NBER WORKING PAPER SERIES

NEGOTIATIONS OF OIL AND GAS AUXILIARY LEASE CLAUSES: EVIDENCE FROM PENNSYLVANIA'S MARCELLUS SHALE

Max Harleman
Pramod Manohar
Elaine L. Hill

Working Paper 30806
http://www.nber.org/papers/w30806

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2022

We would like to thank Ashley Vissing for providing optical character recognition (OCR) Python and Stata code which we adapted to conduct this study. Thanks to National Institutes of Health for partially funding this work under Grant Number: DP5OD021338 (PI:Hill). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research or NIH.

Negotiations of Oil and Gas Auxiliary Lease Clauses: Evidence from Pennsylvania's Marcellus Shale
Max Harleman, Pramod Manohar, and Elaine L. Hill
NBER Working Paper No. 30806
December 2022
JEL No. D23,Q32,Q33,R11

ABSTRACT

Oil and gas lease negotiations provide mineral owners the opportunity to negotiate for both compensation and auxiliary clauses that may protect their health and properties. We use optical character recognition to assemble a novel dataset of compensation and specific clauses in nearly 60,000 leases signed in the Marcellus Shale Play of Pennsylvania. We leverage the dataset to produce three main findings. First, contrary to the standard utility maximization model, we find a positive relationship between compensation and clauses. Second, we find that as development of the shale play progressed over time, compensation rose and leases became more likely to contain environmentally protective clauses. Third, we find that compensation and the presence of clauses have a weak relationship with the geologic productivity of nearby wells. Together, our findings indicate that oil and gas firms simultaneously make concessions by raising compensation and approving clauses, but these concessions do not depend on geologic productivity. This suggests that some mineral owners, such as those that are high-income or from more socially organized communities, have the skills or resources to negotiate for more favorable leases all-around and point to similar environmental justice concerns identified in other shale plays.

Max Harleman
Department of Government and Sociology
Georgia College and State University
410 West Greene Street Campus Box 18
Milledgeville, GA 31061
max.harleman@gcsu.edu

Pramod Manohar
Department of Economics
University of Rochester
Rochester, NY 14642
pmanoha2@u.rochester.edu

Elaine L. Hill
Department of Public Health Sciences
Department of Economics
University of Rochester
265 Crittenden Boulevard
Rochester, NY 14642
and NBER
elaine_hill@urmc.rochester.edu

A data appendix is available at http://www.nber.org/data-appendix/w30806

1 Introduction

The shale oil and natural gas boom has led to income and employment gains in the United States (Hausman and Kellogg, 2015), but it has led to social and ecological disruptions in communities that host drilling (Mason et al., 2015). Both public and private institutions shape the terms of oil and gas extraction in ways that may protect local communities from these disruptions. In the United States, state laws and regulations govern oil and gas extraction by determining where and how companies can drill wells, how they dispose of solid and liquid byproducts, and how they must restore well sites when production stops (Richardson et al., 2013). Private mineral owners can also regulate the industry through a contractual lease agreement that provides the oil and gas company rights to develop the mineral owner's estate.

Each oil and gas lease has two components. First, primary clauses identify the mineral owner, and define the geographic extent of the mineral estate, the duration of the lease, and how the mineral owner will be compensated. Owners of estates where drilling occurs are compensated with one-time bonus payments, and also receive royalties, which are calculated as a percentage of the value of oil or natural gas produced from the well. Second, leases contain auxiliary clauses that shape the terms of development, some of which may protect the health, safety, or aesthetics of the local community. For example, an auxiliary clause may require the company to place a fence around the well site, or to test nearby groundwater sources before and after drilling and provide replacement water if the tests show that groundwater has been contaminated.

We leverage data on auxiliary clauses in Pennsylvania's Marcellus Shale. Several empirical studies of the economic impacts of shale drilling leverage data on lease locations and royalty rates, which are available from private data providers (e.g., Brown et al. (2016, 2019); Harleman and Weber (2017); Ikonnikova et al. (2015)). But only two have leveraged data on auxiliary clauses, and they focus on just one county in Texas (Vissing, 2015; Timmins and Vissing, 2022). The barrier to using auxiliary clause data is that leases are filed at county courthouses as scanned images, which are not machine-readable. In this paper, we use optical character recognition (OCR) to identify the auxiliary clauses within each lease. OCR is a type of computer program that converts images of text documents into machine-readable data.

Our OCR data allows us to examine the prevalence of various clause types across nearly an entire shale play. By connecting the OCR data to lease locations and royalty rates, we explore the determinants of oil and gas leasing outcomes. Specifically, we test the hypothesis that mineral owners treat royalty rates and auxiliary clauses as substitutes when negotiating with oil and gas companies over leasing terms. The standard utility maximization model under efficient Coasean bargaining envisions mineral owners making trade-offs between negotiating for royalties that provide them money and negotiating for clauses that may protect their health, privacy, or enjoyment of their property. If leases provided greater royalty income, mineral owners may be willing to forgo clauses that make drilling and operating wells more costly for oil and gas firms. We find that no such trade-off exists between health-protective clauses and royalty rates. Instead, we find a positive relationship between most clause types and royalty rates, suggesting that some mineral owners have the skills, knowledge, or financial resources to negotiate for leases that are more favorable all-around, while others do not.

If there is no trade-off between clauses and royalty rates, what explains their variation across time and space? We find that 99 percent of the 58,559 leases in our sample have at least one auxiliary clause. But the share of leases that contain any one of 36 unique clause types ranges from less than one percent to 80 percent. One reason that the presence of clauses may vary is that mineral owners learn about the risks of the industry over time. There are numerous examples of consumers learning about health and safety risks through the media, advertising, or word of mouth and making investments to protect themselves, including paying for airbags and seat-belts in cars (Mannering and Winston, 1995; Arnould and Grabowski, 1981), and asking their doctor to prescribe cholesterol-lowering drugs (Calfee et al., 2002). Another explanation is that firms may become less environmentally intensive and therefore willing to include more protective clauses. Consistent with these explanations, we show that as widespread development of the Marcellus Shale progressed over time leases became more likely to contain environmentally protective clauses. Specifically, leases signed later in the study period are more likely to contain clauses that require oil and gas companies to test their water prior to drilling and provide replacement water in the event of contamination, as well as clauses that allow extraction of the oil and natural gas from the mineral estate but no access to the surface for placing well pads or other infrastructure.

Royalty rates also evolve over time, with the average rate remaining constant at the state-mandated minimum of 12.5 percent for leases signed from 2001 to 2007, and increasing to a maximum of 16.8 percent for leases signed in 2010 (just after the commercial potential of the play was realized and widespread drilling took off in 2009). In addition to varying over time, royalty rates may vary across space if oil and gas companies offer greater compensation in areas with greater geologic productivity. We use oil and gas production reports to estimate the first year production of each well, which is the most important predictor of a well's ultimate recovery. We use two measures of geologic productivity: average first year production of all wells within a 2 kilometer (km) radius of a lease, and average first year production of wells that were drilled within a 3 km radius of a lease before the lease was signed. We find that a doubling of the average first year production of these nearby wells leads to only a 1 to 2.2 percent increase in the royalty rate. For the average royalty rate of 14 percent in our sample of leases, this is a 1 to 2 percentage point increase in the share of the value of production that goes to the mineral owner.

We identify a weak relationship between geologic productivity and royalty rates, which is consistent with Brown et al. (2016). They show that low pass-through can be explained by firms exercising market power and uncertainty about the valuation of resource endowments. By using first year production of wells located in close proximity to a lease, we confirm their results which are estimates of the relationship between average royalty rates and average expected ultimate recovery estimates at the county level. Our more localized approach rules out that their finding of limited pass-through at the county level masks relationships between royalty rates and geologic productivity in many highly-productive "sweet spots" that occur within counties.

