The Long-Term Consequences of Oil and Gas Extraction: Evidence from the Housing Market

Mariya Shappo*

Job Market Paper
Latest version available here

November 1, 2020

Abstract

Pennsylvania has a long history of oil and gas extraction. When the production period ends, many wells are left behind and some remain unplugged. Unplugged wells impose serious environmental costs, including explosion hazard and risk of water, atmosphere, and soil contamination. This paper quantifies these costs by estimating the effect of unplugged wells on housing prices. I use rich data on oil and gas drilling, housing market transactions, and lease agreements to conduct difference-in-differences and instrumental variable analyses. I show that well abandonment reduces house prices. Old wells, left behind by oil and gas operators, affect house prices more than active, producing ones. However, this depreciation is reversible: if well site clean-up is completed, house prices recover almost entirely. I show that the benefits of proper well plugging are larger than the costs. This motivates environmental policies aimed at creating incentives for oil and gas producers to plug wells. These policies may include higher bankruptcy insurance requirements or environmental taxes.

JEL Codes: L71, Q35, Q51, R31,
Keywords: Housing, Housing Prices, Hedonic, Environmental Health and Safety, Cost Benefit, Drilling, Fossil Fuels

---

*I would like to thank David Albouy, Alex Bartik, Dan Bernhardt, and Dan McMillen for the support and guidance. I also would like to thank Steven Puller, Peter Maniloff, Dana Shaat, Nazanin Khazra, APPAM 2019 Fall Conference participants, and University of Illinois Graduate Seminar and Applied Micro Research Lunch participants for valuable comments. Please contact the author by email at shappo2@illinois.edu. Any mistakes are my own.

1 Introduction

The literature has been emphasizing the benefits brought by oil and gas extraction to local economies, such as stronger labor markets and increased tax revenues. Feyrer et al. (2017) show that the fracking revolution produces large increases in local wages, royalty payments, and employment. Maniloff and Mastromonaco (2017) attribute 550,000 new local jobs to fracking. Bartik et al. (2019) find fracking-induced improvements in various economic indicators but deterioration in local amenities. Those short- and medium-run benefits are dispersed among many residents, and they are relatively easy to quantify. However, oil and gas extraction could be associated with substantial long-term environmental and public health costs, including explosion hazard and risk of water, atmosphere, and soil contamination. The major source of such risks are unplugged abandoned wells, left behind by oil and gas operators. Such wells are not producing anymore and are not expected to be used in the future. Well plugging — proper closing to prevent leaking — and site clean-up are beneficial because they substantially reduce these environmental costs (Kang et al., 2017). Because there is no explicit market for the environmental risks and site restoration, it is difficult to quantify these costs and benefits.

In this paper, I turn to the housing market to quantify (i) the long-term environmental costs of oil and gas extraction and (ii) the benefits of clean-up, using rich data on oil and gas drilling activity, house sales, and lease agreements in Pennsylvania. Naturally, abandoned wells would have the largest impact on the households that reside in close proximity to them. In the framework suggested by Rosen (1979) and Roback (1982), willingness-to-pay for residing further from potentially hazardous sites is reflected in house prices. The negative amenity, created by abandoned wells, may decrease home values.[1] There are other environmental consequences of abandoned, unplugged wells that are not captured by the housing market, such as adverse climate change effects. Therefore, the housing market reflects the lower bound *perceived* costs.

There are two major challenges in estimating the impact of abandoned wells. First, the location of wells across space could be non-random, and well drilling could be correlated with unobserved neighborhood or house characteristics. This challenge, discussed in Muehlenbachs et al. (2015), emphasizes the importance of using the data on houses sold more than once. In this paper, I use the repeated sales sample, which allows me to control for house and neighborhood time-invariant characteristics that could potentially correlate with well locations.

---

[1] Glaeser and Gyourko (2005) argue that because housing is durable, negative shocks decrease house prices more than population.

Second, both decisions to abandon and plug a well could be correlated with unobservable time-varying house or neighborhood characteristics. This challenge is common in the literature studying the impact of hazardous waste clean-up (Greenstone and Gallagher, 2008, Currie et al., 2011, Gamper-Rabindran and Timmins, 2013): the choice of which site to clean up and which site to delay cleaning up can be non-random. In the context of the abandoned wells problem, there are multiple possibilities of selection. For example, the producers could be more likely to plug wells in wealthier areas where homeowners have more bargaining power, more information, or better means to process this information. Similarly, the Department of Environmental Protection (DEP) may plug more wells in areas with better public service provisions in general.

To address these challenges, I construct an instrumental variable (IV) for well abandonment and plugging, plausibly exogenous to the aforementioned local conditions. This variable is similar in spirit to shift-share instruments. It isolates the changes in the macroeconomic environment driven by gas and oil prices, regulation, and technology. Because the industry is very volatile, I show that these sources provide sufficient variation likely uncorrelated with local economic conditions. I control for royalties and perform additional checks to argue that this instrument is unlikely to affect house prices through other channels.

I find that if wells are left unplugged, the long-term consequences of drilling are negative and substantial. Unplugged wells decrease sale prices of nearby houses by 4%–24%, depending on the number of wells and their proximity. An average treated house located within 2 km of unplugged wells depreciates by 11%, or approximately $15,000. This is a much larger loss in comparison to depreciation caused by new, active drilling.[2] However, when abandoned wells are plugged, house prices restore almost completely. The long-term impact on home values is small and marginally significant if wells are plugged and site restoration is completed. The estimates also imply that the benefits of plugging are larger than the costs.

This paper focuses on older, abandoned wells and sheds light on the long-term housing market impact of oil and gas extraction—not only during the boom but also during the bust. These long-term effects are important, because more wells in the future will lose economic viability and therefore will be left behind. The current literature, however, focuses on the fracking boom. The impact of *new* oil and gas extraction on property values (Boxall et al., 2005, Gopalakrishnan and Klaiber, 2013, Muehlenbachs et al., 2015, Delgado et al., 2016) and mortgage default risk (Cunningham et al., 2020) has been extensively studied. However, the evidence from the boom only captures the short- and medium-run impact of drilling due

---

[2] My analysis shows that the impact of active wells is small. Muehlenbachs et al. (2015) show that groundwater-dependent properties within 1–1.5 km of active wells lose, on average, around $30,167, while piped-water properties gain $4,802.

to its focus on the active production phase. This impact includes both positive effects (higher housing demand and capitalization of royalty payments) and negative effects (environmental and health risks). Little attention is paid to the long-term impact of oil and gas drilling, to when extraction is completed, and to wells left behind. Studies that perform cost-benefit analyses typically ignore the environmental impact of wells after abandoning. Moreover, only direct costs are usually measured due to the difficulty in quantifying the indirect and long-term effects (Considine et al., 2015, Waxman et al., 2020).

In Pennsylvania alone, it is estimated that as many as 300,000 to 760,000 wells have been drilled since 1859 (Pennsylvania Department of Environmental Protection, 2014). A lot of them were drilled prior to modern environmental regulations, and it is hard to find anyone responsible for maintaining or properly closing those wells.[3] According to current laws, all non-producing gas and oil wells must be plugged and well site reclamation must be completed by operators. However, some operators still abandon wells without going through the expensive process of plugging. This is especially more likely when gas and oil prices are low, because it constrains operators' resources. Moreover, low prices push some companies into bankruptcy, in which case all wells operated by the such companies are at risk of being left abandoned without plugging.

This paper emphasizes that oil and gas extraction can be associated with long-term social costs, even after active production is completed. Economists have been interested in how to regulate industries that are associated with such costs. Policy measures, such as insurance requirements for gas and oil producers, may improve environmental outcomes and lead to partial internalization of the costs (Boomhower, 2019). In Pennsylvania, bonding requirements are intended to work as insurance against environmental non-compliance. If a firm cannot plug its wells (e.g., when declaring bankruptcy), the bond is used to perform plugging and clean-up. However, until recently, bond amounts were insufficiently low and not nearly enough to cover plugging costs (Mitchell and Casman, 2011).[4] The Pennsylvania Act 13 of 2012 was adopted to address some of these concerns and to impose stricter environmental requirements, including much higher bond amounts, for producers that use fracking. Moreover, Pennsylvania, as well as many other states, adopted plugging programs for wells without a solvent owner. This paper informs such policies, showing that some environmental costs are paid by households rather than by oil and gas producers. The paper motivates the implementation of policies that create incentives for firms to plug wells, such as higher bonding and penalties for non-compliance with the environmental requirements.

---

[3] Modern plugging requirements were adopted in 1984.

[4] Many other states have adopted bonding requirements. The bond amounts vary, but they are rarely sufficient to cover the reclamation costs.

The estimates also help to identify the group that disproportionately bears the costs in case the producers cannot comply with the regulations.

The rest of the paper is organized as follows. Section 2 provides background on oil and gas drilling in Pennsylvania and discusses the current policies. Section 3 describes and summarizes the data on wells, house sales, and lease agreements. Section 4 discusses the methodology of this study and the sources of variation used. The results are discussed in Section 5, and Section 6 shows how the alternative specifications compare with the baseline. Section 7 argues that the benefits of well plugging, implied by the estimates, are larger than the costs. Section 8 concludes.

2 Background

2.1 Environmental risks of abandoned wells

Abandoned gas and oil wells impose serious environmental and public health risks, both globally and locally. In contrast to active wells, unused sites are not under systematic monitoring and control. This imposes additional risks: if a problem arises on a well site, it might be impossible to fix it promptly.

Abandoned wells impose risks of groundwater and soil pollution with oil, gas, saltwater, and other contaminants. A significant number of water contamination cases were directly linked to abandoned wells; those cases are called ”legacy issues.”[5] Abandoned wells can also leak hydrogen sulfide—a toxic gas with a ”rotten egg” smell. Finally, abandoned well sites can have old equipment, small spills, and other waste. This is not only hazardous but also reduces the local aesthetic value.

According to the recent estimates, abandoned wells are responsible for 5%–8% of total annual anthropogenic methane emissions in Pennsylvania (Kang et al., 2017).[6] Other studies, based on data from Wyoming, Colorado, Utah, and Ohio, found smaller emissions relative to the regional level.[7] While non-toxic when ingested, this gas is explosive when mixed with air. Leaks from abandoned wells are one of the major sources of stray gas that migrates to aquifers or the zone between the groundwater and the surface. If enough of this gas moves into a building (e.g., through a water well, especially if that well is located inside an enclosed structure), it creates the risk of explosion (Ground Water Protection Council, 2011). Methane is also a greenhouse gas with a global warming potential much higher than

---

[5] In Texas and Ohio, 22% of oil-field-related water contamination cases were attributed to abandoned wells (Ground Water Protection Council, 2011).

[6] This corresponds to 0.04–0.07 Mt (10^9 kg) of methane per year.

[7] See Townsend-Small et al. (2016). The rough national estimate is 1.6×10^4 kg per hour.

carbon dioxide.[8] These climate change risks of abandoned wells are very important but will not be captured by a hedonic house price model. Moreover, house sales reflect the *perceived* risks. For that reason, the estimates based on the housing market data should be interpreted as the lower bound of the perceived environmental costs.[9]

While it is hard to find direct evidence, it is likely that local residents are aware of the issues brought by drilling and by abandoned wells specifically. First, this information should be revealed to home buyers. Real estate disclosure laws in Pennsylvania require to reveal information about any potentially hazardous substances on the property before a sale. Loans, insured by the Federal Housing Administration, specifically require inspection and disclosure of any abandoned wells nearby. Second, the issue of potential hazards associated with unplugged wells drew a lot of attention in the media, increasing public awareness; they covered several explosions caused by improperly maintained and abandoned wells.[10] Some of those explosions led to the property damage, deaths, and evacuation of nearby residents.

2.2 Regulation and incentives

The solution to the abandoned wells problem is their proper plugging. It has substantial environmental benefits, reducing the probability of leakage to soil, atmosphere, and water. The process involves placing cement plugs in the boreholes and testing to prevent migration of fluids. It also involves equipment removal and surface reclamation, reducing the associated hazard and improving the aesthetic value of the area.

Ideally, a mineral rights owner leases to an operator who drills, extracts, and then properly closes wells. However, many abandoned wells remain unplugged. According to the Environmental Protection Agency (2020) estimates, there are 2.1 million unplugged abandoned oil and gas wells across the United States. In Pennsylvania alone, where drilling started in 1895, between 100,000 and 560,000 wells are abandoned. Many of those wells were drilled before modern environmental regulations and reporting requirements. As a result, these wells are not accounted for in records and their location and status is unknown. Many of those "legacy wells" are unplugged or improperly plugged. Because there are no well-defined property rights, Coasian solutions do not apply. As a response, regulators in many states, including the DEP in Pennsylvania, adopted abandoned well programs aimed

---

[8] However, methane has a short residence time in atmosphere. Policies aimed at methane reduction will reduce global warming in the short term (Hoegh-Guldberg et al., 2018).

[9] Not all wells are equally dangerous. Relatively few wells emit the majority of methane. While most abandoned wells in Pennsylvania (90%, according to Kang et al. (2017)) emit some methane, unplugged and plugged vented wells are the major sources of emissions.

[10] For example, the explosion in Firestone, Colorado in 2017; in Trinidad, Colorado in 2007; and in Bradford, Pennsylvania in 2011.

at locating and plugging if no responsible party can be identified. Because only a relatively small number of legacy wells were discovered, many of them are unaccounted for in this paper.[11]

In Pennsylvania, modern environmental and reporting requirements were established in the Oil and Gas Act of 1984. According to the current law, well operators are required to properly plug all wells upon abandoning.[12] They also have to restore land surface that has been disturbed during drilling and production. If a well is abandoned and still unplugged, the DEP can enter the site, plug the well, and sell all the equipment left behind.

Current policies also require oil and gas producers to post a bond when applying for a well drilling permit. Bonding is intended to play the role of insurance against non-compliance with environmental requirements. It can be used to cover the cost of plugging and site restoration in case an operator cannot do it—for example, if they filed for bankruptcy. At the same, bonds create incentives for operators to plug wells promptly after they become economically non-viable, because the bond amount is returned only after all environmental requirements are satisfied. Bonding requirements, first introduced in 1984, were $2,500 per well or $25,000 blanket bond covering all the wells operated. This requirement was the same for all wells, regardless of their characteristics. Moreover, no actual cash transfer was required. Instead, owners had the choice to use certificates of deposit or automatically renewable letters of credit from financial institutions.

The size of bond is small and covers only a small fraction of reclamation costs (Mitchell and Casman, 2011). For example, a large producer, Alliance Petroleum, owns 13,228 conventional wells in Pennsylvania.[13] If covered by the blanket bond, the amount is just $1.90 per well. Restoration of sites after oil and gas extraction is very costly and substantially varies by the type and depth of wells. For a single conventional well in the Marcellus shale, the cost is estimated to be around $24,000, and in some cases it may be substantially higher. The blanket bond would therefore cover the cost of plugging only for one well out of more than 13,000.

Because the bond amounts are low, there are insufficient incentives for firms to plug abandoned wells. If a company declares bankruptcy before plugging their wells, bond amount would not be able to cover the full costs. Keeping in mind the large number of legacy wells, plugging is a significant financial liability for the DEP.

---

[11] This paper uses the variation in well status for known wells. As described in Section 4.2, all specifications include house fixed effects and therefore use price variation within the same property. Unknown wells do not change status and therefore will be absorbed by the house fixed effects.

[12] This occurs unless the special temporary inactive status has been granted. However, few wells that were granted this status continue producing.

[13] Conventional drilling uses traditional technology. Wells are typically vertically drilled and are less deep than unconventional ones.

The preliminary data show that COVID-19 increased the number of non-producing wells and pushed many companies into bankruptcy. According to Haynes and Boone, LLP (2020), 13 companies filed for bankruptcy in July–August 2020, which is a 62% increase in comparison to the same period last year. More than 244 producers filed for bankruptcy since 2015, with the total debt nearly $172 billion. However, several states, such as North Dakota, allocated the coronavirus relief funds to plug abandoned wells, which supports employment in the industry (Pew Charitable Trusts, 2020).[14]

In 2012, stricter environmental requirements were adopted for unconventional wells.[15] Unconventional wells are stimulated by hydraulic fracturing and are typically deeper and often horizontally drilled.[16] The measures included increased civil penalty limits for environmental violations and increased impact fees and bonding requirements. Bonding requirements substantially increased for all unconventional well operators. They also became higher for deeper wells, as the cost of plugging directly depends on well depth (see Figure A.4). However, the bond amount remains much smaller than plugging costs: the maximum requirement is $10,000 per well for producers operating less than 25 deep unconventional wells.[17] The cost of plugging for these type of wells is, on average, $100,000.[18]

In addition to the change in the bonding requirements, the new legislation imposes an annual impact fee on unconventional wells. This fee depends on both oil and gas prices and when a well was frilled, and it varies from $60,000 per well for new wells with high gas prices to $5,000 for older wells with low gas prices (see Figure A.3). Importantly, non-producing (i.e., inactive or plugged) wells and marginally producing (so-called stripper[19]) wells are exempt from the fee. This means that this fee should not affect the decision to plug low-producing older wells, because producers could still have incentives to maintain low levels of production instead of bearing large reclamation costs. The majority of the revenues from the fees is transferred to the Public Utility Commission. A substantial portion of those revenues is distributed to county conservation districts, and the rest is distributed

---

[14] Raimi et al. (2020) argue that a federal program to plug abandoned wells would create tens of thousands of jobs.

