP FE	YES	YES	YES	
Year-Month FE	YES	YES	YES	
Robust S.E.	YES	YES	YES	

Results are in the Logged value of the unemployment rate. Column 1 is the fixed effects model results. Columns 2 and 3 are the spatial DID results. Column 2 the sample is restricted to just Base Treat 3 homes and it looks at the effect when a home is initially in the 2 mile zone and then moves to the .5-1 or 0-.5 mile zones because a new well was drilled and the distance to the nearest well decreased. Column 3 is the same but looking at homes that were once in the .5-1 mile zone and now are in the 0-.5 mile zone. The fixed effect model was run first for any home within 2 miles and those outside two miles. The model was then re-run for the three buffer zones, 0.5, 0.5-1, 1-2 miles.

Table 13: Summary Statistics by Buffer Zones

	.5 Mile Zone		.5-1 Mile Zone		1-2 Mile Zone		Above two miles	
	mean	sd	mean	sd	mean	sd	mean	sd
Log house price	12.08	0.89	11.90	0.90	11.94	0.90	12.37	0.88
Log house price per sqft	4.71	0.81	4.56	0.83	4.61	0.83	4.96	0.80
Rooms	3.55	3.33	3.94	3.29	3.86	3.31	4.06	3.48
Bedrooms	3.20	0.89	3.13	0.90	3.09	0.88	3.12	0.92
Sqfeet	1721.13	697.63	1680.60	716.41	1673.29	835.75	1835.49	3917.94
Year built	1983.51	31.43	1975.40	33.06	1973.16	31.02	1976.44	27.47

Table 14: Literature Data and Findings

Paper	Research question	Property Data	Data issues	Geographical universe	Point estimate
Norwood, El Fatmaoui, and Johnson	What effect do the locations of gas and oil wells have on housing market?	Zillow Transaction and Assessment Dataset (ZTRAX)	Small treatment to control ratio on Kansas, Kentucky, Wyoming, or Tennessee	California, Colorado, Oklahoma, Pennsylvania, and West Virginia	Positive 1.2 to 2.9 percent
Weber et al. (2014)	What effect does the shale gas development have on properties value?	Zillow Home Value Index (ZHVI)	Excluded property characteristics	The Barnett Shale in Texas	Positive 5 percent
Weber (2012)	What effect does the neutral boom have on employment and income?	Not applicable	Not applicable	Counties in Colorado, Texas, and Wyoming	Positive 1.5 percent
Apergis (2019)	What effect do fracking activities have on house prices?	Corelogic	No property characteristics	Oklahoma's counties	Positive 6.9 percent
Balthrop and Hawley (2017)	What effect does the New York State's moratorium on hydraulic fracturing have on property value?	The Multiple Listing Service and the Tarrant Appraisal District	Data on property characteristics is available only for one year, 2014	Tarrant County, TX	Negative 1.3 to 3.5 percent
Boslett et al. (2016)	What effect does the New York State's moratorium on hydraulic fracturing have on property value?	Counties' property assessment office and New York's Office of Real Property Tax Services.	Unable to acquire complete data for two counties (Bradford County in PA and Steuben County in NY)	Five counties along the border of NY and PN	Negative 23 percent
Muehlenbachs et al. (2015)	What effect does the shale gas development have on groundwater-dependent properties?	CoreLogic	Over 33 percent of observations missing property characteristics (1).	36 counties in Pennsylvania state	Negative 6.5.0 to 13 percent
Boxall et al. (2005)	What effect does proximity to oil and gas (sour gas in particular) wells have on rural property value?	The Multiple Listing Service (MLS) records of the Calgary Real Estate Board	A small sample of 532 property sales	The City of Calgary in southern Alberta, Canada.	Negative 4 to 8 percent, or a higher magnitude
Mothorpe and Wyman (2021)	What effect does earthquake induced by fracking have on house prices?	Oklahoma County Tax Assessor's database	The data size does not allow for repeat sale consideration	Oklahoma County in Central Oklahoma	Negative 2 to 19.2 percent
Gopalakrishnan and Klaiber (2014)	What effect does the boom in gas exploration have on well water dependent, rural, or highway neighboring properties?	Dataquick	482 non-geocoded observations	Washington County, Pennsylvania	Negative 5.5 to 21.6 percent