Our findings indicate that oil and gas firms simultaneously make concessions by raising royalty rates and approving auxiliary clauses, but that these concessions are not contingent upon or proportionate to the productivity of the mineral resource. This means that some mineral owners are able to acquire atypically favorable leases due to their knowledge or negotiating power. It appears in part that this knowledge is gained by mineral owners who sign leases later into the development of the shale play because they can learn about the potential risks of drilling and include stronger auxiliary clauses. But it does not appear that mineral owners leverage similar knowledge about the productivity of wells on neighboring properties to capture greater royalties. We find that the productivity of wells drilled on nearby properties before the lease was signed bears only a weak relationship with the royalty rate.

Mineral owners with greater information or financial and legal resources to negotiate favorable leases are likely those that are high-income and from more socially organized communities. In our dataset, we cannot distinguish between mineral owners that hold out until later in the study period to learn more about the risks and geologic potential of the shale play from those that are simply not offered a lease until later in the study period. If holding out happens, high-income individuals with a lower marginal utility of bonus and royalty income are more likely to hold out to gain more knowledge about the risks of the industry and the geologic potential of the shale play as a whole. These implications point to similar environmental justice concerns in oil and gas leasing markets as those identified by Vissing (2015) and Timmins and Vissing (2022).

2 Production and Leasing in Pennsylvania's Marcellus Shale

The Marcellus Shale is one of the major shale plays in the United States, and stretches across upstate New York, Pennsylvania, West Virginia, and eastern Ohio. According to the United States Geological Survey, the Marcellus Shale contains approximately 84 trillion cubic feet of technically recoverable natural gas, making it one of the largest shale plays in the United States (US Geological Survey, 2022). The first Marcellus well was drilled in 2002, and widespread commercial drilling began to pick up in Pennsylvania in 2007 (Figure 1). Drilling reached its peak in 2011, and by that year nearly 11,000 square miles have been leased to oil or gas companies for drilling (US Energy Information Administration, 2011). Natural gas production in Pennsylvania's Marcellus has grown each year since 2009, and in 2020 represented 27 percent of domestic shale gas production, second only to Texas (US Energy Information Administration, 2022).

In a rush to secure valuable mineral rights from mineral owners, our data indicate that firms signed around 70 thousand shale oil and gas leases in the Marcellus shale over our study period of 2001 to 2016. Oil and gas leases consist of a set of primary clauses and auxiliary clauses, which are negotiable between the mineral owner and the oil or gas company. Primary clauses are contained within every lease and are made up of a careful description of the minerals being leased, information about royalty and bonus payments paid to the lessor, the span of time covered by the lease, and opportunities for a lease extension. Brown et al. (2019) estimate that payments to shale oil and gas mineral owners have had a substantial impact on private income, amounting to over $66 billion annually in direct and induced private income. Importantly, leases grant oil and gas firms the sole right, but not the obligation to drill for oil and gas over some fixed time frame. If firms drill at least one well that produces oil or natural gas, most leases remain in effect until production ceases. Auxiliary clauses are optional and contain specific language that protects one or both parties. Most of these clauses aim to prevent or remediate possible soil, water, or aesthetic damages on the surface of the lease and on nearby properties. For instance, leases may include surface damage clauses that ensure that the firm must restore the surface of the drilling site back to its original state after drilling is finished.

State and federal regulations govern oil and gas development, but mineral owners have the broadest and most direct ability to shape the terms of oil and gas development and protect themselves from its environmental and public health effects by negotiating favorable auxiliary clauses. The effects on the environment and public health center on water and air contamination. Water quality concerns are related to both surface water and groundwater (Hill and Ma, 2017, 2022; Bonetti et al., 2021). Air quality concerns include elevated levels of particulate matter (Zhang et al., 2023), volatile organic compounds and methane in the air near wells, pipelines, and compressor stations (McKenzie et al., 2012; Macey et al., 2014; Atherton et al., 2017; Payne et al., 2017; Zimmerle et al., 2017). Although the exact causal mechanism is unclear, studies show that these environmental changes lead to declines in health outcomes for individuals living near shale development (Hill, 2018). Black et al. (2021) provide a detailed review of recent economic literature on the environmental, economic, and health implications of shale gas development.

3 Determinants of Lease Negotiation Outcomes

In this section we review four distinct bodies of literature to form predictions about oil and gas leasing outcomes. First, we explore the traditional Coasean bargaining framework which predicts that mineral owners will make trade-offs between negotiating for lease clauses and royalties. Second, we explore a general literature on how firms and individuals learn about and take measures to prevent risk, and a narrower qualitative literature that documents communities learning about the localized impacts of shale development to predict that mineral owners will adopt more protective auxiliary clauses over time. Third, we explore the intuitive predictions that owners of more productive mineral resources will be able to negotiate for higher royalty rates and greater clauses, and that royalty rates will increase over time as unleased acreage grows scarce. We also review recent findings in the nonrenewable resource literature on why these intuitive predictions may be false.

3.1 Coasean Bargaining for Payments and Health Protective Clauses

Coase (1960) states that when there are conflicting property rights, an efficient outcome can be reached after the two parties bargain if the property rights can be well defined, there are zero transaction costs, and the parties are symmetrically informed. A conflict occurs when one party infringes on the other party's rights, such as in the case of a rancher allowing his cattle to graze in a neighboring farmer's wheat field. In this case and in absence of a government enforcing the farmer's property rights, efficiency could be achieved if the farmer paid the rancher to keep the cattle out of his field. More specifically, efficiency depends on a payment that leads to an outcome where the cost of the damaged wheat caused by the last itinerant cow equals the benefit to the rancher of allowing that cow to roam.

In the case of the oil and gas industry in Pennsylvania, there are rarely conflicts over mineral rights themselves. Most mineral resources are owned by private mineral owners, and leases compensate them with payments in the form of royalties and bonuses. More common are conflicts over rights to clean air and water, unmarred scenic landscapes, or disruptions to the peace and quiet of rural communities. Although government regulations attempt to prevent some of these disruptions (i.e., regulations on how wells are constructed are intended to prevent groundwater contamination), others are not addressed bureaucratically. For instance, there are no state-mandated rules about whether loud machinery can be operated at night, whether a tall fence is placed around the well pad, or whether pipelines can be placed on the mineral owner's property. Auxiliary clauses are the primary mechanism for mineral owners to protect themselves (and their neighbors) from unregulated disruptions.

Leases will vary based on the protective clauses that they include and the payments to mineral owners. Consider Figure 2 Panel A, in which the buyer (in this case the oil and gas company) offers the seller (the mineral owner) a lease that includes $P1$ in payments and clauses that prevent $C1$ dollars of damages. If the seller finds this amount acceptable, they will sign the lease.

But perhaps the seller responds to the offer by counter-offering a lease that contains more protective clauses. In Panel B, we rotate the seller's utility curve $S2$ to correspond to a seller that has a greater preference for clauses, relative to curve $S1$, and keep the buyer's offer curve constant. The value of clauses increases to $C2$. Imagine that Panel A represents an efficient outcome, in which the firm is paying the equilibrium price for extracting the minerals and the clauses protect the environment so that the marginal benefit of causing environmental damage equals the marginal cost of the damage. If this were the case, the firm would be unwilling to add more clauses without reducing payments. The buyer curve $B1$ remains fixed, and payments fall from $P1$ to $P2$.

Panels A and B of Figure 2 reveal that with efficient Coasean bargaining the standard utility maximization model predicts mineral owners making trade-offs between clauses and payments. This motivates our first hypothesis, that within a given year royalty rates and the presence of auxiliary clauses that are protective of the environment and human health will be negatively associated. A positive association between royalty rates and clauses would indicate that certain mineral owners have greater ability to negotiate, or greater information about either the impacts of the shale industry, the array of clauses they could potentially include, or the net benefits of drilling experienced by the firm. Panel C of Figure 2 illustrates a case where sellers that are more informed about the industry or the value of their mineral resources, and seek to negotiate for greater payments and more clauses by asking for the combination $(P3, C3)$. If the buyer expects positive profits at this point, they will accept the offer, which reveals that a more informed set of sellers, or those in areas with more productive resources, may be able to simultaneously negotiate for more clauses and higher payments.