[15] See the Pennsylvania Act 13 of 2012. Pennsylvania Act 87 of 2012 amends Act 13 of 2012 and clarifies that bonding requirements for conventional wells stay at the previous level, while the increase was intended for unconventional wells only.

[16] Horizontally drilled wells typically extract gas cover area around one square mile (2.6 square kilometers), depending on geology, lease agreements, and other factors.

[17] Bond amounts are low in other states as well. North Dakota and Wyoming allow for a $100,000 blanket bond, and Texas sets the maximum at $250,000. See Raimi et al. (2020) for a review.

[18] In some cases, the cost can be substantially higher. Cabot Oil reported that they spent $2,190,000 to plug three wells (Cabot Oil Gas Corporation, 2010).

[19] The law defines a stripper well as being ”incapable of producing more than 90,000 cubic feet of gas per day during any calendar month.”

among other agencies.[20] Finally, the new legislation changed the civil penalty limits for unconventional wells operators from \$25,000 plus \$1,000 per day to \$75,000 plus \$5,000 per day.

Kim and Oliver (2017) show that stricter environmental requirements in Pennsylvania lead to fewer permits acquired by firms and higher environmental protection efforts. Black et al. (2017) find that the introduction of an impact fee in Pennsylvania leads to a large decline in leasing but no change in well permitting or drilling. Boomhower (2019) has shown that introduction of bond requirements in Texas coincides with a decrease in the share of operators that left behind orphaned wells from 7% to 3% and a clear decrease in total number of abandoned wells.

3 Data, definitions, and descriptive statistics

3.1 Well data

Data on oil and gas wells are provided by the Pennsylvania DEP. I combine information from several data sets. Spud data provide the spud date of each well (i.e., when the drilling was started), its geographical coordinates, its current status,[21] and the description. Active wells have an issued permit, but the well may or may not be producing or fully drilled. A well is abandoned if it "has not been used to produce, extract, or inject any gas, petroleum, or other liquid within the preceding 12 months; or for which equipment necessary for production, extraction, or injection has been removed; or considered dry and not equipped for production" (Pennsylvania Department of Environmental Protection (2014)). Abandoned wells are expected to be closed permanently,[22] while non-producing wells that are expected to be used in the future have active or temporarily inactive status.[23] Abandoned wells can be plugged by either operators or the DEP. If an abandoned well has been inspected by the DEP, it appears on the Abandoned, Orphan, and DEP Plugged Listing. It also provides the inspection date and plug date for the wells plugged by the DEP, including many old, legacy

---

[20] Pennsylvania Fish and Boat Commission, Pennsylvania Department of Transportation, and Pennsylvania Emergency Management Agency.

[21] The status can be active, abandoned, DEP abandoned list, DEP orphan list, DEP plugged, plugged OG well, regulatory inactive status, operator reported not drilled, and proposed but never materialized. The last two statuses are not of an interest: if the well has not been drilled, it should not affect housing prices through a worsening of environmental conditions. There are very few of these observations, and I exclude them from the sample.

[22] In approximately 20% of cases, an abandoned well starts producing again. This can happen if a well becomes viable because of the technological development or if the ownership was transferred to another operator. In this case, its status changes from abandoned to active.

[23] Wells abandoned prior to April 18, 1985 and generate no economic benefit are called orphaned.

wells. The well description also provides information on its purpose (extraction, injection, storage, etc.), resource extracted (oil, gas, coalbed methane, or combined oil and gas), and configuration (conventional or unconventional).

Production reports are issued monthly, annually, and semi-annually, depending on the type of well and the reporting period. They are available starting from 1980. These reports allow me to track the production history of each well and to determine the periods when it was not producing and when it was abandoned.[24]

DEP data capture only a little over 204,000 wells. According to Pennsylvania Department of Environmental Protection (2014), there are 300,000–760,000 wells drilled in the state, while other sources provide even larger estimates.[25] Section 4.2 explains that this is not an issue for the identification in this paper.

By configuration, only 6.2% of the known wells are unconventional, fracked wells, and the rest are traditional, conventional ones. By resource extracted, 43.9% are gas, 42.3% are oil, and 5.2% combine gas and oil. Coalbed methane and undetermined wells are a small fraction of the sample. In this paper, I define the well type as the combination of configuration and resource, for example, conventional oil and unconventional gas. The majority of known wells are still active. About 30% of all wells were abandoned, 12.1% were left unplugged at some point, and 20.6% of wells are plugged.

Figure 1 shows that drilling is concentrated in the northeastern part of the state, and therefore I focus on only 35 counties in the main analysis.[26] Figure 2 shows that in those 35 counties, the number of active wells has been increasing rapidly since 1990, while the number of unplugged abandoned and plugged wells have been increasing at a slower rate. Well status is correlated with well age. As shown in Figure 3, older wells are more likely to be abandoned. Almost 30% of wells over 35 years old are left unplugged, while only about 5% of newer wells have this status.

3.2 House sales data

I use housing market data from CoreLogic—a private financial and property data provider—that allows me to observe individual transactions. The data include extensive information on property characteristics, such as the exact location, age, number of rooms, square footage,

---

[24] There is also information on the quantity produced by each well over the reporting period. I do not differentiate wells by the production levels.

[25] Kang et al. (2017) estimates there are between 470,000 and 750,000 abandoned gas and oil wells in the state.

[26] The counties are Allegheny, Armstrong, Beaver, Bedford, Blair, Bradford, Butler, Cambria, Cameron, Centre, Clarion, Clearfield, Clinton, Crawford, Elk, Erie, Fayette, Forest, Greene, Indiana, Jefferson, Lawrence, Lycoming, McKean, Mercer, Potter, Somerset, Sullivan, Susquehanna, Tioga, Venango, Warren, Washington, Westmoreland, and Wyoming.

and lot size. These data do not have any demographic or other information about buyers or sellers.

Table 1 shows the summary statistics of the house sales data. The data include over 450,000 housing market transactions that occurred between 1980 and 2017. Because the main analysis is performed on the sample of properties sold more than once (see Section 4.2 for details), the summary statistics uses only repeated sales. In the table, treated houses are within 2 km of any well and are compared with the pure control group, properties 2–20 km of wells. The first row shows that the houses in the control group are slightly more expensive than the treated houses. This is driven by Allegheny County—an urban county where Pittsburgh is located—which accounts for 37% of all observations. Pittsburgh has fewer wells and more expensive houses than the rest of the sample.[27]

Table 2 shows the relative size of the pure control and the treatment groups. Sixty-one percent of properties have at least one well within 2 km. The identifying variation comes from new wells or existing wells changing status, shown in the fourth column. In the sample period between 1980 and 2017, at least one new well was drilled next to 37% of houses, and there are, on average, 15 wells within 2 km of a house, conditional on treatment. Twenty-seven percent of houses have at least one unplugged well within 2 km. The number of unplugged wells changes near 12% of properties. For the treated houses, the average number of this type of well is 2.8. Finally, a little over a half of the properties have at least one plugged well nearby, and 24% of houses have at least one new well plugged during the sample period. This number shows that the size of the treatment groups is sufficient for all treatment types, that is, for drilling, abandoning, and plugging. The distribution of wells across houses is right-skewed, with few outliers having a very large number of nearby wells. As a result, I restrict the sample to the houses that have no more than 100 wells nearby, dropping about 2% of the observations.

3.3 Lease data

If a mineral rights owner (often a landowner) grants the right to develop resources to an operator (i.e., a firm that extracts oil and gas), the terms are set out in a lease agreement. The mineral rights owner's income is called royalties, which is legally required to be at least 12.5% of revenue from selling oil and gas. This royalty percentage is often higher than the legal minimum. In Pennsylvania, natural gas royalty payments in 2017 were $1.15 billion (Independent Fiscal Office, 2020). These payments are very volatile, because they follow the

---

[27] If Allegheny County is excluded, the difference in the pre-treatment mean price is a little over $2,000 and is insignificant. Section 6.3 performs robustness checks and shows that the difference in pre-treatment prices for the full sample is not an issue.

trend of natural gas prices. However, they are a substantial source of income for mineral right owners.

Oil and gas lease data are provided by Enverus, formerly Drillinginfo, a private provider. Drillinginfo provides lease data from 23 Pennsylvania counties.[28] These data contain the name and address of a lease grantor—mineral resource owner—as well as lease terms and royalty percentage. To connect these data to the housing data, I geocode addresses using Geocodio, a commercial service. Doing this provides geographical coordinates of a lease grantor's residence that are in turn matched to the coordinates in the housing market transaction data set. I also use lease data from Allegheny County, provided by FracTracker Alliance. These data can be matched to housing transactions directly using parcel identification numbers. Reliable lease data start from 1995 in both sources. Therefore, all the analysis using these data is restricted to the sample using 1995–2017 housing market transactions in 25 counties. Approximately 2% of houses have an associated lease agreement.

4 Empirical strategy

The baseline estimates are identified using difference-in-differences with continuous treatment and with several treatment types. To define the treated and the pure control groups geographically, I closely follow Muehlenbachs et al. (2015). Figure 4 illustrates these definitions. House A is treated because it is close to a well. A house belongs to the treatment group if the distance between that house and at least one well does not exceed r km. The treatment distance cutoff r is either 2 km, or 1.5 km, or 1 km.[29] Generally, if houses are closer to wells, the negative impacts associated with drilling, such as water contamination or explosion hazard, are expected to be higher. At the same time, landowners can benefit from active production through royalties if wells are very close to or on their property.

House B in the Figure 4 belongs to the pure control group, which includes houses not very close to wells, that is, further than r km. Because these houses are far enough from drilling, it is unlikely that the homeowners face both negative environmental consequences and benefit from royalties. However, these houses need to be close enough to oil and gas extraction to argue that they likely belong to the same local economy and share the labor market. This would help to account for time-varying economic characteristics associated

---

[28] These counties are Armstrong, Blair, Bradford, Cambria, Cameron, Centre, Clearfield, Clinton, Elk, Fayette, Greene, Indiana, Jefferson, Lycoming, McKean, Potter, Somerset, Sullivan, Susquehanna, Tioga, Washington, Westmoreland, and Wyoming.

[29] This definition is rather conservative. As shown in Muehlenbachs et al. (2015), the strongest effect of new production on property prices is observed if the well is approximately 1.5 km away from a property. Appendix B discusses the choice of distance thresholds.

with drilling, such as changes in employment, wages, and local tax revenues. Having these considerations in mind, I include only houses that are no further than 20 km from wells in the pure control group.

The same house can receive several treatment types, depending on the status of nearby wells. A house receives the first treatment when a well is drilled and active. An active well can be producing or temporarily not producing but be capable and equipped for extraction. A house receives the second treatment if an active well becomes abandoned unplugged. If a well is abandoned, it is not expected to produce in the future. Finally, a house receives the third treatment if an abandoned well becomes plugged, which means it is permanently and properly closed. In many cases, houses receive multiple treatments if several wells with different statuses are located nearby.

Because there are multiple treatment types, several pairwise comparisons are possible. For each of these comparisons, one can construct a difference-in-difference estimator, reflecting the house price impact of each treatment type. Those comparisons are shown in Figure 5. The first treatment—drilling an active well—combines several effects. First, house prices may increase because active and producing wells generate royalty income for landowners. Second, drilling introduces environmental risks. Third, active drilling could decrease local amenities since it may cause, for example, noise pollution and increased traffic.[30] Taken together, depending on the relative magnitude of these effects, active wells may increase or decrease prices of nearby houses.

$$\alpha = \Delta Price_{active} - \Delta Price_{no\ wells} = royalties + environment + activity. \tag{1}$$

Equation 1 shows the effects of active production. $\Delta Price_{no\ wells}$ is the change in prices of untreated houses, and $\Delta Price_{active}$ is the change in prices of houses close to active wells.[31] This is a short- to medium run impact of oil and gas extraction as long as wells stay active and generate economic benefits. Let us denote this difference-in-difference coefficient as α.

When an active well becomes abandoned, but not yet plugged, a house receives the second treatment. At this point, production stops without an expectation of renewal. The impact of abandoning an active wells without proper clean-up is described in Equation (2):

$$\beta = \Delta Price_{unplugged} - \Delta Price_{active} = -royalties - activity + monitoring. \tag{2}$$

β comes from comparison houses next to unplugged wells with houses next to active wells.

---

[30] Traffic and noise could be especially concerning during the construction and drilling periods (Goodman et al., 2016, Blair et al., 2018).

[31] These estimates already exist on the literature. See, for example, Boxall et al. (2005), Gopalakrishnan and Klaiber (2013), Muehlenbachs et al. (2015), Bennett and Loomis (2015), Balthrop and Hawley (2017).

All economic benefits disappear; therefore, the landowners lose royalty income. Moreover, possible negative amenities associated with active extraction do not affect housing prices if a well is abandoned. At the same time, because many abandoned wells are not systematically monitored, environmental risks could increase.

Taken together, the long-term impact of gas and oil extraction is measured comparing houses close to unplugged wells with the pure control group, shown in Equation (3). This effect comes from the environmental risks exacerbated by the absence of monitoring at abandoned sites and also includes deterioration of local aesthetic value caused by abandoned sites. The effect is measured by the sum of coefficients α (effect of drilling an active well) and β (effect of abandoning a well without plugging).

$$\alpha + \beta = \Delta Price_{unplugged} - \Delta Price_{no\ wells} = environment + monitoring. \tag{3}$$

The final treatment is plugging an abandoned well; this effect is denoted by γ. When properly plugged, wells impose a much smaller environmental risk and do not deteriorate local aesthetic value. By comparing home prices next to plugged wells with those next to unplugged wells, one can estimate the environmental benefit of proper plugging (reducing environmental risks):

$$\gamma = \Delta Price_{plugged} - \Delta Price_{unplugged} = -environment. \tag{4}$$

Plugged wells pose small risks, and there are also no effects associated with active production. The long-term impact of oil and gas extraction, provided that the environmental requirements are satisfied, is measured by comparing houses next to plugged wells with untreated houses. This measurement is also a sum of the coefficients α (effect of drilling an well), β (effect of abandoning without plugging), and γ (effect of plugging), shown in Equation (5).

$$\alpha + \beta + \gamma = \Delta Price_{plugged} - \Delta Price_{no\ wells} = small\ risks. \tag{5}$$

If plugging completely eliminated negative amenities associated with gas and oil wells, then home values next to plugged wells would be the same as home values in the areas where no drilling ever happened.

4.1 Descriptive event study analysis

Before moving to the main specification, I perform a simple exploratory event study analysis. Instead of a continuous treatment, I use binary treatment variables that indicate if a house has a positive number of wells nearby. The analysis picks up the effect of the first well. I analyze the effects of three events: (i) the first well is drilled and active, measured by coefficient α; (ii) the first well is abandoned without plugging, measured by β; and (iii) the first unplugged well becomes plugged, measured by γ.

To examine the effect of the first active well, I estimate the following equation:

$$\begin{aligned}\ln price_{it} = \sum_{l=-8, l\neq -1}^{8} \alpha_l 1(t = t_i^* + l) \times active_i + \underline{\alpha} 1(t < t_i^* - 8) \times active_i + \\ \bar{\alpha} 1(t > t_i^* + 8) \times active_i + \lambda_{ct} + h_i + \mu_m + \epsilon_{it}. \qquad (6)\end{aligned}$$

In this equation, $active_i$ is the indicator of properties that have at least one active well nearby. The year of drilling the first active well is t_i^*. Time relative to the event of drilling a well is indexed by l. I focus on property sales occurring within 8 years from the beginning of drilling; that is, l varies from -8 to 8. The coefficients of interest are α_l on the interaction of the active well indicator, $active_i$, and the indicator $1(t = t_i^* + l)$. The equation includes county×year fixed effects, λ_{ct}, house fixed effects, h_i, and month of sale fixed effects, μ_m. Because of these fixed effects, I omit $1(t = t_i^* - 1) \times active_i$. The coefficients should be interpreted relative to one year prior to the event. To ensure that the variation comes only from the new drilling, I restrict the sample to houses that either close to new active wells or do not have any changes in the number or status of nearby wells. Observations with other treatment types—that is, houses close to wells switching status from active to abandoned or unplugged to plugged—are excluded.