Table 15: Data Cleaning

Category	Choice
Identify transaction prices reflecting fair market value	Select document types (if yes, include names / codes)
	Select price types (e.g., remove prices flagged as estimated or not full value)
	Select loan types (e.g., remove construction and home equity loans)
	Remove transactions with intra-family sale flags
	Remove transactions with high similarity of owner and buyer names
	Remove transactions involving public sellers or buyers
	Remove transactions of partial property interests (e.g. easements)
	Remove prices below a given threshold (if yes, include threshold)
	Remove prices above a given threshold (if yes, include threshold)
Measure spatial relationship between transacted properties and attribute of interest	Type of spatial location data used (e.g. property coordinates, block group ID)
	Verify geodetic datum of coordinates
	Remove transactions with duplicate coordinates for non-duplicate properties
	Crop coordinates to county and ZIP code boundaries
	Geocode addresses
	Link transactions to property boundary polygons (if yes, describe method)
	Link transactions to geo-located building footprints
	Check for biases from dropping transactions without geo-location data
Link transactions to time-varying property characteristics	Remove transactions if new structure was built after sale
	Remove transactions if structure was remodeled after sale
	Check for biases due to missing data on building and remodeling year
	Constrain the time horizon of the analysis to most recent years
	Use tax assessor data from multiple time periods
Identify property types	Property types included in analysis:
	Identify property types using land use codes in tax assessor data (if yes, include codes or: "codes not reported")
	Check consistency of land use code usage across study region
	Identify property types using land use codes in transaction data
	Verify building presence or absence with data on building valuation
	Verify building presence or absence with data on building characteristics
	Verify building presence or absence with data on building footprints
	Use remote sensing data to identify property types
Account for missing and miss-measured data for standard housing attributes	Housing attributes from tax assessor data included in models:
	Drop transactions with missing data
	Check for potential biases from dropping transactions with missing data
	Use dummy variables to control for missing data
	Impute missing data
	Use neighborhood fixed effects (if yes, name unit, e.g. county, zip code)
	Use unit fixed effects (repeat-sales)
	Remove transactions with implausible values in housing attributes (if yes, please list)
Check property sales history	Remove transactions of the same property within X days after a retained transaction (if yes, report X)
	Remove transactions of the same property at the same price within X days after a retained transaction (if yes, report X)
	Remove transactions of properties that have sold more than X times (if yes, report X)
	Remove transactions with large value increases (e.g., x00% within x year(s) for 1ÂšxÂš10)

Zillow Housing Data

Considering the size and scope of the Zillow ZTRAX repository, it is necessary to document the data cleaning process used for this research. However, in order to create a dataset that is both national and representative, some adjustments were made to the import process. In general, the effort follows Zillow's own script which creates a hedonic dataset.[18] The end product results in a dataframe in which each row is a home transaction and each column reflects home and transaction characteristics. The files are initially imported state-by-state and then appended together to make a master file.

The process goes as follows. First, three tables are imported from the Assessment repository: Main, Building, and Building Areas. These three tables combine to provide house characteristics, as well as information about the type of property exchanged in a given transaction. For example, the variable "PropertyLandUseStndCode" in the Building table details whether a property is a single-family residence, used in industry, is a farm, et cetera. We erred on the side of inclusively when filtering for these variables during import, as reporting standards across counties and states vary widely. The properties included are described as follows in Zillow's documentation:

1. Residential General
2. Single-Family Residences
3. Rural Residences
4. Mobile Home
5. Townhouse
6. Cluster Home
7. Condominium
8. Cooperative
9. Row House
10. Planned Unit Development
11. Residential Common Area
12. Seasonal, Cabin, Vacation Residence
13. Bungalow
14. Zero Lot Line
15. Manufactured, Modular, Prefabricated Homes

[18] The original file is publicly available on the firm's ZTRAX GitHub repository: `https://github.com/zillow-research/ztrax/blob/master/ExampleRcode_UsingZTRAXtoCreateHedonicDataset.R`

16. Patio Home

17. Garden Home

18. Landominium

19. Inferred Single-Family Residential

Also, following the logic described by Zillow, we filter the "BuildingAreaStndCode" from the BuildingAreas table in order to get as accurate a measure of total square footage as possible. Again, different counties have different reporting standards as to what is included in their square footage calculations, so to ensure consistency we have included only those options which enumerate the buildings on the property, not the land itself. These two filters – for "PropertyLandUseStndCode" and "BuildingAreaStndCode" – are the only two at this point in the process. Once this is complete, the three assessment tables are merged to create a single assessment file with all the necessary housing characteristic variables to be used in analysis.

The second set of data comes from the Transaction repository. Included are the PropertyInfo and Main tables. All the information provided here reflects the transaction itself, not any characteristics of the home. This includes variables like the price of the transaction, the date of transfer, and the type of transfer. The only filtering that occurs in this step is in regard to the variable "DataClassStndCode," which details the type of transaction occurring. Since the subject of study are property transactions, only deed transfers and deed transfers with concurrent mortgages are included. This excludes other types of transactions, including foreclosures and inter-family transfers as in the case of inheritances. These two tables are appended together to make a single transaction file. Finally, the transaction and assessment files are combined to make a single master file for a given state. The states files are then appended together to make a national-level dataset which is then used for analysis.

The master file is filtered to exclude extreme observations, as well as define the period of study. To ensure that results are not being driven but incorrect or implausible observations, we drop transactions which had sales prices below \$10,000 and above \$10,000,000, similar to Cheng et al (2018). On the lower end it is unlikely that transactions with prices below \$10,000 occurred on the market, and may have slipped through the "DataClassStndCode" filter. Prices above \$10,000,000 are extraordinary and in some cases are likely the result of data entry errors. Similarly, house characteristics are filtered to exclude observations that are in the top thousandth or top ten-thousandth percentile. Doing so, for example, eliminated an observation with over 1000 bedrooms. This process removed a large number of observations in states which do not require counties to report the home characteristics, leaving small states like Maine with just 11,000 transaction observations. To guarantee a representative sample, we then dropped states which did not have at least 100,000 observations. That is an arbitrary standard, but by doing so we can more confidently argue that each states' market is properly represented. Finally, prices were adjusted to reflect 2018 prices using the Federal Reserve's Consumer Price Index.