3.2 Learning about Industrial Innovation

Several studies show that as firms gained experience with shale oil and gas development, they learned to select more profitable drilling technologies and locations (Covert, 2015; Fitzgerald, 2015; Fetter et al., 2018; Agerton, 2020; Levitt, 2011). Others show that since the start of the shale boom, firms have become more environmentally responsible as measured by recorded violations per inspection by an environmental regulator (Kim and Oliver, 2017). This improved environmental performance is perhaps caused by firms learning to apply safer practices, but could also be due to stronger regulations and leases (Rahm et al., 2015), or the once diffuse industry consolidating into a smaller number of larger firms.

Just as firms learn over time, there are numerous examples of individuals learning to protect themselves from dangerous products or services. For example, Mannering and Winston (1995) document the rapid adoption of airbags in automobiles in the 1990s, and show that consumers became increasingly willing to pay for airbags as they learn of their life-saving abilities through media coverage and word of mouth. This is analogous to Panel B of figure 2—as drivers demand more safety created by airbags, they are willing to pay more for vehicles in exchange for this increased safety. But there are also cases where consumers are slow to learn. For instance, in the early 1980s, seat belts were used by less than 20 percent of automobile occupants, despite their life-saving abilities (Arnould and Grabowski, 1981). Further, direct consumer advertising of the statin class of cholesterol-reducing drugs did not have significant effects on new statin prescriptions, despite the proven efficacy of the drugs (Calfee et al., 2002).

Much qualitative research documents that through exposure to shale development, affected residents are learning about specific industry practices and health concerns (e.g., Brasier et al. (2015); McElroy et al. (2020)). For example, Sangaramoorthy et al. (2016) interviewed residents who developed concerns about the health impacts of shale development by witnessing their neighbors experience psychological distress, nosebleeds, sore throats, rashes, asthma, and headaches. Furthermore, their interview data reveal that the residents associate health concerns with specific industry practices and infrastructure, such as compressor stations, storage and condensation tanks, and truck traffic. Outside of the spread of information in affected communities by word of mouth, media and internet coverage of the impacts of shale development and "fracking" soared in the late 2010s. These qualitative findings motivate our second hypothesis, that mineral owners will negotiate for more protective lease clauses in later years of the development of a shale play.

We also explore the evolution of royalty rates over time as the shale play is developed. Intuitively, would expect royalty rates to rise over our study period. In 2001 modern shale production technologies were just beginning to be developed and expectations and knowledge about the productivity of the Marcellus shale were low. Widespread adoption of the technologies and more competition among firms to lease acreage in areas where early drilling proved most productive should translate into rising royalty rates over the study period. Nevertheless, Agerton (2020) present an alternative hypothesis that royalties may increase initially as unleased acreage becomes scarce, but decrease later in the development of a play because firms must undergo costly search to locate holdout mineral owners. We follow the more intuitive explanation to motivate our third hypothesis, that royalty rates will rise over the study period. If this hypothesis holds and like Brown et al. (2016) we fail to find that owners of more productive mineral resources are able to negotiate for higher royalty rates, improved expectations and knowledge about the profitability of the shale would be the primary driver of growing royalty rates.

3.3 Natural Resource Productivity and Payments to Mineral Owners

There is significant heterogeneity in resource abundance across space, including within similar formations (Ikonnikova et al., 2015). As a consequence of spatial variation in resource abundance, some locations are more profitable to drill from than others. Mineral owners in more profitable areas may thus capture more rents than their counterparts in less profitable areas, in the form of higher royalty payments and a larger presence of lease clauses. Panel C of figure 2 illustrates that as drilling becomes more profitable, oil and gas firms should sign leases with higher royalty payments and more auxiliary clauses. This theoretical expectation motivates our fourth hypothesis, that owners of more productive mineral resources will be able to negotiate for higher royalty rates and greater clauses.

Despite this theoretical expectation, Brown et al. (2016) find limited pass-through of resource abundance into royalty rates. Specifically, they find that a doubling of the expected ultimate recovery of the average well in a county increases royalties paid to mineral owners by between 1 and 2 percent. They create an economic model that predicts that in perfect competition (with free entry into leasing markets among firms with equivalent information about geologic productivity) a one percent increase in resource abundance will lead to a one percent increase in royalty rate paid to mineral owners. They provide two explanations for their empirical finding of pass-through that is much lower than the perfectly competitive scenario. First, a single firm may benefit from first-mover advantages or spatial economies of scale in a given geographic area. In such a monopsonistic market with an upward-sloping supply of mineral rights, the monopsonist offers low royalty rates because increasing the rate allows it to acquire too few additional leases to compensate for offering infra-marginal owners a higher rate. Second, leases are signed before drilling begins but remain in effect until production ends, meaning that a mineral owner cannot renegotiate a higher royalty rate if the resource proves to be productive. Uncertainty about the richness of mineral resources means that firms with superior geologic information can offer non-renegotiable royalty rates that are lower than in markets with perfect competition and symmetrical information across firms, as well as between firms and mineral owners.

While Brown et al. (2016) estimate pass-through of geologic productivity to county average royalty rates, we estimate pass-through at a finer scale. Our data enables us to assess the relationship between royalty rates and the average production of all wells within a 2 km radius of the lease. Our approach accounts for the fact that there is significant variation in geologic productivity within counties, and observing no relationship between production and royalties at the county level may mask stronger relationships between these variables within counties. Our more disaggregated data also allow us to test whether greater production of wells drilled *before* the lease was signed and within a nearby radius has an effect on royalty rates. Over time, both firms and mineral owners may learn about the productivity in a given geographic area and the royalty rates offered in that area. Mineral owners may be able to leverage this information to negotiate higher royalty rates and stronger clauses. Thus, we expect leases that are near highly productive wells and are signed after those wells are drilled should display greater pass-through of localized productivity than leases with less productive wells nearby.

4 Data

4.1 Leases

Our analysis leverages 185,968 lease documents which were filed and scanned at county courthouses in Pennsylvania. The private data provider DrillingInfo makes these documents available as scanned PDF and image documents, which we purchased in 2017. To identify the universe of clauses within the leases, we read through lease documents to saturate a list of clauses and variations of the language used by each drilling company to describe the clauses. Altogether, we identified a list of 36 unique clauses, which were expressed with 360 unique regular expressions. A regular expression is a sequence of symbols and characters that expresses patterns that can be used to search for the clauses in each lease document. The number of regular expressions is significantly greater than the number of unique clause types because a single clause type can be expressed in several different ways. For example, a clause that an oil and gas firm cannot inject liquid waste into the ground on the leased property could be stated as "no water injection" by one firm, "no water disposal" by another, and "not allow for any injection well" by yet another. We selected at random and read through 275 leases before saturating a list of regular expressions associated with the 36 clause types. Reading an additional 25 for a total of 300 read leases yielded no further clause types or regular expressions.

Next, we used the optical character recognition (OCR) tool Tesseract to obtain text files that contain the contents of each lease. Of the 185,968 documents, Tesseract was successfully able to "read" and create text documents for 172,616. We employed regular expression analysis in Python which searched for the presence of our 360 expressions in the text files. The regular expression analysis created a dataset containing all of the 172,616 documents as rows, with 36 binary variables indicating whether each unique clause type was present in the document.

DrillingInfo also provides machine-readable data on lease polygons (the geospatial boundaries of each mineral lease), and also contains information about the document type, the date it was signed, and the royalty rate of each lease. We merged 148,889 scanned documents with the polygon data, with the unmatched leases largely being due to some counties not being reflected in the polygon data.[1] Of the matched documents, we keep those that meet three inclusion criteria. First, we keep 63,082 that are the "lease" document type, and drop lease amendments, extensions, and memos, which do not contain complete information on auxiliary clauses and royalty rates. Second, we drop 4,265 leases with less than the state minimum royalty rate of 12.5 percent and greater than the 99th percentile of 25 percent, because we believe leases with rates outside of this range represent data entry errors. Third, we drop 258 leases that were signed before the year 2001, which predates leasing for shale development, and those signed after 2016, because we have incomplete data for 2017, the final year covered by the scanned records. We are left with 58,559 leases, which represents our full sample.