These event study analyses are useful in several ways. First, they allow me to test the parallel trend assumption of differences-in-differences analysis. If this assumption is satisfied, the coefficients α_l^s should be similar for all pre-treatment observations, that is, for each $l < 0$. Second, they help me evaluate the persistence of house price changes that follows new drilling. This could suggest what effects are more important for home values. For example, houses depreciating only temporarily could indicate a decrease in local amenities due to the process of well drilling and the associated noise, dust, and traffic. On the contrary, house prices that persistently decline can indicate that the loss of royalties and environmental risks drive this effect.

Figure 6(a) plots the coefficients α_l from Equation (6). The standard errors are clustered at the census-tract-by-year level. There are no statistically significant differences in the house price trends between the treated and the untreated houses prior to the event. If anything, homes close to wells slightly appreciate prior to treatment, possibly due to royalty expectations. However, after the first well is drilled, house prices decrease. This decrease becomes slightly larger with time, which could be due to declining productivity of older wells and therefore lower royalty payments. This decrease can also indicate how the local environment deteriorates with time.

I estimate similar equations to examine the effect of the first well becoming abandoned and unplugged. Here I am interested in the coefficients on the interaction of time indicators and the unplugged well indicator. To ensure that the variation comes only from one channel—abandoning without plugging—I restrict the sample to houses receiving only that treatment. Observations with either new drilling or where wells become plugged are excluded. The results are shown in Figure 6(b). Visually, pre-trends are flat and do not seem to be concerning. Once wells become abandoned and unplugged, houses close to them become cheaper relative to houses next to active wells.

The final event study examines the effect of plugging the first abandoned well. The equation is estimated on the sample of houses either with no change in the treatment status or where unplugged wells become plugged. Houses receiving any other types of treatment are excluded. Figure 6(c) shows the estimation results. Even though the standard errors are large, the figure shows that house prices increase following well plugging relative to houses where wells remain unplugged. However, the prices may increase even one period prior to the event. This increase could indicate measurement error when homeowners have information about the plans of the DEP or operators to plug wells before the completion of those requirements is officially reported. Another potential explanation is that the decision to plug could be endogenous: wells could be sealed in the areas where housing prices are moving upward. The main specification addresses these concerns, using the IV.

The results of these analyses should be interpreted with caution. First, because of the sample restrictions, the sample size is significantly smaller than the full model. In practice, several events often happen during the time period we observe housing market transactions; therefore, a house can be treated multiple times. Those treatments could be different by type: for example, in the time between the observed house sales, several wells could be drilled and other wells could become abandoned. Because event study analysis relies on isolating just one treatment type, all observations receiving multiple treatments are omitted from the analysis. Second, the analysis uses only the first well as the event, while the average number of drilled wells (conditional an treatment) is around 15. Finally, the event study analyses do not address endogeneity concerns, discussed in Section 4.3. Even though event studies can provide useful insights about the dynamics of home prices around the events of interest, the estimates are imprecise.

4.2 Baseline specification

The main specification combines all variation coming from well drilling, abandoning, and plugging. Under the assumption that the treatment effect of each well having the same status is identical,[32] I estimate the following equation:

$$\ln price_{it} = \alpha D_{it} + \beta A_{it} + \gamma P_{it} + \lambda_{tc} + h_i + \mu_{im} + \epsilon_{it} = \tag{7a}$$

$$\alpha ACT_{it} + (\alpha + \beta) UP_{it} + (\alpha + \beta + \gamma) P_{it} + \lambda_{tc} + h_i + \mu_{im} + \epsilon_{it}. \tag{7b}$$

Equation (7a) includes three main explanatory variables. D_{it} is the number of drilled wells next to house i at time t, including all wells regardless of their status (active, unplugged, and plugged). A_{it} is the number of abandoned wells, including plugged and unplugged ones. P_{it} is the number of plugged wells. The equation also controls for county-by-year fixed effects, λ_{ct}, house fixed effects, h_i, and month of sale fixed effects, μ_m. Writing the equation in this way allows me to directly estimate housing price responses to changes in well status. The coefficient β should be interpreted as the effect of abandoning relative to active wells, as in Equation (2). This coefficient reflects higher environmental risks due to no monitoring along with the loss of royalties and disamenities from active production. Analogously, coefficient γ is interpreted as the effect of plugging relative to unplugged wells. Equation (4) shows that this coefficient captures a reduction in environmental risks and an increase in local aesthetic value due to land reclamation.

This equation can be easily re-written in a form where only mutually exclusive well categories (i.e., statuses) are used. In this case, the coefficients should be interpreted relative to the pure control group—houses without wells nearby. In Equation (7b), explanatory variables are the number of active wells, ACT_{it}, abandoned unplugged wells, UP_{it}, and plugged wells, P_{it}. The impact of unplugged wells relative to untreated houses is $\alpha + \beta$, as in Equation (3). This is interpreted as the market valuation of the environmental risks from unplugged wells, reflected in the housing prices. Finally, the coefficients $\alpha + \beta + \gamma$ reflect the difference in prices of houses close to plugged wells and the untreated group—small environmental risks. This measurement is the long-term impact of oil and gas extraction if environmental requirements are satisfied.

---

[32] In Section 6.1, I show this assumption, and therefore the linear specification is appropriate.

4.3 Instrument

In literature studying the effects of hazardous waste clean-up, for example, assessing the benefits of the Superfund program (Greenstone and Gallagher, 2008), identification of causal effects is usually challenging. This is because the decision of which sites are cleaned up and which ones are not could be endogenous.

A similar identification challenge arises in assessing the impact of well abandonment and plugging. The timing of abandonment and the likelihood of clean-up depends on several groups of factors. First, the macroeconomic environment, such as gas and oil prices and changes in regulation, is important. When prices are low, more wells may be left behind because they are not economically viable anymore or because some operators declare bankruptcy. Technological development may change the productivity of one type of well relative to another type. Changes in the regulatory environment—such as Act 13 of 2012 in Pennsylvania, discussed in Section 2.2—can affect firms' incentives to abandon or plug. Those factors typically affect all firms (but may differ by well type), and they are likely exogenous to local economic conditions. Macroeconomic environment and tax revenues can also affect the financial capabilities of the DEP to plug wells that do not have a solvent owner.

Second, well abandonment and plugging also depends on well-specific factors, such as its age and production capabilities. These factors can be correlated with the local housing market unobservable characteristics. Older, less productive wells may generate fewer royalties than newer, more productive ones. Moreover, firms could have information about well quality that is not observed by a researcher. They can plug wells that are already problematic or have higher environmental risks if left unplugged and can decide not to plug or postpone plugging wells viewed as less dangerous. This selection is problematic, as more dangerous wells can affect home prices through the decline in environmental quality even before plugging.

Third, well abandonment and plugging depends on operator-level factors, such as firm productivity and their environmental record. The producers could be more likely to plug wells in wealthier areas, where households may have more information, better means to process this information, or more bargaining power. Or more productive firms could offer better terms of lease agreements to homeowners and have better environmental records, which likely correlates with home values while wells are still active. At the same time, such firms may be more likely to abandon and properly plug wells. Less productive firms, on the contrary, may be more likely to file for bankruptcy, leaving unplugged wells. Finally, the DEP might be more likely to plug wells in the areas where public service provision is better in the first place.

In addition to these endogeneity concerns, measurement error could arise from misreporting or failure to report well status and attenuate the ordinary least squares (OLS) estimation toward zero. This issue could be especially serious for houses with few wells nearby.

To address these concerns, I construct an instrument for well abandonment and plugging intensity. This instrument uses the idea that well status depends on macroeconomic conditions, such as technology, regulation, and gas and oil prices, and separates these factors from other local well status determinants. The assumption is that conditional on county-by-year fixed effects, this macroeconomic environment is independent on *local* housing market unobservable characteristics. That is, the changes in macroeconomic conditions are not correlated with the quality of neighborhoods, local wells, and firms.

These changes in the macroeconomic environment may affect wells disproportionately by type. For example, stricter environmental requirements, adopted in 2012, apply only to unconventional wells. Fast technological change could also make conventional, traditional wells less economically viable, leading to higher abandonment rates. Similarly, changes in oil prices affect the economic viability of oil wells, while changes in gas prices affect the economic viability of gas wells. These changes allows me to use the geographical distribution of wells by type as a measure of exposure to those macroeconomic determinants of well abandonment and plugging. The proposed IV combines the type-specific county-level time variation in well abandonment and plugging with the local distribution of wells by type.

This instrument is similar in spirit to a well-known shift-share IV (Bartik, 1991), where the local "share" is the distribution of wells by type and "shift" is the county-level changes in abandoning and plugging of wells that belong to this type. This class of instruments has drawn a lot of attention in recent applied econometrics literature (Jaeger et al., 2018, Goldsmith-Pinkham et al., 2019, Borusyak et al., 2019).

To construct the instrument, I proceed in two steps. The first step isolates the changes in the macro environment that drives well abandonment and plugging. This is the global "shift"—the probability that wells are abandoned or plugged, varying at the county-type-year level. The second step combines these "shifts" with local "shares"—counts of wells by type near each treated house.

In the first step, I estimate linear probability models that use well status indicators as the outcome variables. It essentially decomposes a well status into (i) factors common for all wells within county-type-year and (ii) well-specific factors, which are also the local factors:

$$\Pr(abandoned)_{jt} = \underbrace{\sum_{k=1980}^{2017} \phi_k^a 1(t=k) \times type_j \times county_j}_{\Pr(\widehat{abandoned})_{jt}} + \delta_j + u_{jt}. \tag{8}$$

$$\Pr(plugged)_{jt} = \underbrace{\sum_{k=1980}^{2017} \phi_k^p 1(t=k) \times type_j \times county_j}_{\Pr(\widehat{plugged})_{jt}} + \delta_j + u_{jt}. \tag{9}$$

The outcome variable in Equation (8) is the indicator equal to one if well j is abandoned at time t. I consider the sample of wells that were already drilled by t. The interactions of indicators for year, $1(t = k)$, well type, $type_j$, and county, $county_j$, capture the factors common for all wells within the same county-year-type combination. The regression also includes well fixed effects, δ_j. Equation (9) has the same right-hand side but uses the indicator for plugged status as the outcome variable.

As the result of the first step, I construct the predicted status of each drilled well driven by global factors, $\Pr(\widehat{abandoned_{jt}})$ and $\Pr(\widehat{plugged_{jt}})$. Those global factors can include technology, gas and oil prices, changes in regulation, and other macroeconomic factors. Importantly, this predicted probability does not include the well fixed effects, as it is necessary to avoid any local predictors of well abandonment and plugging. Those fixed effects absorb all well- and firm-specific time-invariant characteristics, such as the year of drilling (and therefore well age), depth, production capabilities, and the quality of well. Well fixed effects can also potentially absorb some location time-invariant characteristics, such as geography and well visibility. Because those characteristics could be correlated with unobserved housing and neighborhood attributes, it is important to exclude them from the instrument. By construction, the predicted probabilities are identical for all wells j' in the same county-year-type: $\Pr(\widehat{abandoned})_{j't} = \Pr(\widehat{abandoned})_{c,t,type}$, where c is county and t is year.

In the main analysis, the explanatory variables of interest are A_{it} and P_{it}—the predicted *numbers* of abandoned and plugged wells close to house i. Those variables are at the house level, while the predicted probabilities, obtained in step one, are at the well level. The second step of the IV construction aggregates the predicted probabilities to the house level, using well locations. This yields the predicted number of wells. To do that, for each house i, I simply sum up the predicted probabilities of being abandoned for nearby wells, drilled by the previous period:

$$\hat{A}_{it} = \sum_{j\ close\ to\ i\ at\ t-1} \Pr(\widehat{abandoned})_{jt} = \tag{10a}$$

$$\sum_{type=1}^{T} D_{i,t-1,type} \times \Pr(\widehat{abandoned})_{c,t,type}. \tag{10b}$$

Equation (10a) yields the predicted number of abandoned wells located close to house i. Because the predicted probabilities are identical for each county-type-year combination, the predicted number of abandoned wells of each type is the count of such wells drilled last period, $D_{i,t-1}$, multiplied by the type-specific probability of being abandoned, $\Pr(\widehat{abandoned})_{c,t,type}$. This gives the predicted number of abandoned wells near house i by type. The total predicted number of abandoned wells is the sum of those type-specific counts over well types (Equation (10b)).

This instrument is a shift-share IV, where $D_{i,t-1}$ are local "shares" and $\Pr(\widehat{abandoned})_{c,t,type}$ is the global "shift." To see this, note that one can write the actual observed count of abandoned wells as the number of drilled wells by type, multiplied by the fraction of abandoned among those wells (summed up together over types): $A_{it} = \sum_{type=1}^{T} D_{i,t,type} \times Fraction\ abandoned_{i,t,type}$. The fraction of abandoned wells is house specific and is therefore indexed by i, t, and $type$. As shown in Equation (10b), the instrument replaces this local fraction with the county-level predicted probability. At the same time, the "share" part of the IV—the number of drilled wells by type—is lagged one period.

Similarly, I construct the predicted number of plugged wells:

$$\hat{P}_{it} = \sum_{j\ close\ to\ i\ at\ t-1} \Pr(\widehat{plugged})_{jt} = \tag{11a}$$

$$\sum_{type=1}^{T} D_{i,t-1,type} \times \Pr(\widehat{plugged})_{c,t,type}. \tag{11b}$$

The predicted number of abandoned wells, $\hat{A}_{it}$, and the predicted number of plugged wells, $\hat{P}_{it}$, are used as the instruments. The first stage of the two-stage least squares (2SLS) estimation is in Equations (12)–(13):

$$A_{it} = \beta^a \hat{A}_{it} + \gamma^a \hat{P}_{it} + \mu_{tc} + h_i + m_{im} + v^a_{it}. \tag{12}$$

$$P_{it} = \beta^p \hat{A}_{it} + \gamma^p \hat{P}_{it} + \mu_{tc} + h_i + m_{im} + \upsilon^p_{it}. \tag{13}$$

Additionally, I estimate the reduced-form model. The estimating equation is similar to (7a), where the observed counts of abandoned and plugged wells are replaced with the instruments:

$$\ln price_{it} = \alpha D_{it} + \beta \hat{A}_{it} + \gamma \hat{P}_{it} + \mu_{tc} + h_i + m_{im} + \epsilon_{it}. \tag{14}$$

The results can be interpreted causally under the assumption that the instrument is uncorrelated (conditional on the set of controls) with unobservable time-varying housing or neighborhood characteristics.

There are two obvious threats to the exclusion restriction. First, the changes in the macroeconomic environment that drives the instrument may affect the housing market through other channels. Specifically, declines in gas and oil prices may not only cause higher well abandonment rates but may also increase local unemployment and reduce income in areas where drilling is common. The changes in those variables in turn can affect house prices. That would violate the exclusion restriction, as it requires that the instrument affects house values only through well abandonment and plugging and not through any other channels. I address this concern in two ways. All specifications include county×year fixed effects, and therefore the variation comes from comparing houses within the same county and year. These fixed effect would absorb the variation in many potentially confounding factors, such as labor market outcomes.[33] Additionally, I directly test the correlation between the instrument and unemployment rate, median income, and share of population below the poverty line in Appendix B. The coefficients are both statistically and economically insignificant.

Second, the macroeconomic environment may affect house prices through changes in royalties from active wells. For example, a decrease in gas and oil prices can simultaneously increase the number of abandoned wells and decrease royalty payments for leaseholders. This is problematic because lower royalty payments (or their expectation) can directly reduce the prices of properties associated with lease agreements. To address this, I use the lease data and restrict the sample only to houses that do not have lease agreements, disabling this royalty channel. The results are in Section 5.2.

---

[33] Labor markets are usually geographically defined as commuting zones, which typically include more than one county.

5 Results

5.1 Baseline specification

Table 3 shows estimation results of Equation (7a). The sample is restricted to houses that have no more than 100 wells nearby. Because house prices could be correlated within a neighborhood, and because the instrument varies by county-year-type, the standard errors are clustered at the county×year level.