In Table 1, we display the number of leases with each clause and without each clause. Of the 58,559 leases total leases, 99 percent of these leases contain at least one auxiliary clause. Each of the auxiliary

[1]The most notable counties without polygon data are Butler county in Western Pennsylvania and Susquehanna county in Northeastern Pennsylvania, both of which have substantial leasing and drilling.

clauses is grouped into one of six clusters—surface protection, externalities, water protection, legal protection, favorable to producer, and optional clauses. The first five follow and build upon Vissing's (2015) typology of lease clauses. The surface protection cluster includes clauses negotiated between the lessor and lessee that are meant to minimize disturbances at the surface of the drilling site. This cluster includes clauses such as a requirement to bury or prohibitions on placing pipelines on the surface, or requirements that oil and gas companies compensate mineral owners for any damage they cause to land, trees, or crops. Clauses in the externalities cluster protect mineral owners against negative externalities such as noise pollution, traffic congestion, and solid waste disposal. The water protection cluster includes clauses meant to prevent water contamination. Clauses in the legal protection cluster include clauses that indemnify mineral owners or affect how mineral owners are compensated. The favorable to producer cluster groups together clauses that benefit the drilling company, such as free access to surface or groundwater on the mineral owner's property. We add an "optional" cluster to Vissing's typology to capture additional options that are typically included in an "addendum" to an oil and gas company's standard leasing document. Table 1 shows that 29 percent of leases in our sample contain addenda, which typically include clauses that provide free household gas to mineral owners, restrictions that the lease cannot be transferred to a firm that is smaller than the company that signs the lease, and restrictions on placing certain infrastructure on the leased property.[2]

In addition to the clause clusters, in our main analysis we focus on nine specific auxiliary clause types that we expect to either be particularly attractive to mineral owners for the protection of their property, or for which we expect increasing media attention over the study period will increase their prevalence. The first type is a clause that prohibits the firm from storing gas in the mineral owner's subsurface, which could prolong the firm's use of the estate if they intermittently use the property to store gas in order to bring natural gas to market at times of high prices. Second, we explore a clause that explicitly prohibits the firm from accessing the mineral owner's surface, meaning that the firm must drill laterally into the mineral estate from an adjacent property. Third, we explore a clause that requires the firm to come to a mutually agreeable development plan before utilizing the mineral owner's surface to produce oil and gas. Development plans could stipulate where well pads, access roads, or fences are placed. Fourth, we explore a clause that specifies that the firm will test the mineral owner's water supply prior to drilling to establish a baseline water quality measure against which to compare future tests in the event of possible contamination. Fifth, we explore a clause that prohibits oil and gas companies from deducting post-production costs from royalty payments (such as transportation costs to deliver gas to buyers). Sixth, we explore a clause that requires firms to replace the mineral owners drinking water supply if drilling is found to contaminate it. Seventh, we explore the numerical value of the "setback" included in the lease, which is the minimum distance within which a firm cannot drill a well near a residential building or water well. To construct the setback clause binary variable, we generate an indicator that equals 1 if the setback in the lease is greater than the statewide

[2]As mentioned above, altogether the 36 clauses in the six clustered are defined 360 unique regular expressions that we identified by reading leases. As a supplement to this paper, we have compiled a database that documents each of the 360 regular expressions, which serves to precisely define each of the auxiliary clauses. The database also includes the regular expression search terms that we coded in python to capture the clauses from lease text files that were recovered by the OCR process. We will make the database available with the final published version of this paper, and it is available upon request from the authors prior to publication.

minimum setback. To account for a legislative change, for leases signed prior to 2012 we use the minimum setback of 200 feet, and for leases signed in 2012 and after we use a minimum of 500 feet. Eighth, we examine a clause that prohibits placing pipelines on the mineral owner's surface. Ninth and finally, we explore a group of "additional infrastructure" clauses that prohibit placing compressor stations, natural gas and waste storage tanks, and hazardous waste storage receptacles on the mineral owner's surface.

4.2 Oil and Gas Well Data

Data on the characteristics, location, and production data of unconventional oil and gas well permits come from the Unconventional Natural Gas Well Geodatabase published by the Carnegie Museum of Natural History (2022). We utilize the data on wells drilled between 2008 to 2021 to estimate the average first year production of wells within either a 2 km or 3 km radius of a lease. To do this, we use a methodology similar to Harleman (2021), which utilizes oil and gas well production reports to estimate a decline curve model for each well, and takes the integral of oil and gas production (in trillion btus) over the first 365 days of the well's lifespan.[3] We utilize the estimated first year production as a proxy for the expected ultimate recovery, which is not a known value for unconventional wells in the Marcellus because the vast majority remain in production to date. First year production is highly correlated with expected ultimate recovery, since daily natural gas production is highest when a well is initially drilled and declines over time without further stimulation. Data from the US Energy Information Administration (2021) suggests that over 40 percent of a typical Marcellus well's ultimate production comes in its first year.

The Appendix describes in detail the steps that we take to spatially connect wells to leases and estimate the average first year production and total depth across all wells near a lease. In total, we have data on 8,445 leases that have at least ten wells within 2 km. Table 2 presents descriptive statistics for the average royalty rate for the complete sample of 58,559 leases. It also presents descriptive statistics for the 8,445 leases near wells with complete data for estimated first year production and average total drilled depth across all wells within 2 km of the lease. In Appendix section A.3 we show that our first year production estimates for each well are consistent with other estimates from Harleman (2021) and the US Energy Information Administration (2021). Table 2 shows that leases in our sample are near wells that produce an average of 1.63 trillion British thermal units (btu) in their first year, which is also consistent with these sources.[4]

Table 2 also shows that our 8,445 leases are near wells with an average total depth of 12,627 feet. We utilize this depth variable as an instrumental variable to account for measurement error in our first year production estimates, as described in the following section. Data on the total depth of each well, our instrument variable, comes from the Pennsylvania Department of Conservation and Natural Resources

[3] We estimate first year production as a proxy for expected ultimate recovery because for the vast majority of wells we do not observe cumulative production on day 365 (e.g., the first production report of a well may cover the first 60 days of a year, while the second production report of that well may cover the entire following calendar year).

[4] Harleman (2021) uses a decline curve production model from the US Energy Information Administration (2021) to show that the typical Marcellus well produces around 1.5 trillion btu in its first year (see Harleman (2021), Online Appendix Page 9). Our estimate here that the average lease is near wells that produce an average of 1.63 trillion btu is slightly higher because there is a greater concentration of wells in places with more productive wells.

(2022). We describe how we assign each lease an average depth across all nearby wells in Appendix sections A.4 and A.5. The depth variable from the Department of Conservation and Natural Resources (DCNR) is the total depth to which the well has been drilled. We utilize this variable, and not a more incomplete variable measuring the deepest depth at which the oil and gas driller extracts oil or gas, because it is non-missing for nearly every well in our sample. The two depth variables are highly correlated, with a Pearson correlation coefficient of .98. Figure 3 Panel A displays a map of the 7,336 wells that are within 2 km of a lease in our sample. Panel B displays a map of the 58,559 leases in our full analytical sample.

5 Empirical Approach

Our empirical approach is in three parts. First, to test whether a trade-off exists between health-protective clauses and royalty rates, we estimate the following model with ordinary least squares (OLS):

$$RoyaltyRate_{lct} = \beta_0 + \beta_1 Clause_{lct} + \tau_t + \lambda_c + \alpha + \varepsilon_{lct} \tag{1}$$

Our outcome variable is $RoyaltyRate_{lct}$, which is the royalty rate contained in lease l which was signed in county c and year t. Our explanatory variable of interest in the first specification is $Clause_{lct}$, which is a dummy variable for whether a given type of auxiliary clause is present in lease l. If mineral owners treat royalty rates and auxiliary clauses as substitutes when negotiating leasing terms, we would expect β_1 to be negative. If instead some mineral owners are able to negotiate for leases that are more favorable all-around, β_1 would be positive. In our preferred specifications we include year (τ_t), county (λ_c), and oil and gas firm (α) fixed effects to account for unobserved geographic, temporal, and firm-specific characteristics that may lead to correlations between clauses and royalty rates that are not reflective of trade-offs within a given leasing negotiation.

In the second part of our empirical approach, we examine how clauses evolve over time by plotting the proportion of leases signed in year t that contain a given auxiliary clause. We also plot average royalty rates over time. However, these simple time plots are not an unbiased reflection of mineral owners negotiating over time if different temporal waves of leasing are driven by certain operators or occur in certain counties. This is important because many oil and gas companies have proprietary leasing templates, and if that company signs many leases in a given year it would give the false impression that mineral owners began favoring a particular clause. To overcome this threat to validity, we estimate the following model with OLS:

$$RoyaltyRate_{lct} = \beta_0 + \sum_{\tau=2002}^{2016} \beta_\tau \tau_t + \lambda_c + \alpha + \varepsilon_{lct} \tag{2}$$

In some specifications, we replace the outcome with a binary variable that indicates the presence of a particular auxiliary clause ($Clause_{lct}$). We plot our annual estimates of β_τ, which represent the difference in the royalty rate (or likelihood of clause adoption) in year t relative to the omitted year, 2001. Because we control for county fixed effects (λ_c) and firm fixed effects (α), the estimates represent changes over time that are not due to idiosyncratic firm characteristics such as leasing templates.

Our third and final model to examine the determinants of oil and gas royalties and leases builds off the approach for measuring the pass-through of geologic productivity into royalty rates developed by Brown et al. (2016). They develop a theoretical model in which the natural logarithm of the royalty rate offered by a firm is a function of the natural logarithm of the expected ultimate recovery of oil and gas at a given well, the time-varying price of natural gas and time-varying market return on capital, and fixed expenditures to drill a well. Because they do not have a measure of expected ultimate recovery for each lease, their model relies on relating a time-constant average royalty rate in a county to a time-constant estimate of the expected ultimate recovery of the typical well in the county. We estimate a more localized relationship between royalty rates and geologic productivity:

$$ln(RoyaltyRate_{lct}) = \beta_0 + \beta_1 ln(FirstYearProduction_{lct}) + \tau_t + \lambda_c + \alpha + \varepsilon_{lct} \tag{3}$$

Where $ln(FirstYearProduction_{lct})$ is the estimated amount of oil and gas produced in the first year in trillions of British thermal units (trillion btu), averaged across all wells within a 2 km radius of the lease. In this model, we only include leases that have at least 10 wells drilled within the 2 km radius, and use the average first year production across the wells as a proxy for expected ultimate recovery.[5] As in the other three models, using lease-level data allows us to include fixed effects for year, county, and oil and gas firm. By including these fixed effects we effectively control for the time-varying market conditions that Brown et al. (2016) control for by calculating averaging annual interest rates across time with the weight on each year determined by the acreage-weighted share of leases signed in the county in that year. In addition to including $ln(RoyaltyRate_{lct})$ as the outcome in equation 3, we also estimate the model with $Clause_{lct}$ as the outcome, varying the definition of the variable across models to examine different auxiliary clauses. Doing so allows us to determine whether owners of more productive oil and gas resources are able to negotiate for more favorable leasing conditions.

In separate specifications of model 3, we estimate the first year production variable using only wells that were drilled *before* the lease was signed. In these specifications, we expand the radius to 3 km to retain a suitable sample size of leases with at least 10 wells nearby. This specification allows us to test whether the pass-through of productivity is stronger when mineral owners or firms learn about the productivity or royalty rates offered in a given geographic area. If this type of learning occurs, we would expect the production of wells drilled before the lease was signed to have a stronger relationship with royalties.

Because $FirstYearProduction_{lct}$ is estimated for each well (see Appendix A.3) and averaged across all wells within 2 km of the lease, it is inevitably measured with error. To account for this measurement error that would result in a downward bias in the relationship between geologic productivity and leasing terms, we utilize the average total drilled depth of all wells within 2 km as an instrument for average first year production. Data on the total depth of each well, which we link to our data on well locations using

[5]Wells experience a sharp decline in production after their first year, and their initial rate of production is one of the strongest predictors of how much oil and gas they ultimately produce (Ikonnikova et al., 2015). Appendix A.3 details how we use data from the Pennsylvania Department of Environmental Protection to estimate first year production for each well near an oil and gas lease.

each well's unique permit number, comes from the Pennsylvania Department of Conservation and Natural Resources (2022). Depth should be a relevant instrument because deeper shale resources produce greater amounts of oil and gas (Marchand and Weber, 2020), and is the same variable used to instrument for county-level expected ultimate recovery by Brown et al. (2016).

6 Results

6.1 Correlations Between Royalty Rates and Clauses

Table 3 shows the correlations between royalty rates and six "clusters" of clause types described in Section 4.1. Robust standard errors are reported in parenthesis in this table and in all subsequent regression tables. The equation 1 models are estimated with and without fixed effects for the oil and gas firm that signed the lease, to explore whether idiosyncratic firm leasing templates affect the relationship. The coefficients labeled "cluster" can be interpreted as the association between the presence of any one clause in the cluster and royalty rates, which are captured as a proportion (i.e., the state minimum rate of 12.5 percent is coded as .125). For example, in Column 1, the presence of any one of the eight clauses in the surface protection cluster is associated with a .3 percent higher royalty rate on average. This relationship changes little with the inclusion of firm fixed effects. The table suggests that there is a positive, but weak correlation between clauses aimed to protect the mineral owner's surface and water, while the other four clusters have weak negative or zero correlations. Adding the six coefficients from our preferred specifications (the even Columns) suggests no overall relationship between clauses and royalty rates.

Table 4 shows the correlation between royalty rates and nine individual clause types discussed in Section 4.1. Altogether, the results indicate a positive and weak correlation between royalties and these clauses. The strongest associations are for the clauses that prohibit gas storage and additional infrastructure (compressor stations, storage tanks, and hazardous waste storage) on the property, which correspond to nearly a 1 percent higher average royalty rate. Interestingly, leases that contain clauses to prohibit firms from deducting post-production costs from the value of production are associated with a .6 percent higher average royalty rate. This suggests that some mineral owners are more informed or more capable of negotiating better initial royalty rates and protections to ensure their payments are maximized. The exception is the presence of a clause that prohibits surface use, which is associated with a .4 percent lower average royalty rate. This suggests that firms are willing to provide fewer royalties to mineral owners that prohibit drilling, placing pipelines, or building roads on their surface.

Taken together, the results from Tables 3 and 4 indicate that there is a weak positive relationship between royalty rates and many individual clause types aimed at environmentally or monetarily benefiting mineral owners. Our first hypothesis relied on the standard utility maximization model under efficient Coasean bargaining to predict that royalty rates and the presence of auxiliary clauses are negatively associated. Under this hypothesis, mineral owners view royalties and clauses as substitutes when negotiating leases. We reject our first hypothesis, as our results suggest that mineral owners simultaneously negotiate for higher royalty payments and more protective clauses. This positive association between royalty rates and clauses aligns with Panel C of Figure 2, and suggests that certain mineral owners have greater ability to negotiate, or greater information about the impacts of the shale industry.

6.2 Trends of Royalty Rates and Clauses Over Time

The left side of Figures 4 through 6 display simple bar graphs of the percentage of leases signed in a given year that contained a specific clause type. The solid lines on the right side are the coefficients estimated by model 2, which contain firm and operator fixed effects and represent trends in leasing that are not driven by certain firms or counties being disproportionately represented in a given year. The dotted lines are 95 percent confidence intervals, calculated using robust standard errors.

Turning first to Panel A of Figure 4 for exposition, 1 percent of leases signed in 2001 contained a clause prohibiting gas storage, which grew up to 28.9 percent for leases signed in 2015. Conditional on firm and county leasing characteristics, leases signed in 2008 are 7 percentage points more likely to contain the gas storage clause, relative to leases signed in 2001, but no clear pattern emerges afterward. A clearer pattern emerges in Panel B, with the no surface use clause being virtually absent from leases signed before 2009, yet reaching a maximum of 19.6 percent penetration in 2012. No surface use clauses may in part be crowding out development plan clauses (Panel C), which grow in their prevalence through 2008, up to a maximum of 53.5 percent of leases signed in 2008, and subsequently fall later in the study period.