Columns (1)–(3) show the OLS results, and columns (4)–(6) show the 2SLS results. The distance cutoff used to define the treatment group varies from 2 km in columns (1) and (4) to 1 km in columns (3) and (6). The OLS results in column (1) show that each active well leads to a 0.1% decline in house prices within 2 km. This measurement is the coefficient α. Even though the price decline is statistically significant, it is small. An average treated house has 15.2 drilled wells nearby; therefore, the price decline is 1.52%. Because an average treated house price is $\$139,347$, the loss is $\$139,347 \times 0.0152 = \$2,118$.[34]

Coefficient β shows that when a well changes status from active to abandoned, house prices decline by 1.3%. This coefficient comes from the comparison of houses close to active wells with houses close to unplugged ones. The long-term impact of oil and gas extraction without proper plugging is measured by $\alpha + \beta = -0.014$ per well, and an average treated house has 2.8 unplugged wells. According to the estimates, this leads to a 3.92%, or $\$5,462$, decline in home prices.

Remarkably, plugging an abandoned well restores prices almost entirely. The coefficient for the number of plugged wells, γ, is 0.012, which should be interpreted relative to all abandoned wells. The long-term impact of oil and gas extraction is therefore $\alpha + \beta + \gamma = -0.001$. This effect is not statistically significant.

Columns (2) and (3) in Table 3 show that when the cutoff moves closer to houses, the coefficients also become smaller by magnitude and statistically insignificant. This attenuation could potentially be explained by higher measurement error. A narrower definition of the treatment group leads to both fewer observations that are close to wells changing status and a smaller number of those wells (Table 2). This could attenuate the results if the status of wells is reported imprecisely. In column (3), when a 1 km distance cutoff is used, the coefficient for the number of plugged wells even becomes negative, though insignificant. This is due to the increased role of royalties for houses very close to drilling, as shown in Section 5.2.

---

[34] Muehlenbachs et al. (2015) show that groundwater-dependent properties within 1–1.5 km of active wells lose, on average, around \$30,167, while piped-water properties gain \$4,802. My specification does not make the distinction by water source and therefore represents the average result.

The coefficients in the 2SLS regressions are much larger in magnitude but suggest the same big picture as the OLS results. The positive coefficient for drilled wells, α, in columns (4)–(6) is explained by higher royalties for houses close to active wells (Section 5.2). Naturally, this coefficient is larger in magnitude in column (6), because royalties are more important if active wells are closer.

If an active well is abandoned without plugging, prices of houses within 2 km decline by 3.9%. An average treated house therefore depreciates by 10.9%, or $\$15,188$. This effect becomes larger in magnitude for houses closer to wells—within 1.5 and 1 km—likely because of higher perceived environmental risks and visibility. For an average treated house within 1.5 km, the loss is 19.7%, or $\$27,465$, and for a house within 1 km the loss is 39%, or $\$54,500$. Interestingly, for houses within 1–1.5 km of wells, plugging does not restore prices completely. Even though the coefficients are marginally significant, houses within 1.5 km of plugged wells are cheaper by 10%, and within 1 km by 12%, than properties without close wells. This is an improvement in comparison to unplugged wells, but the effect is still is sizable in magnitude.

The estimates imply that the OLS results are likely to be biased toward zero. This is consistent with the measurement error story when abandonment and plugging might be either improperly reported or reported with a delay. This problem could be substantial because, on average, houses have few abandoned wells nearby. Measurement error is much smaller at a more aggregated level, such as a county. The instrument uses this fact and relies on the variation in abandonment and plugging at the county level instead of the individual house level.

First-stage results for a 2 km distance cutoff are reported in Table 4. The instruments are predictive of the actual counts of both abandoned and plugged wells. The reduced-form estimation in Table 5 suggests a similar big picture to 2SLS coefficients, but the results are noisier and somewhat smaller in magnitude.

5.2 The role of royalties

Taking royalties into account is important for two reasons. First, these data provide an additional way to disentangle two channels through which well abandonment affects home prices: the loss of royalties and worsening environment. Second, these data help to support the use of the instrument. As discussed in Section 4.3, it is hard to justify that the exclusion restriction holds for houses associated with lease agreements. This is because gas and oil prices, correlated with the instrument, also affect home prices through changes in royalty payments. To avoid this, I estimate the 2SLS specification on the sample of houses that *are not* associated with a lease agreement.

To evaluate the role of royalties in the observed response to well abandonment and plugging, I introduce an indicator variable $Lease_i$ that takes a value of one if house i was ever associated with a lease agreement. This characteristic is time invariant. I estimate the following triple differences model:

$$\begin{aligned} \ln price_{it} = \alpha D_{it} + \beta A_{it} + \gamma P_{it} + \alpha_L D_{it} \times Lease_i + \beta_L A_{it} \times Lease_i + \gamma_L P_{it} \times Lease_i \\ + \mu_{tc} + h_i + m_{im} + \epsilon_{it}. \end{aligned} \tag{15}$$

This specification builds on the baseline, adding the interactions between the lease agreement indicator and the numbers of drilled, abandoned, and plugged wells. The coefficient α_L is the house price impact of nearby active wells on properties with leases, relative to properties without them. Similarly, coefficient β_L shows the difference in housing price responses to well abandonment without plugging, and γ_L shows the difference in responses to well plugging. Due to lease data limitations, I estimate this regression on a smaller sample, which includes 25 out of 35 counties used in the full sample and sales only after 1995. I estimate this regression using OLS, and the results are in Table 6.

Columns (1)–(3) are the estimation results of Equation (7a) on the restricted sample for reference. Because the sample is much smaller, the results are noisier than the baseline. Columns (4)–(6) estimate Equation (15), changing the distance cutoff. Active wells decrease house prices by 0.3%–0.4% per well (depending on the distance cutoff) for houses without royalties. Not surprisingly, active wells *increase* the price of properties associated with leases by 1%–2.2% per well.

When a well becomes abandoned and unplugged, the house price reduction is larger for properties with leases because the landowners lose the royalty income. The total long-term impact of drilling if wells are left unplugged is also more negative for such properties, even though this difference is not statistically significant. This difference between houses with and without leases is measured by $\alpha_L + \beta_L = -0.02$ for properties within 2 km of wells. The difference likely exists because such houses are very close to wells and therefore may have larger environmental risks.

Next, I estimate the baseline specification in Equation (7a), using 2SLS and the sample of houses without an associated lease agreement. The sample is also restricted to 25 counties and the time period after 1995. The results are in Table 7. If anything, the results are stronger, suggesting the 2SLS estimates in the baseline are not driven by royalties.

6 Robustness

I perform several robustness checks to examine how the results change for different subsamples and definitions of the treatment and control groups. I also argue that the linear functional form is an appropriate approximation. While the magnitude of the coefficients slightly changes depending on the specification, the main conclusions of the analysis are robust.

6.1 Functional from

To show that the linear estimating equation is appropriate, I change the specification to allow for a more flexible functional form. Instead of using the numbers of wells as the right-hand-side variables, I use indicators for each well count:

$$\ln price_{it} = \sum_{k=1}^{9} \alpha_k 1(D_{it} = k) + \bar{\alpha} 1(D_{it} \geq 10) + \sum_{k=1}^{9} \beta_k 1(A_{it} = k) + \bar{\beta} 1(A_{it} \geq 10) + \sum_{k=1}^{9} \gamma_k 1(P_{it} = k) + \bar{\gamma} 1(P_{it} \geq 10) + \mu_{ct} + h_i + m_{im} + \epsilon_{it}. \quad (16)$$

Equation (16) regresses log house price on the following set of indicators. $1(D_{it} = k)$ takes a value of one if exactly k wells were drilled close to house i, provided that k is less than 10. $1(D_{it} \geq 10)$ is the indicator for houses with ten or more drilled wells. Indicators for abandoned wells counts, $1(A_{it} = k)$, and plugged well counts, $1(A_{it} = k)$, are defined analogously. The coefficients for abandoned wells, β_k, and for plugged wells, γ_k, could potentially vary depending on the number of wells, allowing for the non-constant marginal effect of abandoning and plugging.

The OLS results are in Figure 7, and the reduced-form estimation is in Figure 8. The figures plot the coefficients from Equation (16) on the y-axis, and the number of wells is on the x-axis. The results suggest that the linear functional form is appropriate for this analysis, at least if the number of drilled wells is not too large.[35]

6.2 Definition of the control and treatment

I change the definition of the control and treatment groups to allow foe the distance buffer between them. For this analysis, the single pure control group is limited to houses located

---

[35] As mentioned in Section 4.2 above, all analysis restricts the sample to houses with no more than 100 wells.

within 20 km of wells but no closer than 2 km. There are two treated groups: (i) houses within 1.5 km and (ii) houses within 1 km of wells. Because the control group is fixed for this analysis, houses within 1.5–2 km (or 1–2 km in the second case) are dropped from the sample, creating a buffer between the treated and untreated observations. The results, shown in Table A.1, are very similar to the main specification.

6.3 Timing of treatment

Selection at the house level is possible: drilling may not be located randomly in space. Not only well status can be endogenous but also the decision whether to drill and well location. Houses close to wells can be different from houses far from them in a way not controlled for by property fixed effects. For example, wealthier homeowners may oppose drilling in their neighborhoods, while less wealthy residents may welcome it, expecting higher income from royalties. To address this, I eliminate the pure control group from the analysis. The variation therefore comes only from the timing of drilling, abandonment, and plugging, while all observations are treated within the sample period. The results are less precise due to a much smaller sample size but are essentially unchanged (see Table A.2).

6.4 Other sample restrictions

The sample for the baseline analysis includes only houses with no more than 100 wells nearby. Even with this restriction, the distribution of wells across houses is right skewed with the mean equal to 15 and median equal to 7. To make sure that the results are not driven by the right tail of the well distribution, I restrict the sample to houses with no more than 75, 50, and 25 wells nearby. As expected, the results are slightly weaker (see Table A.3), but the big picture is unchanged.

While many wells are drilled in rural Pennsylvania, Allegheny County is the heart of the Pittsburgh metropolitan area.[36] As discussed in Section 3.2, pre-treatment prices for control and treated houses are different in Allegheny County but not in the rest of the sample. To see how much this affects the results, I estimate Equation (7a) on the sample of houses excluding Allegheny County. The results, shown in Table A.5, are very similar to the main specification.

---

[36] There are multiple definitions of urban areas. Even though several other counties are included in metropolitan areas by the Office of Management and Budget (OMB), only Allegheny County stands out as the most populated one.

6.5 Alternative instruments

The instrument in the baseline specification uses the geographical distribution of wells by type one year prior to the observed house sale as the "shares." This version of the instrument provides the strong first stage. However, this comes at a cost: it is possible that recent well drilling could be correlated with house prices in a way not captured by the controls. To decrease the likelihood of this, I use the distribution of wells five years prior and ten years prior to house sales.

$$\hat{A}^{l}_{it} = \sum_{type=1}^{T} D_{i,t-l,type} \times \Pr(\widehat{abandoned})_{c,t,type}. \tag{17}$$

Equation (17) shows the construction of the predicted number of abandoned wells, using the earlier distribution of wells across houses. The lag used to construct the "shares," $D_{i,t-l,type}$, is denoted by l and is equal to either 5 or 10. The predicted number of plugged wells is constructed similarly. While the likelihood of correlation between these well counts and housing market unobservables is lower, the first stage of such an instrument is weaker. The estimation results are in Table A.4. The standard errors are larger, but the coefficients are similar to the baseline.

7 Back-of-the-envelope cost-benefit analysis

The results of the hedonic housing price model provides an estimate of how much the residents are willing to pay for being away from abandoned unplugged wells. At the same time, it estimates how the homeowners perceive the benefits of plugging. In this section, I compare these benefits with the cost of plugging.

The cost of plugging depends on many factors, but, on average, it is $\$24,000$ for conventional wells and $\$100,000$ for unconventional wells. Those estimates are taken from Mitchell and Casman (2011) and are in 2011 dollars. In 2019 dollars, the costs are $\$28,145$ and $\$117,274$. OLS estimates imply that if one well is plugged, each house within 2 km of that well gains 1.2% of its value, or $\$139,347 \times 0.012 = \$1,672$. However, there is typically many houses near a well. In 2017, there were, on average, 95 houses within 2 km of each drilled well, allowing me to calculate the benefit of plugging for all nearby houses, which is $\$1,672 \times 95 = \$158,840$.

Estimates of 2SLS show that the benefit of plugging a well is 3% of a house value. Similar calculations imply that the dollar value (in 2019 dollars) is $\$4,141$ per house per well, or

$393,435 for all houses.

I interpret these values as the lower bound estimate of the benefits of plugging one well. There are others, not captured by the hedonic estimates, such as climate change benefits from methane emissions reduction. As pointed out by Raimi et al. (2020), well plugging also creates advantages to local economies through job creation. Even though one might expect higher total benefits of plugging, my lower bound estimates suggests that the gains are larger than its costs.

8 Conclusion

The future of oil and gas extraction sector is unclear, and it has been a widely discussed political topic. This paper evaluates the housing market impact of oil and gas extraction in the long term. I analyze how drilling affects property values throughout the well life cycle, starting from the active phase and ending with abandonment and plugging.

I find that well abandonment without proper clean-up has a large negative impact on house prices. This impact quantifies the long-term cost of oil and gas extraction if environmental requirements are ignored. However, if land reclamation and well plugging is performed, house prices restore almost completely, quantifying the benefits of plugging. This conclusion is robust to different specifications and methods.

My estimates imply that the benefits from proper clean-up are larger than its costs. My results have important policy implications. First, they support publicly funded well plugging programs, such as the Pennsylvania Abandoned Wells Program. Proper well site clean-up is important not only from a long-term environmental standpoint, but it also helps local households that partially bear the costs of drilling. Second, this paper motivates environmental policies aimed at creating incentives for well owners to plug wells. These policies may include higher bonding requirements, fines for non-compliance, or environmental taxes.

References

Balthrop, A. A. 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):105–55.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? *W.E. Upjohn Institute for Employment Research.*

Bennett, A. and Loomis, J. (2015). Are housing prices pulled down or pushed up by fracked oil and gas wells? a hedonic price analysis of housing values in weld county, colorado. *Society and Natural Resources*, 28(11):1168–1186.

Black, K. J., McCoy, S. J., and Weber, J. G. (2017). When externalities are taxed: The effects and incidence of pennsylvania's impact fee on shale gas wells. *Journal of the Association of Environmental and Resource Economics*, 5(1).

Blair, B. D., Brindley, S., Dinkeloo, E., McKenzie, L. M., and Adgate, J. L. (2018). Residential noise from nearby oil and gas well construction and drilling. *Journal of Exposure Science Environmental Epidemiology*, 28.

Boomhower, J. (2019). Drilling like there's no tomorrow: Bankruptsy, insurance, and environmental risk. *American Economic Review*, 109(2):391–426.

Borusyak, K., Hull, P., and Jaravel, X. (2019). Quasi-experimental shift-share research desings. *NBER Working Paper*, (24997).

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*, 27:248–269.

Cabot Oil Gas Corporation (2010). Summary of Cabot's Good Faith Efforts.

Considine, T. J., Considine, N. B., and Watson, R. (2015). Economic and environmental impacts of fracking: A case study of the marcellus shale. *International Review of Environmental and Resource Economics*, 9(3-4):209–244.

Cunningham, C., Gerardi, K. S., and Shen, L. (2020). Double trigger for mortgage default: The fracking boom. *Available at SSRN:*.

Currie, J., Greenstone, M., and Moretti, E. (2011). Superfund cleanups and infant health. *American Economic Review*, 101:435–441.

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.

Environmental Protection Agency (2020). Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2018. Technical report.

Ericson, S. J., Kaffine, D. T., and Maniloff, P. (2020). Costs of increasing oil and gas setbacks are initially modest but rise sharply. *Energy Ploicy*, 146.

Feyrer, J., Mansur, E. T., and Sacerdote, B. (2017). Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review*, 107(4):1313–34.

Gamper-Rabindran, S. and Timmins, C. (2013). Does cleanup of hazardous waste sites raise housing values? evidence of spatially localized benefits. *Journal of Environmental Economics and Management*, 65:345–360.

Glaeser, E. L. and Gyourko, J. (2005). Urban decline and durable housing. *Journal of Political Economy*, 113(2):345–375.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2019). Bartik instruments: What, when, why and how. *NBER Working Paper*, (24408).

Goodman, P. S., Galatioto, F., Thorpe, N., Namdeo, A. K., Davies, R. J., and Bird, R. N. (2016). Investigating the traffic-related environmental impacts of hydraulic-fracturing (fracking) operations. *Environmental International*, 89-90.

Gopalakrishnan, S. and Klaiber, H. A. (2013). Is the shale energy boom b bust for nearby residents? evidence from housing values in pennsylvania. *American Journal of Agricultural Economics*, 96(1):43–66.

Greenstone, M. and Gallagher, J. (2008). Does hazardous waste matter? evidence from the housing market and the superfund program. *Quarterly Journal of Economics*, 123.

Ground Water Protection Council (2011). State Oil and Gas Agency Groundwater Investigations And Their Role in Advancing Regulatory Reforms. A Two-State Review: Ohio and Texas. Technical report, Ground Water Protection Council.