Figure 5 Panel A shows large and steady increases in the prevalence of the clause that requires baseline water quality testing prior to drilling, which is virtually absent in 2001 and is in nearly half of all leases signed in the final two years of the study period. Clauses that require firms to replace the mineral owner's drinking water if oil and gas extraction is found to cause contamination also follow an upward, albeit much weaker trend. These trends in water testing and replacement clauses hold when accounting for county and firm fixed effects. Panel C shows a weak trend in clauses that require greater than the state-wide minimum setback from residential buildings and water wells, and this type of clause became virtually absent from leases after a legislative change in 2012 that extended the setback from 200 to 500 feet.

Figure 6 Panel A shows a downward trend in clauses that prohibit placing pipelines on a mineral owner's property, conditional on the fixed effects. This is likely explained by the development of several high-volume pipeline projects and their associated collection lines that were planned in the state throughout the 2010s (State Impact Pennsylvania, 2018). For example, the Mariner East pipeline was completed in 2014, and carries natural gas from western Pennsylvania to industrial complexes over 300 miles away in eastern Pennsylvania and Delaware (Energy Transfer, 2022). As the industry moved toward transporting natural gas via pipelines rather than ground transportation, oil and gas firms became less likely to offer pipeline prohibitions. Conversely, as the pipeline infrastructure was built out, mineral owners appear to have become more likely to negotiate restrictions on placing additional infrastructure on their properties, such as compressor stations and natural gas and waste storage tanks (Panel B).

The results in the three figures are mixed. The use of development plans, greater than minimum setbacks, and pipeline prohibitions fall by the end of the study period. But water-protection clauses, gas storage prohibitions, and no surface use clauses appear to grow. But concerns about what quality were the dominant concern in the public discourse surrounding fracking, and the no surface use clauses provides the most overarching security against its environmental impacts. For these reasons, we have weak support for our second hypothesis that mineral owners sign leases with more protective clauses over time as the shale play develops. This may be due to some combination of affected residents learning about specific industry practices and health concerns through word of mouth or the media, or firms becoming less environmentally intensive and therefore willing to include more protective clauses.[6]

Turning to the left side of Figure 6, Panel C shows that annual average royalty rates remained flat at the state minimum of 12.5 percent until widespread drilling picked up in 2008, rose to a maximum of 16.8 percent in 2010, and subsequently fell year-over-year to around 13 percent by 2016. The right side shows that these trends hold when accounting for different leasing waves across firms and counties. The results are generally consistent with our third hypothesis that that royalty rates rise over the study period. Prior to 2008, low expectations and knowledge about the productivity of the Marcellus shale resulted in firms offering and mineral owners accepting state-mandated minimum rates. Going forward, a combination of competition among firms to secure productive mineral rights and growing knowledge among mineral owners that their resources were valuable led to elevated rates. That rates peaked in 2010 is likely due to relatively high natural gas prices through 2010 that made drilling more profitable and leases more desirable (US Energy Information Administration, 2022). Beyond 2010, royalty rates appear to fall with future expectations of natural gas prices as the shale gas boom led to elevated supply. Rates may also be falling because firms must undergo costly search to locate holdout mineral owners in later years of development (Agerton, 2020). Similarly, in Panel D firms appear to become less willing to prohibit royalty deductions in times of low natural gas prices, as exhibited by declines in the use of such clauses after 2010.[7]

6.3 Pass-through of Geology into Royalty Payments and Clauses

To estimate the extent that greater geologic productivity passes through into royalty rates, we begin by exploring the first stage relationship between lease-level averages of estimated first year production (our independent variable) and total drilled depth (our instrumental variable). As discussed in 5, we follow Brown et al. (2016) by using depth as an instrument for production because our independent variable, which is an estimate, is measured with error that would lead to a downward bias in its relationship with royalty rates. The first stage results are in Table 5. The fixed effects vary across the first three columns, which include our full analytical sample of 8,445 leases that have at least 10 wells within a 2 km radius. The fourth

[6] For example, unconventional natural gas well laterals, the portion of the well that runs horizontally through the shale formation, have become longer over time. Longer laterals mean that oil and gas firms can drill fewer wells and construct fewer well pads to access a fixed amount of natural gas. This technological improvement would make the firms more willing to sign no surface use clauses.

[7] That clauses prohibiting post-production cost deductions from royalty rates appear to rebound in prevalence in 2016 could be due to lawsuits brought by the state of Pennsylvania against two major oil and gas companies in 2015 and 2016 (Cocklin, 2012). If operators may have been attempting to assuage mineral owners or shield themselves against similar suits by stating clearly in the lease that they would not deduct the costs.

column includes a subset of 1,367 leases with at least 10 wells drilled within a 3 km radius *before* the lease was signed. Across all four columns, we have positive and statistically significant coefficients on the depth instrument, and F-statistics indicative of a strong instrument.

The estimation results for the pass-through of first year production into royalty rates are shown in Table 6. The full analytical sample of 8,445 leases (wells with 2 km) are in Columns 1 through 6, and the subset of 1,367 leases (wells with a 3 km radius drilled before the lease) are in Columns 7 and 8. The OLS estimates are in the odd columns, and the two-stage least squares (2SLS) estimates are in the even columns. The first two columns include no controls. We add county and year fixed effects in Columns 3 and 4, and additionally add firm fixed effects in Columns 5 through 8. In all columns, we observe a weak positive relationship between first year production and royalty rates. The exception is the OLS estimate in Column 3, which becomes positive when instrumenting for first year production in Column 4. Our preferred estimates are the fully controlled 2SLS models in Columns 6 and 8. Both are positive, with the estimate in Column 8 being slightly shy of statistical significance (p=0.121) due to the smaller sample size.

We use the coefficients to understand how a doubling of $ln(FirstYearProduction)$ affects royalty rates. We find that a doubling of first year production—an increase of 0.70 log points—leads to a 1 to 1.6 percent ($= 0.70 \times 1.42\%$, $0.70 \times 2.24\%$) increase in the share of the value of production going to the mineral owner. Brown et al. (2016) find a nearly identical relationship, with a doubling of expected ultimate recovery at the county level leading to a 1 to 2.2 percent increase in the share of value of production going to the mineral owner. Our findings are consistent with limited pass-through of resource abundance into royalty rates, even when accounting for hyper-localized variation in geologic productivity. Further, our separate specification that tests whether greater production of wells drilled *before* the lease was signed suggests that mineral owners do not learn about localized geological productivity from prior drilling, or that they are unable to leverage this learning to negotiate for higher royalties. Based on these results, we reject our third hypothesis that that owners of more productive mineral resources negotiate for higher royalty rates and greater clauses.

As an extension of our pass-through analysis, in Table 7 we examine whether mineral owners are able to leverage greater geological productivity to negotiate for more protective clauses. OLS estimates are in the odd columns, and 2SLS estimates are in the even columns. Focusing on the preferred 2SLS specifications, a doubling of the natural logarithm of first year production leads to roughly a .75 and 3 percentage point change in the likelihood of a clause being adopted for the gas storage, development plan, royalty deduction, and greater than minimum setback clause types.[8] All other clause types have an insignificant relationship with first year production, except for the no surface use clause, which is roughly 2.65 percentage points less likely to be adopted with a doubling of first year production at nearby wells. These estimates suggest a weak relationship between the quality of the mineral resource underlying the leased acreage and the quality of auxiliary clauses included in leases, lending further evidence to our previous finding that mineral owners are not trading off royalty rates for more protective clauses. In other words, the limited pass-through of geological productivity into royalty rates does not appear to be offset by its pass-through in auxiliary clauses.

[8] Unlike the royalty analysis in Table 6 which is in log-log functional form, Table 7 is in level-log form. This range of positive coefficients is from the setback clause $.011 \times .7 = 0.0077$ or a .77 percentage point increase, and for the no royalty deduction clause $.048 \times .7 \times 100 = 0.033$ or a 3.3 percentage point increase.

7 Discussion

Two of our findings suggest that firms exhibit market power in leasing markets. First, our results suggest that some mineral owners are able to negotiate for all-around better leases, which means that firms expect positive profits even when making concessions on both royalty rates and clauses. This mirrors Panel C of Figure 2, where mineral owners with greater information and resources are able to achieve the combination of $(P3, C3)$. This outcome is in contrast to a perfectly competitive leasing market, where competition among firms to secure leased acreage would result in equilibrium combinations of royalties and clauses that would require mineral owners to make trade-offs across the two instruments when negotiating leases. Second, with Brown et al. (2016) we find limited pass-through of geologic productivity that is much lower than the perfectly competitive scenario in which a one percent increase in resource abundance leads to a one percent increase in the royalty rate. They offer two explanations for this result: that oil and gas firms are monopsonistic in leasing markets, or that uncertainty about productivity along with non-renegotiable royalty rates lock in mineral owners to initial leases even if the resource proves to be very productive. We provide some evidence in support of the former explanation because publicly available production data from wells very close to the leased acreage and drilled before the lease was signed bears little relationship with royalties.

If firms do behave as monopsonists, it is likely to be mineral owners from the highest-income or the most socially organized communities that have the skills, knowledge, and resources to negotiate for all-around better leases. This may point to a distributive environmental justice concern, such as those previously documented in a small oil and gas leasing literature. Vissing (2015) studies auxiliary clauses in Tarrant County, Texas, and shows that leases in Census tracts with greater shares of minority, non-English speaking, and less-wealthy households have lower royalty rates and are less likely to contain clauses that benefit mineral owners. Similarly, Timmins and Vissing (2022) find that wealth, race, ethnicity, and language affect the prevalence of beneficial lease clauses.

Policies that provide greater information to mineral owners or standardize leasing templates across a shale play may engender more equitable outcomes across race, ethnicity, and income groups. McFarland (2022) recommends that leasing guides, which include a checklist of clauses that can be incorporated in the lease, be provided to mineral owners. Another policy mechanism that could mitigate information asymmetries is making the terms of prior leases that were recently negotiated available to mineral owners who are currently negotiating lease terms (Vissing, 2015). State-mandated uniform leasing templates, with a set of standardized auxiliary clauses that provide greater than minimum protection against the externalities associated with drilling, could also protect the health of households and communities.

8 Conclusion

This paper studies the determinants of oil and gas leasing outcomes in the Marcellus Shale in Pennsylvania, and produces three main findings. First, we show that there is a positive association between royalty rates and environmentally protective clauses. Second, we show that royalty rates and some of the most environmentally protective clauses become more prevalent as the shale play develops over time. Third, consistent with past findings at a more coarse level of geographic aggregation, we show that a hyper-localized measure of geologic productivity bears only a weak positive relationship with the royalty rate received by the landowner. Productivity also bears a weak relationship with the likelihood of a lease including beneficial lease clauses. That owners of more productive mineral resources do not receive higher royalty rates or stronger leases indicates that improved expectations and knowledge about the profitability of the shale play as a whole is the primary driver of leasing outcomes. This suggests that mineral owners who hold out to sign leases later in the development of the shale play, who are likely to be high-income individuals with lower marginal utility of income, are likely to sign the better leases. High-income mineral owners are also more likely to have greater skills, knowledge, and resources to negotiate for stronger leases and greater compensation, which may motivate policies to provide uniform information to all landowners or to mandate standardized leasing templates in order to engender more equitable outcomes in leasing markets.

References

Agerton, M. (2020). Learning where to drill: Drilling decisions and geological quality in the haynesville shale.

Arnould, R. J. and H. Grabowski (1981). Auto safety regulation: An analysis of market failure. *The Bell Journal of Economics*, 27–48.

Atherton, E., D. Risk, C. Fougere, M. Lavoie, A. Marshall, J. Werring, J. P. Williams, and C. Minions (2017). Mobile measurement of methane emissions from natural gas developments in northeastern british columbia, canada. *Atmospheric Chemistry and Physics 17*(20), 12405–12420.

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

Bonetti, P., C. Leuz, and G. Michelon (2021). Large-sample evidence on the impact of unconventional oil and gas development on surface waters. *Science 373*(6557), 896–902.

Brasier, K. J., L. Davis, L. Glenna, T. W. Kelsey, D. K. McLaughlin, K. Schafft, K. Babbie, C. Biddle, A. DeLessio-Parson, D. Rhubart, et al. (2015). Communities experiencing shale gas development. In *Economics of unconventional shale gas development*, pp. 149–178. Springer.

Brown, J. P., T. Fitzgerald, and J. G. Weber (2016). Capturing rents from natural resource abundance: Private royalties from us onshore oil & gas production. *Resource and Energy Economics 46*, 23–38.

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

Calfee, J. E., C. Winston, and R. Stempski (2002). Direct-to-consumer advertising and the demand for cholesterol-reducing drugs. *the Journal of Law and Economics 45*(S2), 673–690.

Carnegie Museum of Natural History (2022). Pennsylvania unconventional natural gas well geodatabse. `https://maps.carnegiemnh.org/index.php/projects/unconventional-wells/download/`.

Coase, R. H. (1960). The problem of social cost. *The Journal of Law and Economics 56*(4), 837–877.

Cocklin, J. (2012). Pennsylvania supreme court deals blow to state's case against natural gas royalty deductions. `https://www.naturalgasintel.com/pennsylvania-supreme-court-deals-blow-to-states-case-against-natural-gas-royalty-deductions/`

Covert, T. (2015). Experiential and social learning in firms: the case of hydraulic fracturing in the bakken shale. *Available at SSRN 2481321*.

Energy Transfer (2022). Mariner pipeline. `https://marinerpipelinefacts.com/`.

Fetter, T. R., A. L. Steck, C. Timmins, and D. Wrenn (2018). Learning by viewing? social learning, regulatory disclosure, and firm productivity in shale gas. Technical report, National Bureau of Economic Research.

Fitzgerald, T. (2015). Experiential gains with a new technology: An empirical investigation of hydraulic fracturing. *Agricultural and Resource Economics Review 44*(2), 83–105.

Harleman, M. (2021). Compensating communities for industrial disamenities: The case of shale gas development. *Economic Inquiry*.

Harleman, M. and J. G. Weber (2017). Natural resource ownership, financial gains, and governance: The case of unconventional gas development in the uk and the us. *Energy Policy 111*, 281–296.

Hausman, C. and R. Kellogg (2015). Welfare and distributional implications of shale gas. Technical report, National Bureau of Economic Research.

Hill, E. and L. Ma (2017). Shale gas development and drinking water quality. *American Economic Review 107*(5), 522–25.

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

Hill, E. L. and L. Ma (2022). Drinking water, fracking, and infant health. *Journal of Health Economics 82*, 102595.

Ikonnikova, S., J. Browning, G. Gülen, K. Smye, and S. Tinker (2015). Factors influencing shale gas production forecasting: Empirical studies of barnett, fayetteville, haynesville, and marcellus shale plays. *Economics of Energy & Environmental Policy 4*(1), 19–36.

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

Levitt, C. J. (2011). Learning through oil and gas exploration. Technical report, Working Paper, November.

Macey, G. P., R. Breech, M. Chernaik, C. Cox, D. Larson, D. Thomas, and D. O. Carpenter (2014). Air concentrations of volatile compounds near oil and gas production: a community-based exploratory study. *Environmental Health 13*(1), 1–18.

Mannering, F. and C. Winston (1995). Automobile air bags in the 1990s: market failure or market efficiency? *The Journal of Law and Economics 38*(2), 265–279.

Marchand, J. and J. G. Weber (2020). How local economic conditions affect school finances, teacher quality, and student achievement: Evidence from the texas shale boom. *Journal of Policy Analysis and Management 39*(1), 36–63.

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

McElroy, J. A., C. D. Kassotis, and S. C. Nagel (2020). In our backyard: Perceptions about fracking, science, and health by community members. *New Solutions: A Journal of Environmental and Occupational Health Policy 30*(1), 42–51. PMID: 32028857.