Haynes and Boone, LLP (2020). Oil Patch Bankruptcy Monitor. Technical report.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., J.Guiot, Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S., A. Thomas, R. W., , and Zhou, G. (2018). *Impacts of 1.5^{o}C Global Warming on Natural and Human Systems*, chapter 3. Global Warming of 1.5°C. An IPCC

Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte et. al. (eds.)].

Independent Fiscal Office (2020). Natural Gas Royalties Increase in 2017. Technical report, Independent Fiscal Office.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. *NBER Working Paper*, (24285).

Kang, M., Christian, S., Celia, M. A., Mauzerall, D. L., Bill, M., Miller, A. R., Chen, Y., Conrad, M. E., Darrah, T. H., and Jackson, R. B. (2017). Identification and characterization of high methane-emitting abandoned oil and gas wells. *Proceedings of the National Academy of Sciences of the United States of America*, 113(48):13636–13641.

Kim, B. and Oliver, M. E. (2017). Taming drillers through legislative action: evidence from pennsylvania's shale gas industry. *Resource and Energy Economics*, 50:15–35.

Lewis, C., Greiner, L., and Brown, D. R. (2018). Setback distances for unconventional oil and gas development: Delphi study results. *PLoS One*, 13.

Maniloff, P. and Mastromonaco, R. (2017). The local employment impacts of fracking: A national study. *Resource and Energy Economics*, 49(1):62–85.

Manson, S., Schroeder, J., Van Riper, D., and Ruggles, S. (2019). Ipums national historical geographic information system: Version 14.0 [database].

Mitchell, A. L. and Casman, E. A. (2011). Economic incentives and regulatory framework for shale gas well site reclamation in pennsylvania. *Environmental Science and Technology*, 45(22):9506–9514.

Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, 105(12):3633—3659.

Pennsylvania Department of Environmental Protection (2014). Abandoned and Orphan Oil and Gas Wells and the Well Plugging Program. Technical report, Pennsylvania Department of Environmental Protection, 8000-FS-DEP1670.

Pew Charitable Trusts (2020). In Slumping Energy States, Plugging Abandoned Wells Could Provide an Economic Boost. Technical report.

Raimi, D., Nerurkar, N., and Bordoff, J. (2020). Green stimulus for oil and gas workers: Considering a major federal effort to plug orphaned and abandoned wells. *Columbia University CGEP*.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 90(6).

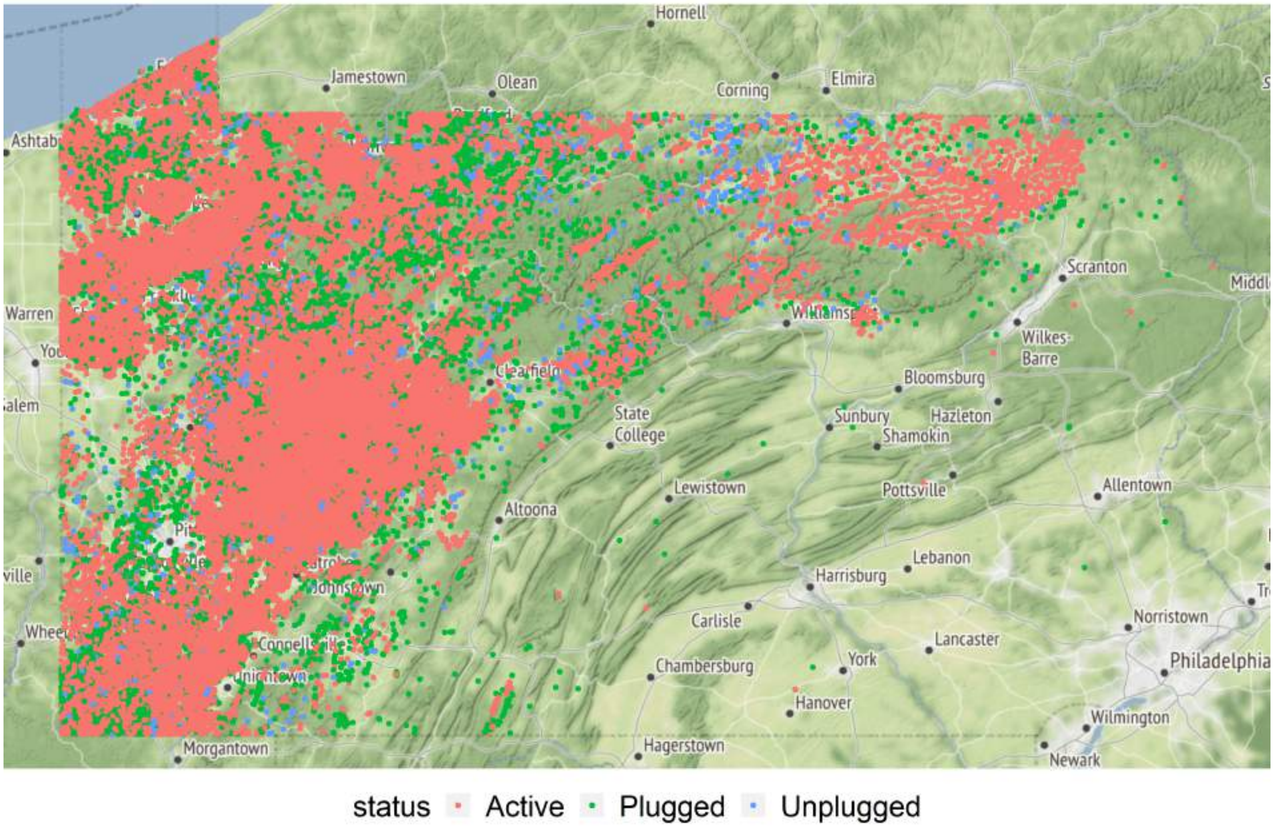
Rosen, S. (1979). Wage-based indexes of urban quality of life. In Mieszkowski, P. and Strassheim, M., editors, *Current Issues in Urban Economics*, pages 74–104. Johns Hopkins University Press, Baltimore, Maryland.

Townsend-Small, A., Ferrara, T. W., Lyon, D. R., Fries, A. E., and Lamb, B. K. (2016). Emissions of coalbed and natural gas methane from abandoned oil and gas wells in the united states. *Geophysical Research Letters*, 43:2283–2290.

Waxman, A. R., Khomaini, A., Leibowicz, B. D., and Olmstead, S. M. (2020). Emissions in the stream: estimating the greenhouse gas impacts of an oil and gas boom. *Environmental Research Letters*, 15(1).

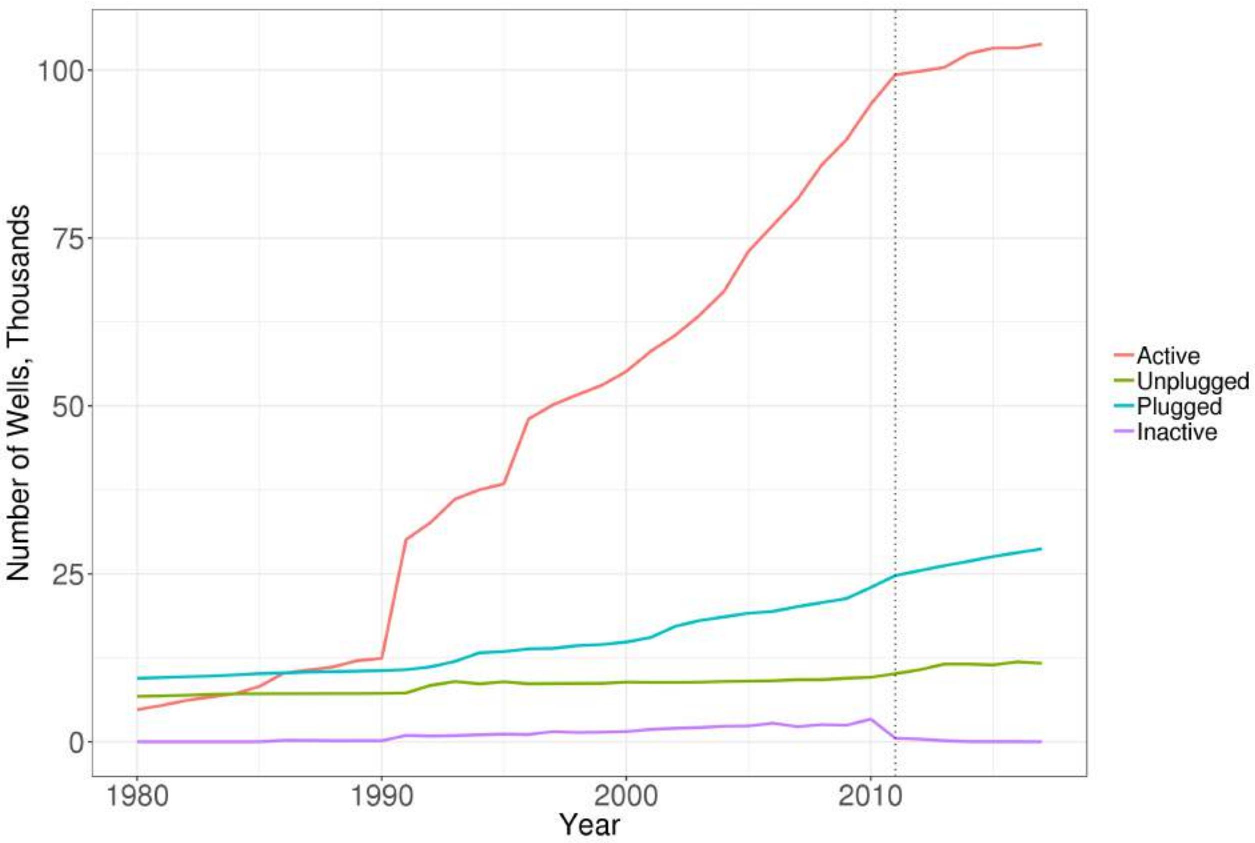
Figures

Figure 1: Drilled Wells in Pennsylvania, by Status



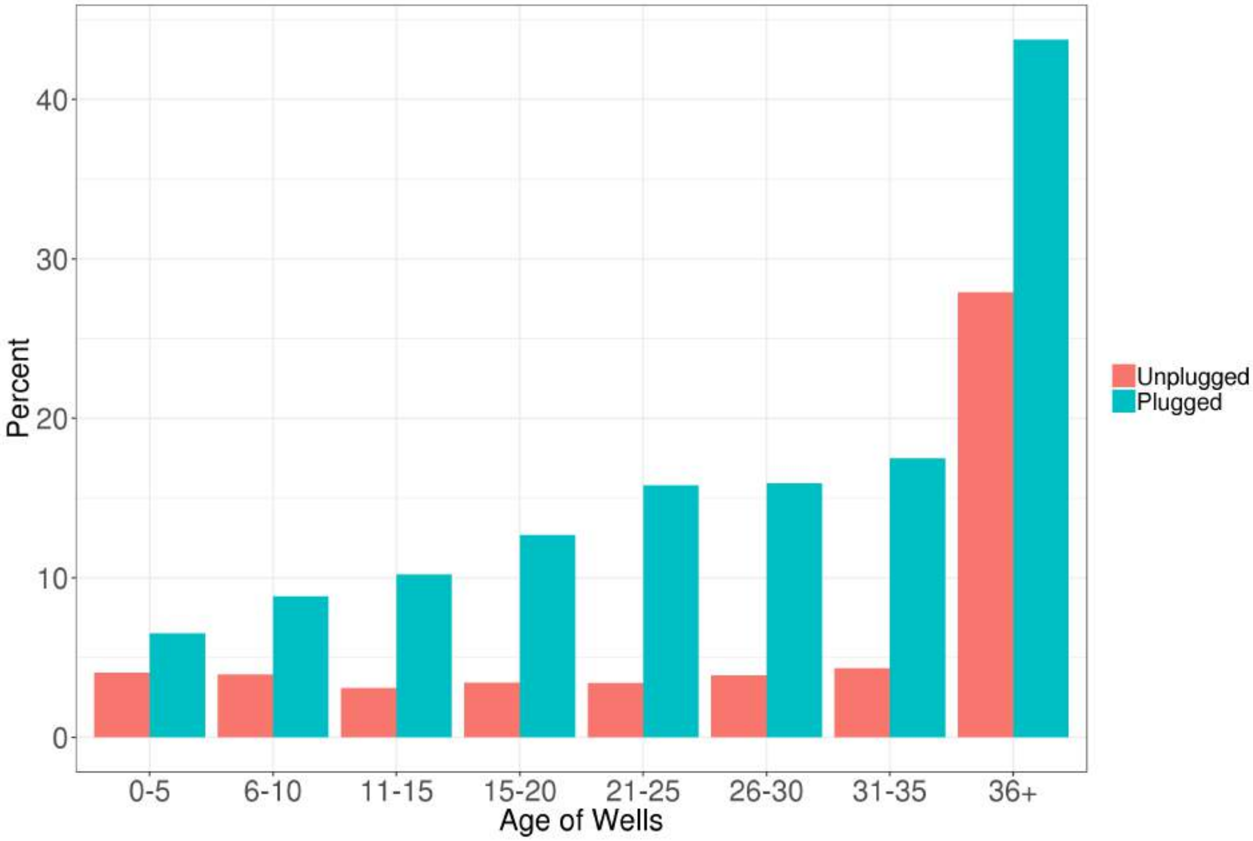
Note: Source of well data: Department of Environmental Protection Spud database. The number of known wells: 204,300 (February 2019). The location of 59,235 wells is unknown (not on the map). Of the wells with an unknown location, 100% are conventional, 3.2% are unplugged, 61.1% are plugged, and 35.6% are active.

Figure 2: Summary Statistics of Known Wells, by Status



Note: Source of well data: Department of Environmental Protection Spud database and production reports. Abandoned wells are not used for production for at least 12 months or are not equipped for production. Plugged wells are properly closed to prevent leaking, and unplugged wells are abandoned without plugging. Inactive wells are temporarily not producing but can produce.

Figure 3: Summary Statistics of Known Wells, by Age



Note: Source of well data: Department of Environmental Protection Spud database and Production Reports, 1980–2017. Approximately 26.3% of wells were ever abandoned, 18.2% were ever plugged, and 10.2% were ever left unplugged.

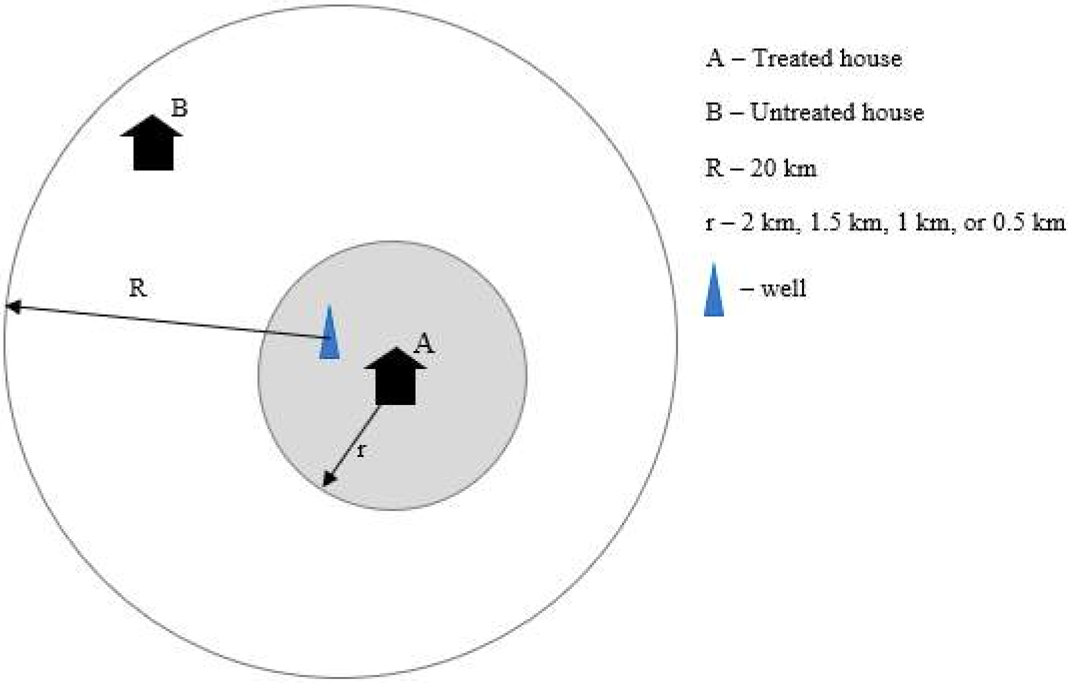


Figure 4: Treatment and Pure Control Groups: Definition

Note: House A is treated, because it is close (i.e., within r km) to a well. The treatment type is determined by the status of nearby wells (drilled, abandoned, plugged). The treatment intensity is determined by the number of wells. House B is the pure control group, because it is further than r km but closer than 20 km to a well.

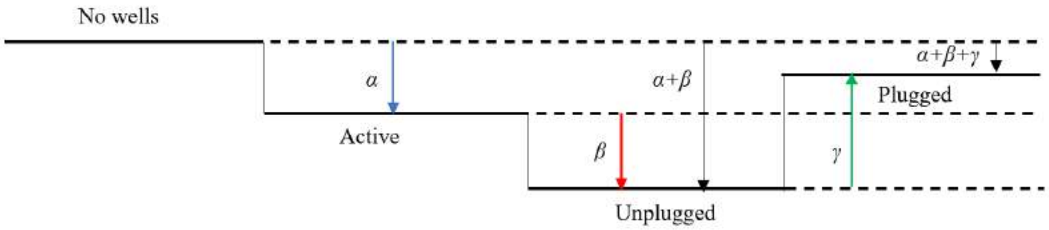
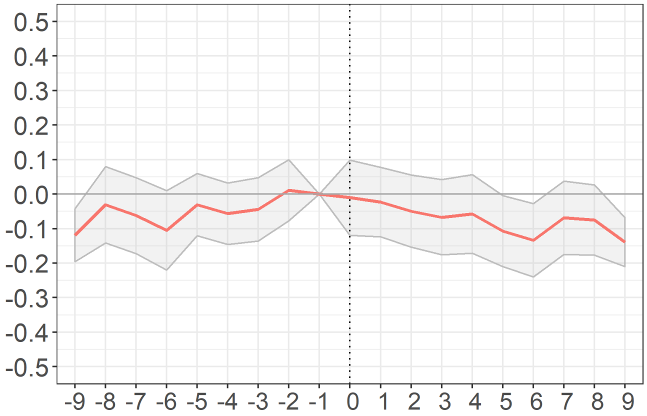


Figure 5: Pairwise Difference-in-Differences Estimators

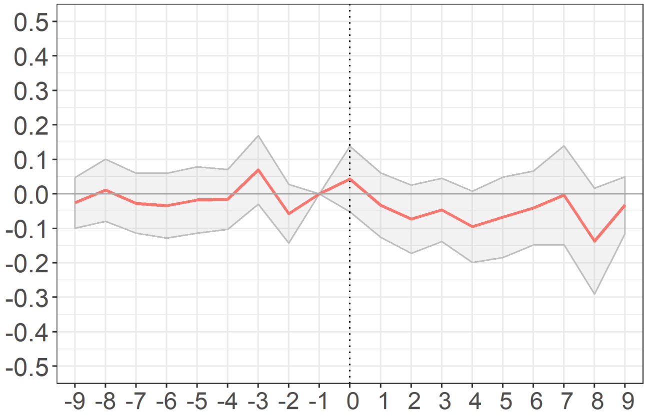
Note: Possible pairwise comparisons of untreated houses (no wells) and houses with different treatment types. α captures the effect of proximity to active wells—a combination of positive impact from royalties, negative impact from the environmental risks, and decrease in amenities from active drilling. β captures the effect of abandoning an active well without proper plugging. The positive effect of royalties and the disamenities associated with active drilling go away, but the environmental risks increase. $\alpha + \beta$ shows the long-term impact of oil and gas extraction if wells are not properly closed. γ captures the effect of plugging that reduces the environmental risks. $\alpha + \beta + \gamma$ is the long-term impact of oil and gas extraction if wells are properly closed.

Figure 6: Event Study Analysis: Effect of Change in Well Status on Houses within 2 km

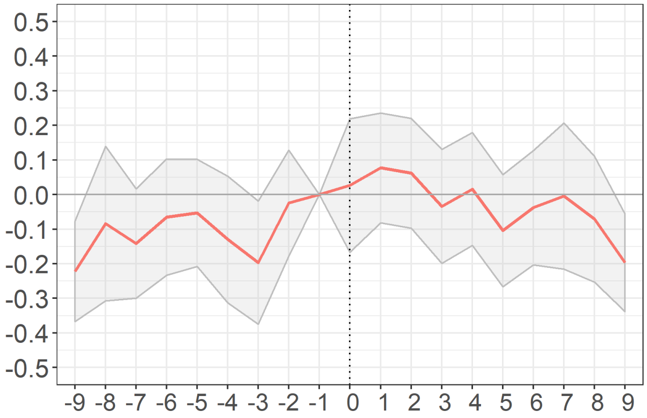
(a) From No Wells to Active



(b) From Active to Unplugged

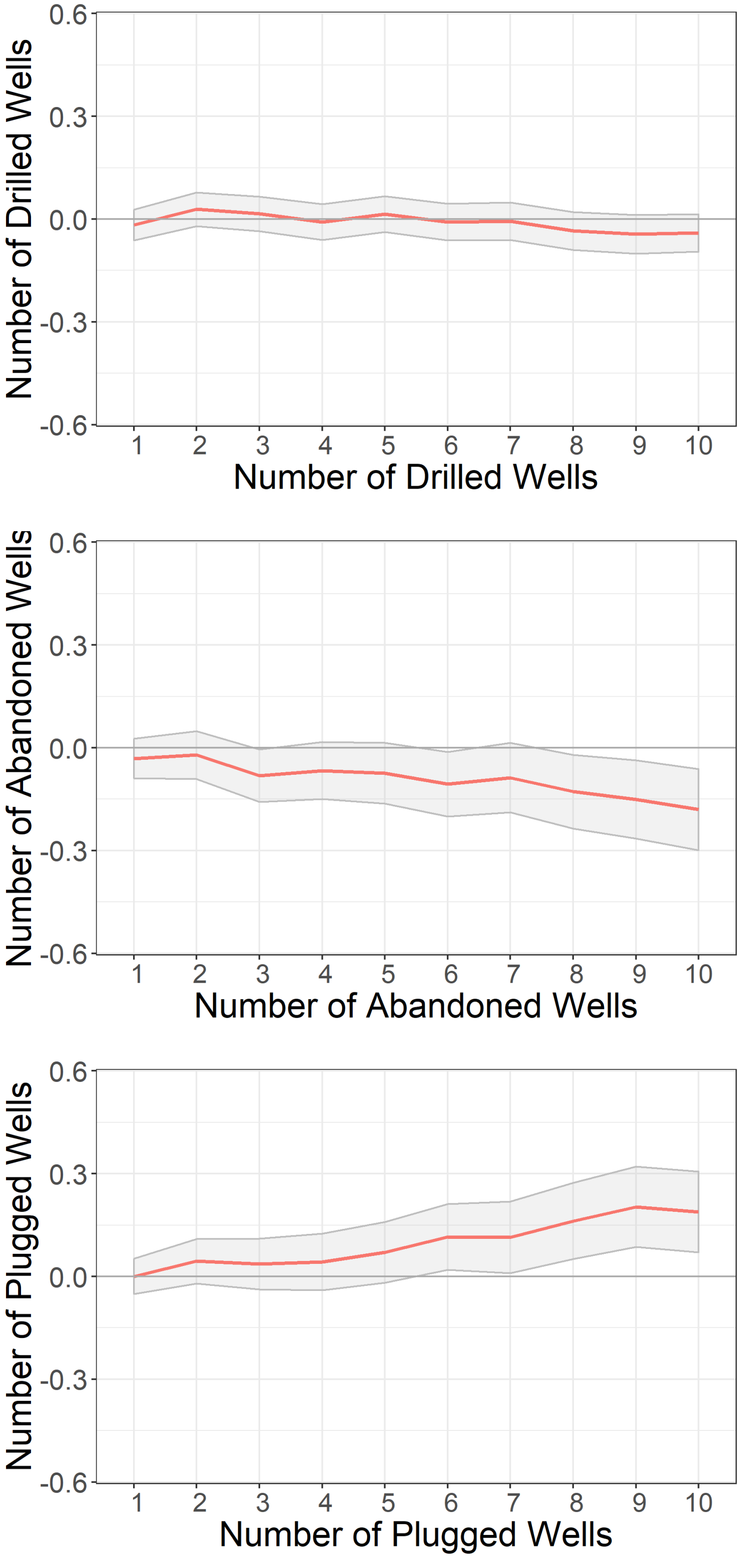


(c) From Unplugged to Plugged



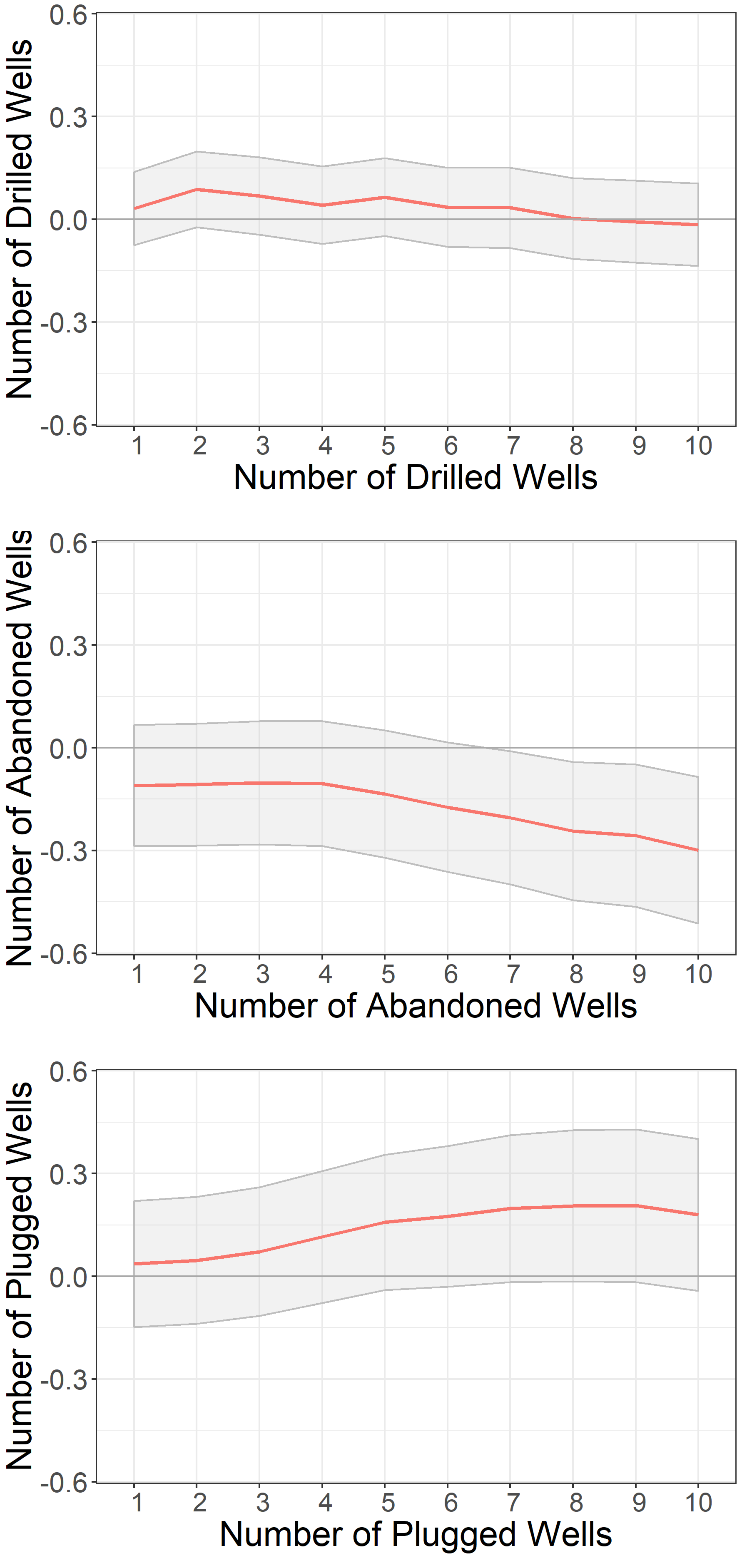
Note: Figures plot the coefficients from Equation (6). The coefficient for one year prior to the event is normalized to zero. Treated houses are within 2 km of a well. Shaded areas are 95% confidence intervals, clustered at the county-by-year level.

Figure 7: Regression Results: Flexible Functional Form, OLS



Note: Figures plot the coefficients from equation (16). Shaded areas are 95% confidence interval, clustered at the county×year level.

Figure 8: Regression Results: Flexible Functional Form, Reduced Form



Note: Figures plot the coefficients from equation (16), using the predicted numbers of abandoned wells, $\hat{A_{it}}$, and plugged wells, $\hat{P_{it}}$, as the explanatory variables (reduced form). Shaded areas are 95% confidence intervals, clustered at the county×year level.

Tables

Table 1: Summary Statistics of House Sales

	Control Mean	Treatment Mean	P-value
Price (pre-treatment, K 2019 USD)	145.5	139.2	0.01
Bedrooms	3.00	3.00	0.70
Basement	0.91	0.95	0.00
Brick	0.36	0.34	0.00
Year Built	1946	1951	0.00
Central Air	0.35	0.51	0.00
Fire	0.27	0.28	0.00
Building Size (K sqft)	16.8	16.2	0.00
Lot Size (K sqft)	27.9	33.6	0.35
Stories	1.62	1.55	0.00

Notes: Sample of houses sold at least twice between 1980 and 2017 in 35 Pennsylvania counties. The control group includes houses further than 2 km from wells. Treated houses have at least 1 (but no more than 100) well within 2 km drilled within the sample period. *P*-values are for a simple t-test for differences in means.

Table 2: Summary Statistics of House Sales and Matched Wells

Type of well		Share of houses		Number of wells if treated	
		At least one well	Number of wells changed	Mean	Median
Drilled within	2 km	0.61	0.37	15.18	7
	1.5 km	0.53	0.30	9.82	5
	1 km	0.41	0.21	5.59	3
Abandoned within	2 km	0.56	0.28	6.15	4
	1.5 km	0.46	0.20	4.30	2
	1 km	0.32	0.13	2.82	2
Unplugged within	2 km	0.27	0.12	2.76	2
	1.5 km	0.19	0.08	2.19	1
	1 km	0.12	0.05	1.70	1
Plugged within	2 km	0.54	0.24	5.07	3
	1.5 km	0.43	0.18	3.61	2
	1 km	0.29	0.11	2.43	2
Observations			450,468		

Notes: Sample of houses sold at least twice in 35 Pennsylvania counties between 1980 and 2017. The sample is restricted to houses that have no more than 100 wells nearby.

Table 3: Regression Results: Linear Specification

	Dependent variable:					
	ln(*price*)					
	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	-0.001^{*}	-0.001	0.001	0.0002	0.003^{**}	0.013^{**}
α	(0.0005)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)
Number Abandoned	-0.013^{***}	-0.007	0.002	-0.039^{***}	-0.093^{***}	-0.251^{***}
β	(0.004)	(0.006)	(0.011)	(0.013)	(0.019)	(0.094)
Number Plugged	0.012^{**}	0.004	-0.015	0.030^{**}	0.061^{**}	0.190^{**}
γ	(0.005)	(0.007)	(0.012)	(0.015)	(0.026)	(0.081)
Dist. cutoff	2 km	1.5 km	1 km	2 km	1.5 km	1 km
Observations	431,757	435,008	436,361	431,757	435,008	436,361
Adjusted R^2	0.560	0.559	0.559	0.560	0.558	0.557
$\alpha + \beta$	-0.014	-0.008	0.003	-0.039	-0.090	-0.236
p-val.	0.002	0.206	0.77	0.006	<0.001	0.020
Avg. unplugged	2.76	2.19	1.70	2.76	2.19	1.70
$\alpha + \beta + \gamma$	-0.001	-0.004	-0.011	-0.009	-0.028	-0.048
p-val.	0.501	0.090	0.001	0.532	0.058	0.055
Avg. plugged	5.07	3.61	2.43	5.07	3.61	2.43

Notes: Houses with no more than 100 wells nearby. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4: First-Stage Results

	Dependent variable:	
	Abandoned	Plugged
	(1)	(2)
Number of Drilled Wells	-0.139^{***}	-0.127^{***}
D	(0.012)	(0.009)
Predicted Number of Abandoned Wells	0.772^{***}	-0.177^{***}
$\hat{A}$	(0.07)	(0.031)
Predicted Number of Plugged Wells	0.097	1.277^{***}
$\hat{P}$	(0.115)	(0.068)
R^2 (proj. model)	0.184	0.158
F Stat. (excl. instr.)	144.9	145.3

Notes: Estimation results of Equation (7a) for houses with no more than 100 wells nearby. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. The treatment distance cutoff is 2 km. The predicted number of abandoned wells is as in Equation 10a. The predicted number of plugged wells is as in Equation 11a.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5: Reduced-Form Results

	Dependent variable:		
	ln(*price*)		
	(1)	(2)	(3)
Number of Drilled Wells	0.002	0.006^{***}	0.013^{***}
α	(0.002)	(0.002)	(0.004)
Predicted Number of Abandoned Wells	-0.036^{***}	-0.070^{***}	-0.078^{***}
β	(0.011)	(0.014)	(0.027)
Predicted Number of Plugged Wells	0.035^{*}	0.061^{**}	0.037
γ	(0.018)	(0.021)	(0.040)
Observations	431,757	435,008	436,361
Adjusted R^2	0.560	0.559	0.559
Dist. cutoff	2 km	1.5 km	1 km
$\alpha + \beta$	−0.034	−0.069	−0.065
p-val.	0.007	<0.001	0.027
Avg. unplugged	2.76	2.19	1.70
$\alpha + \beta + \gamma$	0.001	−0.003	-0.028
p-val.	0.995	0.967	0.222
Avg. plugged	5.07	3.61	2.43

Notes: Houses with no more than 100 wells nearby. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment.

*p<0.1; **p<0.05; ***p<0.01

Table 6: Regression Results: By Lease Agreement, OLS

	Dependent variable:					
	ln(*price*)					
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Drilled Wells	-0.002^{***}	-0.003^{***}	-0.003	-0.003^{***}	-0.004^{***}	-0.004^{*}
α	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Number of Abandoned Wells	-0.007	-0.011	-0.017	-0.006	-0.010	-0.017
β	(0.006)	(0.008)	(0.012)	(0.006)	(0.009)	(0.014)
Number of Plugged Wells	0.007	0.009	0.017	0.005	0.009	0.018
γ	(0.006)	(0.008)	(0.013)	(0.006)	(0.010)	(0.015)
Number Drilled × Lease				0.010^{***}	0.018^{***}	0.026^{**}
α_L				(0.004)	(0.005)	(0.010)
Number of Abandoned Wells × Lease				-0.031^{**}	-0.030	-0.010
β_L				(0.014)	(0.022)	(0.037)
Number of Plugged Wells × Lease				0.022	0.009	-0.015
γ_L				(0.022)	(0.032)	(0.051)
Dist. cutoff	2 km	1.5 km	1 km	2 km	1.5 km	1 km
Observations	266,423	267,212	267,394	266,423	267,212	267,394
Adjusted R^2	0.638	0.637	0.637	0.638	0.637	0.637
$\alpha + \beta$	-0.009	-0.014	-0.020	-0.008	-0.013	-0.021
(p-val.)	0.201	0.256	0.325	0.592	0.556	0.524
Avg. unplugged no lease	1.90	1.67	1.47	1.90	1.67	1.47
$\alpha + \beta + \gamma$	-0.003	-0.005	-0.003	-0.003	-0.005	-0.003
(p-val.)	0.480	0.365	0.917	0.687	0.627	0.993
Avg. plugged no lease	4.18	3.03	2.17	4.18	3.03	2.17
$\alpha_L + \beta_L$				-0.02	-0.013	0.016
(p-val.)				0.614	0.994	0.998
Avg. unplugged w/lease				2.37	1.90	1.44
$\alpha_L + \beta_L + \gamma_L$				0.002	-0.003	0.001
(p-val.)				0.999	0.999	0.999
Avg. plugged w/lease				5.64	3.68	2.41

Notes: Standard errors are in parentheses. Drilled wells are wells for which drilling has been completed—sum of the number of active, unplugged, and plugged wells. Abandoned wells are permanently non-producing and may be plugged or unplugged. Distance cutoff is the maximum distance between a house and a well to be considered treated. The sample includes houses with no more than 100 wells nearby from 25 counties and in the time period between 1995 and 2017.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 7: Regression Results: Houses without Lease Agreements, IV

	Dependent variable:		
	ln(*price*)		
	(1)	(2)	(3)
Number of Drilled Wells	-0.0003	-0.0004	0.006
α	(0.001)	(0.001)	(0.004)
Number of Abandoned Wells	-0.115^{***}	-0.141^{***}	-0.318^{***}
β	(0.030)	(0.036)	(0.087)
Number of Plugged Wells	0.097^{***}	0.117^{***}	0.269^{***}
γ	(0.035)	(0.043)	(0.096)
Dist. cutoff	2 km	1.5 km	1 km
Observations	266,423	267,212	267,394
Adjusted R^2	0.636	0.636	0.635
$\alpha + \beta$	-0.116	-0.142	-0.312
p-val.	<0.001	<0.001	<0.001
Avg. unplugged	1.90	1.67	1.47
$\alpha + \beta + \gamma$	-0.019	-0.025	-0.043
p-val.	0.290	0.274	0.275
Avg. plugged	4.18	3.03	2.17

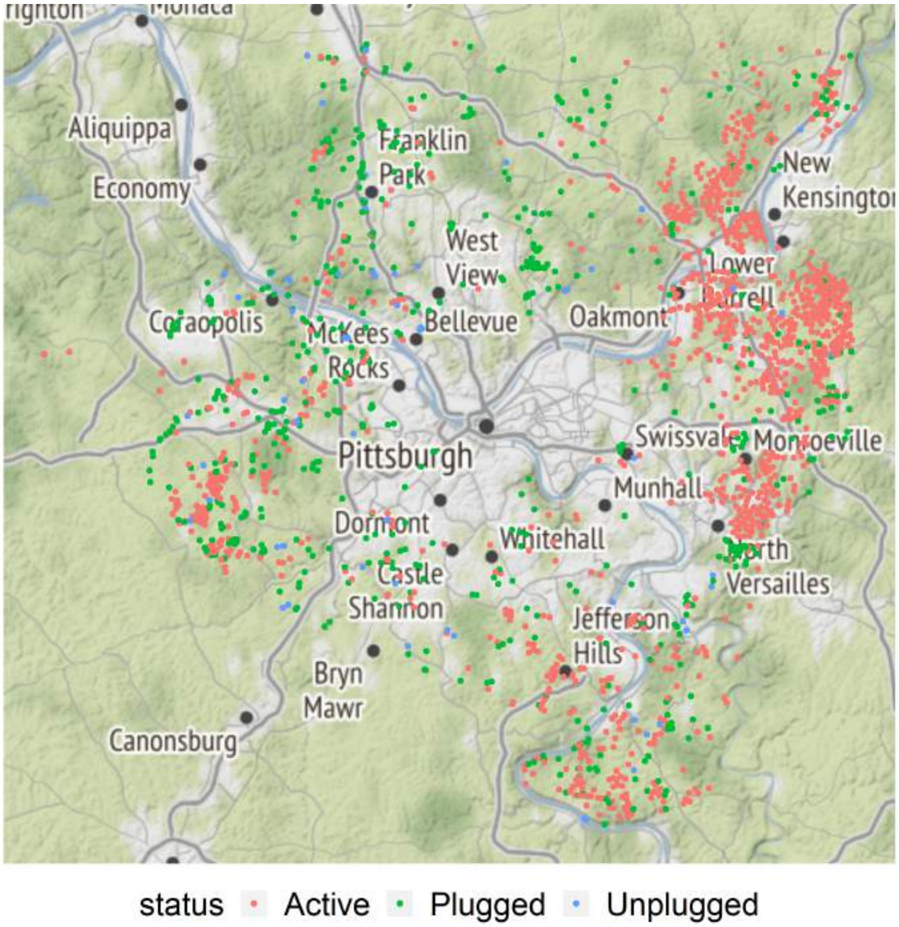
Notes: Houses with no more than 100 wells nearby and do not have an associated lease agreement. The sample includes houses sold between 1995 and 2017. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Appendix

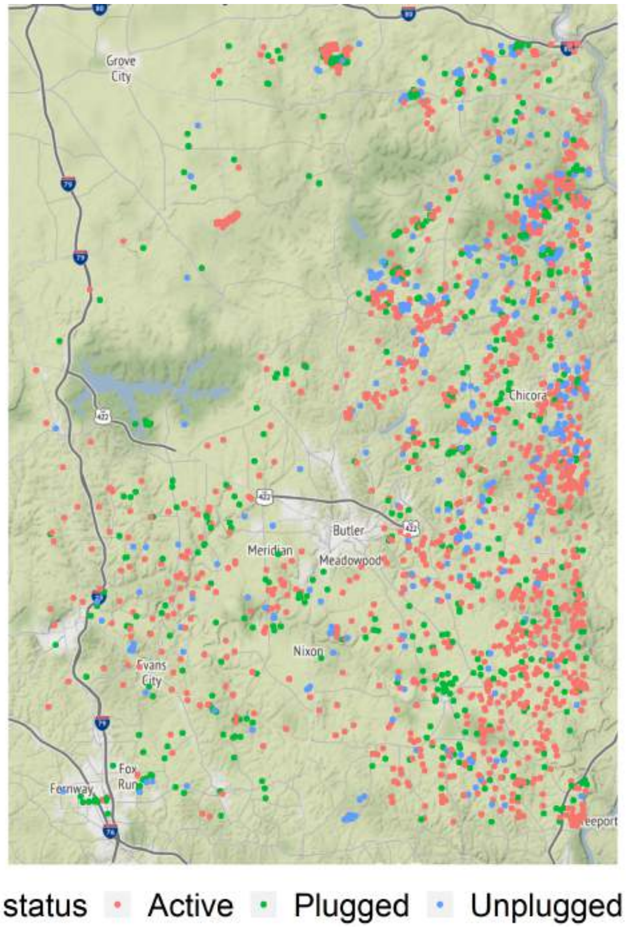
A Additional Tables and Figures

Figure A.1: Drilled Wells in Allegheny County, Pennsylvania, by Status



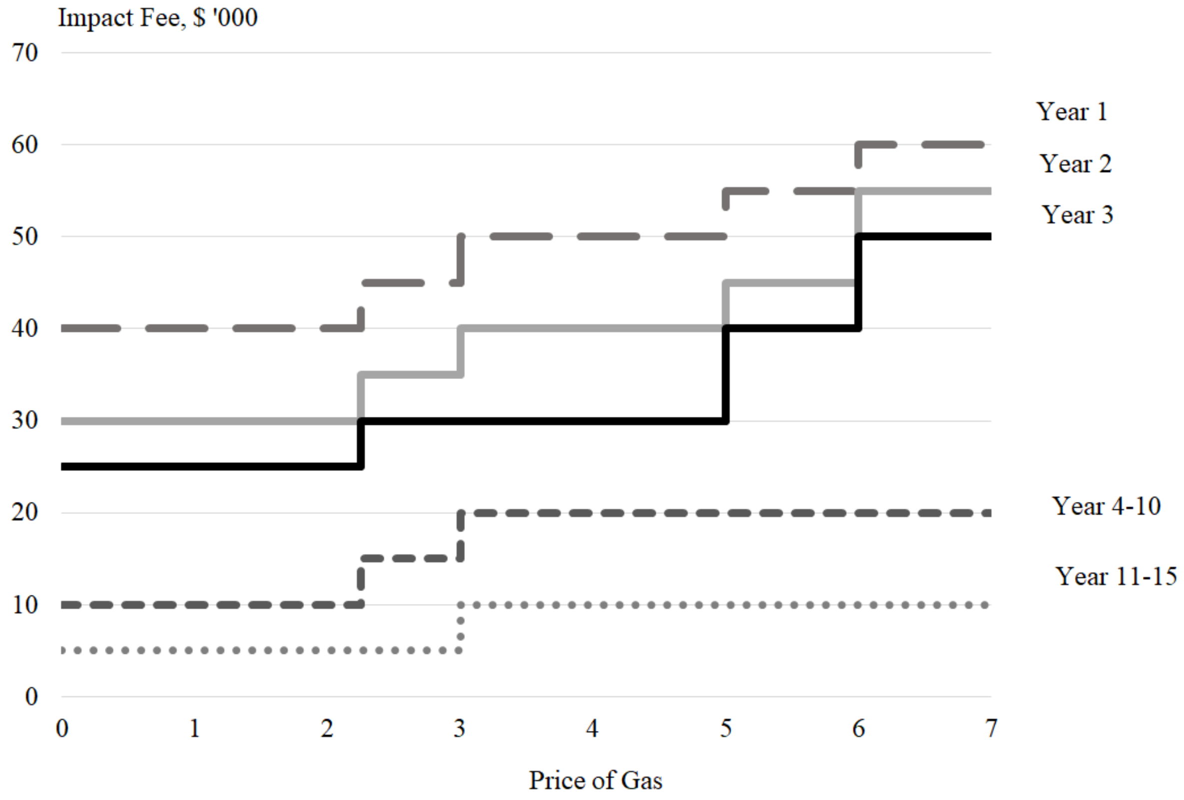
Note: Source of well data: Department of Environmental Protection Spud database. The number of known wells in Allegheny County, PA: 2,012 (February 2019). The location of 40 wells is unknown (not on the map).

Figure A.2: Drilled Wells in Butler County, Pennsylvania, by Status



Note: Source of well data: Department of Environmental Protection Spud database. The number of known wells in Butler County, PA: 3,392 (February 2019). The location of 465 wells is unknown (not on the map).

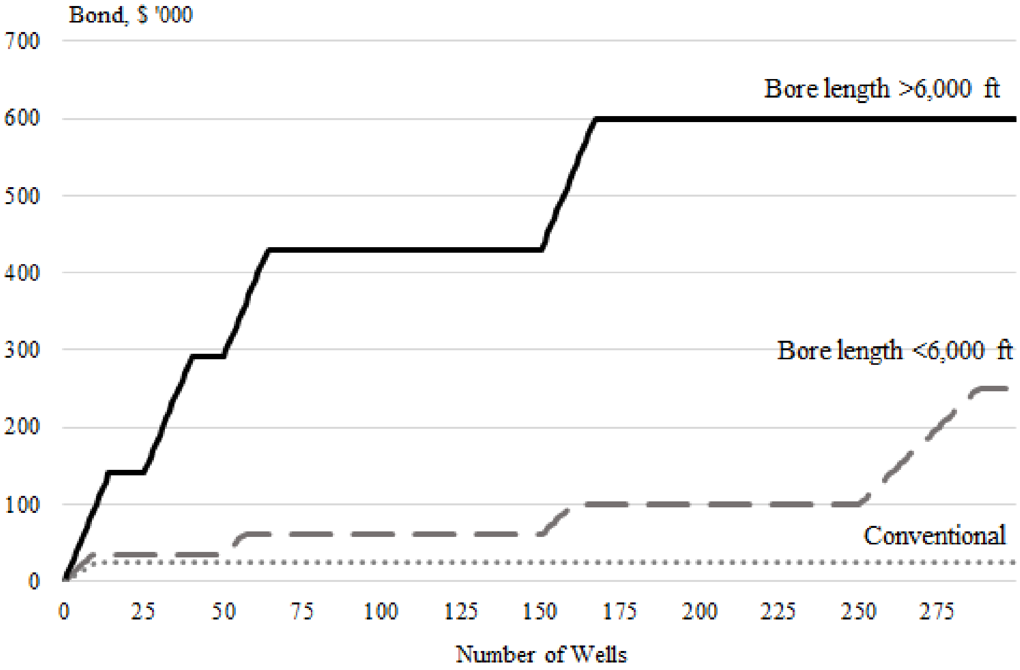
Figure A.3: Impact Fee for Unconventional Wells by Year since Spud Date, 2012 to Present



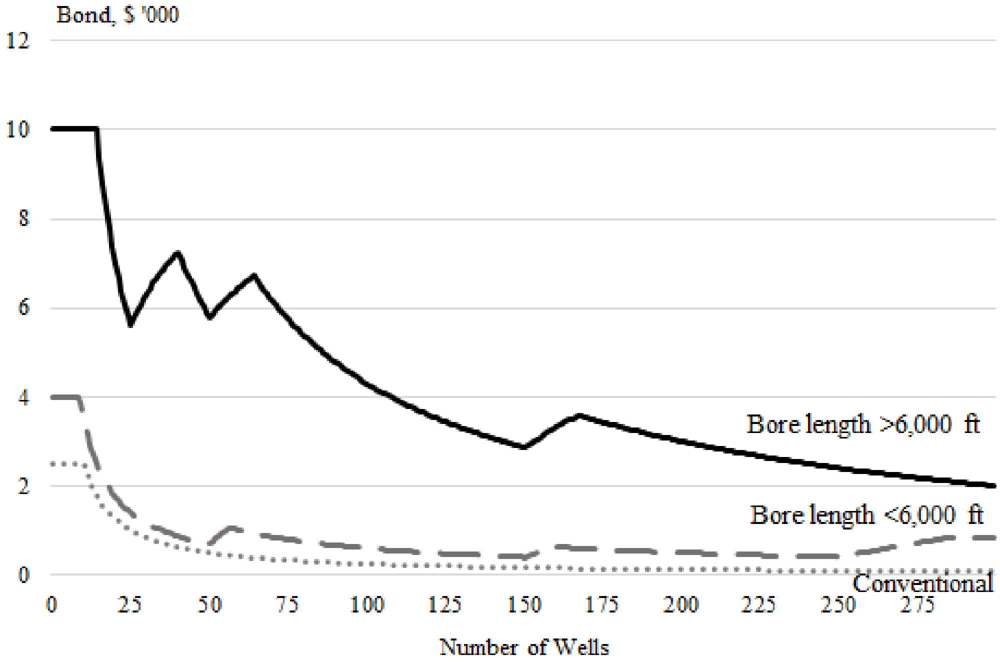
Note: Impact fee is imposed on each producing unconventional well. The amount depends on the price of gas (x-axis) and on the year since spud date (separate lines).

Figure A.4: Bonding Requirements, 2012 to Present

(a) Total



(b) Per Well



Note: Bond is required to be posted as a part of well permit applications since 1984. The PA Act 13 of 2012 introduced new bonding requirements for unconventional wells, while regulation for conventional wells remains unchanged. Starting from 2012, requirements for unconventional wells additionally distinguish those with bore length more and less than 6,000 feet.

Table A.1: Regression Results: Distance Buffer between Control and Treatment

	Dependent variable:					
	log(price)					
	OLS		2SLS		Reduced Form	
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	−0.001	0.0005	0.004**	0.015**	0.016***	0.007***
α	(0.001)	(0.001)	(0.002)	(0.006)	(0.004)	(0.002)
Number Abandoned	−0.007	0.002	−0.094***	−0.285***	−0.086***	−0.071***
β	(0.006)	(0.009)	(0.022)	(0.096)	(0.026)	(0.015)
Number Plugged	0.002	−0.017*	0.056**	0.205**	0.030	0.056**
γ	(0.007)	(0.010)	(0.025)	(0.087)	(0.038)	(0.023)
Observations	395,366	344,440	395,366	344,440	344,440	395,366
Adjusted R^2	0.562	0.558	0.561	0.555	0.558	0.562
Distance cutoff	1.5 km	1 km	1.5 km	1 km	1.5 km	1 km
$\alpha + \beta$	−0.008	−0.003	−0.009	−0.270	−0.064	−0.071
p-val.	0.463	0.986	<0.001	0.008	<0.001	0.017
$\alpha + \beta + \gamma$	−0.006	−0.014	−0.035	−0.064	−0.008	−0.040
p-val.	0.062	<0.001	0.021	0.020	0.778	0.053

Notes: Houses with no more than 100 wells nearby. The pure control group includes houses within 2–20 km of wells. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment. *p<0.1; **p<0.05; ***p<0.01

Table A.2: Regression Results: Variation from Timing of Treatment

	Dependent variable:					
	ln(*price*)					
	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	-0.001	-0.0003	0.002	0.0004	0.004^{**}	0.013^{**}
α	(0.0005)	(0.001)	(0.001)	(0.001)	(0.002)	(0.005)
Number Abandoned	-0.010^{**}	-0.004	0.006	-0.024^{**}	-0.065^{***}	-0.190^{**}
β	(0.005)	(0.006)	(0.009)	(0.011)	(0.020)	(0.084)
Number Plugged	0.012^{**}	0.001	-0.018^{*}	0.016	0.029	0.125^{*}
γ	(0.005)	(0.007)	(0.010)	(0.015)	(0.024)	(0.076)
Dist. cutoff	2 km	1.5 km	1 km	2 km	1.5 km	1 km
Observations	279,767	243,395	192,469	279,767	243,395	192,469
Adjusted R^2	0.573	0.579	0.575	0.573	0.578	0.572
$\alpha + \beta$	-0.011	-0.005	0.008	-0.024	-0.062	-0.177
p-val.	0.051	0.848	0.773	0.083	0.005	0.062
$\alpha + \beta + \gamma$	0.001	-0.003	-0.010	-0.008	-0.033	-0.052
p-val.	0.905	0.563	0.013	0.691	0.046	0.032

Notes: Houses with no more than 100 wells nearby. Only houses treated between 1980 and 2017 are included. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A.3: Regression Results: Smaller Number of Wells

	Dependent variable:					
	log(price)					
	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	-0.002^{***}	-0.003^{***}	-0.002^{**}	-0.002^{***}	-0.003^{***}	-0.003^{**}
α	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number Abandoned	-0.011^{**}	-0.006	-0.008	-0.056^{***}	-0.036^{***}	-0.021
β	(0.005)	(0.005)	(0.006)	(0.012)	(0.012)	(0.014)
Number Plugged	0.014^{**}	0.009	0.012^{*}	0.061^{***}	0.044^{***}	0.031^{**}
γ	(0.006)	(0.006)	(0.007)	(0.013)	(0.013)	(0.016)
Obs.	423,533	408,090	382,617	423,533	408,090	382,617
Adjusted R^2	0.564	0.569	0.577	0.564	0.569	0.577
Max. wells	75	50	25	75	50	25
$\alpha + \beta$	-0.012	-0.11	-0.010	-0.047	-0.078	-0.150
p-val.	0.027	0.055	0.196	0.003	<0.001	0.013
$\alpha + \beta + \gamma$	0.001	-0.001	0.006	0.001	0.008	0.058
p-val.	0.999	0.975	0.992	0.007	0.797	0.017

Notes: Houses with no more than 100 wells nearby. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A.4: Regression Results: Alternative Instruments

	Dependent variable:					
	$\ln(price)$					
	2SLS: $t-5$			2SLS: $t-10$		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	−0.002**	−0.002	−0.001	−0.002	−0.002	−0.003
α	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.006)
Number Abandoned	−0.025**	−0.062***	−0.115**	−0.021**	−0.054***	−0.129**
β	(0.010)	(0.018)	(0.058)	(0.011)	(0.018)	(0.055)
Number Plugged	0.027*	0.057**	0.092	0.017	0.039	0.103*
γ	(0.014)	(0.023)	(0.061)	(0.016)	(0.026)	(0.058)
Dist. cutoff	2 km	1.5 km	1 km	2 km	1.5 km	1 km
Observations	431,766	435,019	436,372	431,766	435,019	436,372
Adjusted R^2	0.560	0.558	0.558	0.560	0.558	0.558
F-stat Number Abandoned	128.8	122.9	87.9	87.94	91.4	50.1
F-stat Number Plugged	98.9	111.9	75.1	53.9	63.1	56.3
$\alpha+\beta$	−0.027	−0.064	−0.117	−0.023	−0.056	−0.132
(p-val.)	0.024	< 0.001	0.112	0.075	<0.004	0.035
$\alpha+\beta+\gamma$	0.000	−0.007	−0.025	−0.005	−0.017	−0.029
(p-val.)	0.999	0.821	0.075	0.904	0.54	0.213

Notes: Notes: Houses with no more than 100 wells nearby. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. The distance cutoff is 2 km.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$ $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A.5: Regression Results: Rural Areas

	Dependent variable:					
	ln(*price*)					
	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Number Drilled	0.001	0.001	0.003^{**}	0.001	0.004^{**}	0.012^{***}
α	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.005)
Number Abandoned	-0.014^{***}	-0.009	0.002	-0.027^{**}	-0.069^{***}	-0.180^{**}
β	(0.005)	(0.007)	(0.011)	(0.011)	(0.020)	(0.075)
Number Plugged	0.015^{**}	0.005	-0.016	0.026^{*}	0.045^{*}	0.128^{*}
γ	(0.006)	(0.008)	(0.012)	(0.015)	(0.023)	(0.070)
Dist. cutoff	2 km	1.5 km	1 km	2 km	1.5 km	1 km
Observations	262,117	265,364	266,717	262,117	265,364	266,717
Adjusted R^2	0.458	0.457	0.457	0.458	0.457	0.456
$\alpha + \beta$	-0.013	-0.007	0.005	-0.026	-0.065	-0.167
(p-val.)	0.028	0.657	0.947	0.048	0.003	0.051
$\alpha + \beta + \gamma$	0.001	-0.002	-0.010	-0.001	-0.020	-0.039
(p-val.)	0.960	0.845	0.021	0.999	0.294	0.073

Notes: Houses with no more than 100 wells nearby, excluding Allegheny County. Standard errors (in parentheses) are clustered at the county-by-year level. House fixed effects, county×year fixed effects, and month of sale fixed effects are included in all columns. Drilled wells include all wells regardless of their status. Abandoned wells have not been producing or have been used in other ways for at least 12 months and may be plugged or unplugged. A well is plugged when environmental and plugging requirements are satisfied. Distance cutoff is the maximum distance between a house and a well to be considered treated. Average numbers of unplugged and plugged wells are conditional on treatment.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

B Choice of distance cutoffs to define treatment

The distance cutoff to determine the treatment and the control group is chosen based on several factors. First, other studies, for example, Muehlenbachs et al. (2015), use similar distance cutoffs, which ensures the comparability of the results. Second, the choice of the distances is supported by numerous other studies, summarized in Table A.6. This table describes at what distance the effect of well proximity was identified.

There is no consensus in the literature about the setback distances to adequately protect public health and ensure safety. Most studies find that health outcomes, pollution, and other risks are most significant within approximately 0.5–2.5 km. Additionally, those studies find that these risks are higher within close proximity of well sites.

In the US, many states require wells to be drilled with at least 200–1,000 ft (61–304 m). In Colorado, for example, it was proposed to increase the setback distance to 2,500 ft (762 m), but this proposal failed to pass in a ballot referendum. Pennsylvania law does not allow conventional wells to be drilled closer than 200 ft (≈61 m) and unconventional wells closer than 500 ft (≈152 m). However, even these rather lenient setback requirements can be waived. The US Department of Housing and Urban Development (HUD) requires at least 300 ft (≈91 m). These setback distances are often viewed as inadequate, and many recommend the setback distance between 0.25 and 2 miles (0.4–3.2 km) (Lewis et al., 2018). However, the cost of requiring the setback distance of 2,500 ft (762 m) is high, and health benefits should be quite large to justify it (Ericson et al., 2020).

Table A.6: Identified Effects of Gas and Oil Extraction Proximity by Distance to Wells, Summary

Study	Outcome	Effect Identified within
Hill (2012)	Infant health	2.5 km
Haley (2016)	Evacuation due to blowout Sulfur pollution	0.8-4 km (0.5-2.5 mi) 0.43-3.6 km (0.3-2 mi)
Kassotis et al. (2014)	Elevated levels of endocrine disrupting chemicals in water sources	1.6 km (1 mile)
Rabinowitz et al. (2015)	Respiratory and dermal conditions	1-2 km
McKenzie et al. (2012)	Pediatric sub-chronic non-cancer hazard and chronic hazard indices	<800 m

C Correlation of the instrumental variables with other outcomes

The baseline identification uses a shift-share IV. The shift components of the instrument are shocks to well abandonment and plugging specific to well types by resource extracted (oil, gas, oil and gas, coalbed methane) and configuration (conventional or unconventional). The identification is based on the assumption that these shocks are exogenous to the local economic conditions. Moreover, exclusion restriction requires that the IV affects the outcome only through the exogenous variable of interest. The concern is the instrument might pick up changes in the macroeconomic environment that are correlated with other outcomes. For example, low gas and oil prices may simultaneously drive up well abandonment intensity and unemployment rates and reduce household income in a way that is not absorbed by county-year fixed effects in the main specification. That would be worrisome, as this would violate the exclusion restriction. While it is impossible to test directly if these assumptions are satisfied, correlation between the instrument and other economic variables provides suggestive evidence.

I collect data at the census tract level using the National Historical Geographic Information System (NHGIS) (Manson et al., 2019). I use the data from the 1980, 1990, and 2000 decennial censuses and 2008–2012 and 2013–2017 American Community Surveys (ACS).[37] I construct the following outcome variables: percent of population (at least 25 years old) with a college degree, employment to population rate, unemployment to population rate, natural logarithm of median income, and percent of population below poverty line. All variables, except median income, are in percentages.

Well data are aggregated at the census tract level. Wells within 2 km of census tract boundaries are also assigned to that tract.[38] To examine how well the instruments are correlated with these variables, I estimate the following regressions:

$$y_{tract,t} = \alpha D_{tract,t} + \beta \hat{A}_{tract,t} + \gamma \hat{P}_{tract,t} + \mu_{county,t} + \delta_{tract} + \epsilon_{tract,t}, \tag{A.1}$$

where $y_{tract,t}$ is the census tract level outcome at time t, $D_{tract,t}$ is the number of drilled wells, and $\hat{A}_{tract,t}$ and $\hat{P}_{tract,t}$ are the predicted number of abandoned and plugged wells related to that census tract. To be consistent with the main specification (1), the regression includes county×year fixed effects, $\mu_{county,t}$, and census tract fixed effects, δ_{tract}. Standard errors are clustered at the county-by-year level.

Table A.7 shows the estimation results of the specification (A.1) on the full sample. Table A.8 differs by the sample used. To be consistent with the baseline, census tracts with more than 500 wells are excluded, which is approximately 2.5% of the observations. None of the variables are significant at the 5% level. The predicted number of abandoned wells could be associated with the lower unemployment rate, and this coefficient is significant at the 10% level. Regardless of the statistical significance, all the coefficients are very small in the magnitude and are therefore *economically* insignificant. Correlation between the predicted number of abandoned wells and the employment rate is the most concerning. However, the average predicted number of abandoned wells (both plugged and unplugged) is 4.6 per census tract and 10.5 conditional on having at least one abandoned well. That is, on average, the decrease in employment is less than 1%. At the same time, there is no corresponding change in the unemployment rate and an increase in the share of people with a college

---

[37] To consistently define census tracts, I use Census Tract Relationship Files, retrieved from ICPSR and the US Census Bureau website.

[38] The results do not change significantly if only wells that are within census tract boundaries assigned to the tract

education. Taken together, the effect is (if any) at the labor market participation margin. Such a small change in the employment-to-population ratio is unlikely to explain the large decrease in home prices near abandoned wells. I therefore conclude that the variation in the macroeconomic factors, which could potentially be correlated with the instruments, is absorbed by the county-by-year fixed effects.

Table A.7: Correlation between Instruments and Census Tract Characteristics, Full Sample

	Dependent variable:				
	Percent college (1)	Employment rate (2)	Unemployment rate (3)	Median income, logarithm (4)	Percent below poverty line (5)
Number of Drilled Welld	−0.001	−0.001	−0.0005	−0.0001	0.002
	(0.001)	(0.001)	(0.001)	(0.0001)	(0.003)
Predicted Number of Abandoned Wells	0.008	−0.006	0.002	−0.0002	0.006
	(0.006)	(0.006)	(0.003)	(0.0004)	(0.010)
Predicted Number of Plugged Wells	−0.008	0.011*	−0.004	0.001	−0.037*
	(0.006)	(0.006)	(0.004)	(0.001)	(0.021)
Observations	5,484	5,485	5,485	5,478	5,482
Adjusted R^2	0.914	0.759	0.433	0.584	0.802

Notes: Standard errors (in parentheses) are clustered at the county-by-year level. Columns (1)–(3) and (5) are in percentages. The predicted number of abandoned wells is the instrument from (10a). The predicted number of plugged wells is the instrument from (11a). All columns include county-by-year fixed effects and census tract fixed effects.
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A.8: Correlation between Instruments and Census Tract Characteristics, $\leq$500 Wells

	Dependent variable:				
	Percent college (1)	Employment rate (2)	Unemployment rate (3)	Median income, logarithm (4)	Percent below poverty line (5)
Number of Drilled Wells	−0.010	0.011	−0.002	0.00003	0.015*
	(0.006)	(0.008)	(0.002)	(0.0005)	(0.009)
Predicted Number of Abandoned Wells	0.077*	−0.092*	0.0003	−0.005	0.063
	(0.045)	(0.054)	(0.014)	(0.005)	(0.045)
Predicted Number of Plugged Wells	−0.054	0.050	−0.005	0.009	−0.234***
	(0.063)	(0.075)	(0.017)	(0.006)	(0.068)
Observations	5,824	5,825	5,825	5,818	5,822
Adjusted R^2	0.914	0.758	0.430	0.583	0.806

Notes: Standard errors (in parentheses) are clustered at the county-by-year level. Columns (1)–(3) and (5) are in percentages. The predicted number of abandoned wells is the instrument from (10a). The predicted number of plugged wells is the instrument from (11a). All columns include county-by-year fixed effects and census tract fixed effects.
*p<0.1; **p<0.05; ***p<0.01.