McFarland (2022). Checklist for negotiating an oil and gas lease. `https://www.landcan.org/pdfs/jbm-ogleasechecklist_163158826.pdf`.

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

Payne, B. F., R. Ackley, A. Paige Wicker, Z. L. Hildenbrand, D. D. Carlton, and K. A. Schug (2017). Characterization of methane plumes downwind of natural gas compressor stations in pennsylvania and new york. *Science of The Total Environment 580*, 1214–1221.

Pennsylvania Department of Conservation and Natural Resources (2022). Exploration and development well information network. `https://edwin.hylandcloud.com/`.

Pennsylvania Department of Environmental Protection (2022). Oil and gas reports. `https://www.dep.pa.gov/DataandTools/Reports/Oil%20and%20Gas%20Reports/Pages/default.aspx`.

Rahm, B. G., S. Vedachalam, L. R. Bertoia, D. Mehta, V. S. Vanka, and S. J. Riha (2015). Shale gas operator violations in the marcellus and what they tell us about water resource risks. *Energy Policy 82*, 1–11.

Richardson, N., M. Gottlieb, A. Krupnick, and H. Wiseman (2013). The state of state shale gas regulation. *Washington, DC: Resources for the Future*.

Sangaramoorthy, T., A. M. Jamison, M. D. Boyle, D. C. Payne-Sturges, A. Sapkota, D. K. Milton, and S. M. Wilson (2016). Place-based perceptions of the impacts of fracking along the marcellus shale. *Social Science and Medicine 151*, 27–37.

State Impact Pennsylvania (2018). Your guide to pipelines. `https://stateimpact.npr.org/pennsylvania/tag/pipelines/`.

Timmins, C. and A. Vissing (2022). Environmental justice and coasian bargaining: The role of race, ethnicity, and income in lease negotiations for shale gas. *Journal of Environmental Economics and Management 114*, 102657.

US Energy Information Administration (2011). Review of emerging resources: Gas, us shale and plays, shale oil.

US Energy Information Administration (2021). Production decline curve analysis in the annual energy outlook. https://www.eia.gov/analysis/drilling/curve_analysis/.

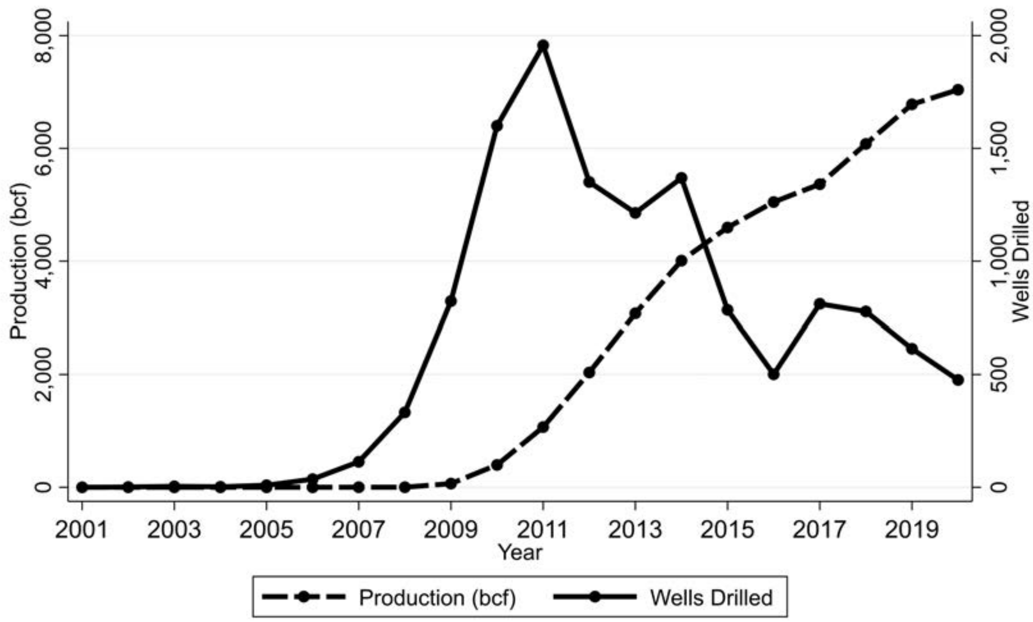
US Energy Information Administration (2022). Henry hub natural gas spot price. https://www.eia.gov/dnav/ng/hist/rngwhhdA.htm.

US Geological Survey (2022). Usgs releases new assessment of gas resources in the marcellus shale. https://pubs.usgs.gov/fs/2011/3092/pdf/fs2011-3092.pdf.

Vissing, A. (2015). Private contracts as regulation: A study of private lease negotiations using the texas natural gas industry. *Agricultural and Resource Economics Review 44*(2), 120–137.

Zhang, R., H. Li, N. Khanna, D. Sullivan, A. Krupnick, and E. Hill (2023). Air pollution impacts of shale gas development in pennsylvania. *Journal of the Association of Environmental and Resource Economists*.

Zimmerle, D. J., C. K. Pickering, C. S. Bell, G. A. Heath, D. Nummedal, G. Pétron, and T. L. Vaughn (2017, 11). Gathering pipeline methane emissions in Fayetteville shale pipelines and scoping guidelines for future pipeline measurement campaigns. *Elementa: Science of the Anthropocene 5*. 70.



Note: Data for on wells drilled are from the Pennsylvania Department of Environmental Protection (2022). Data on production are from the US Energy Information Administration (2022).

Figure 1: Unconventional Natural Gas Drilling and Production in the Marcellus

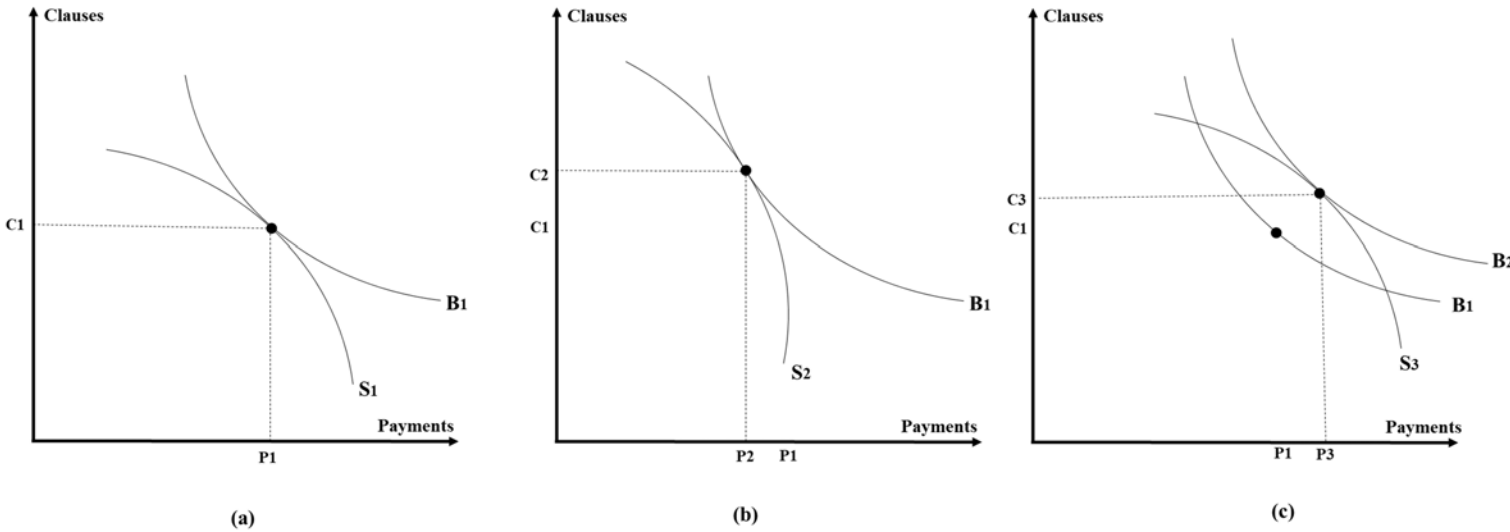
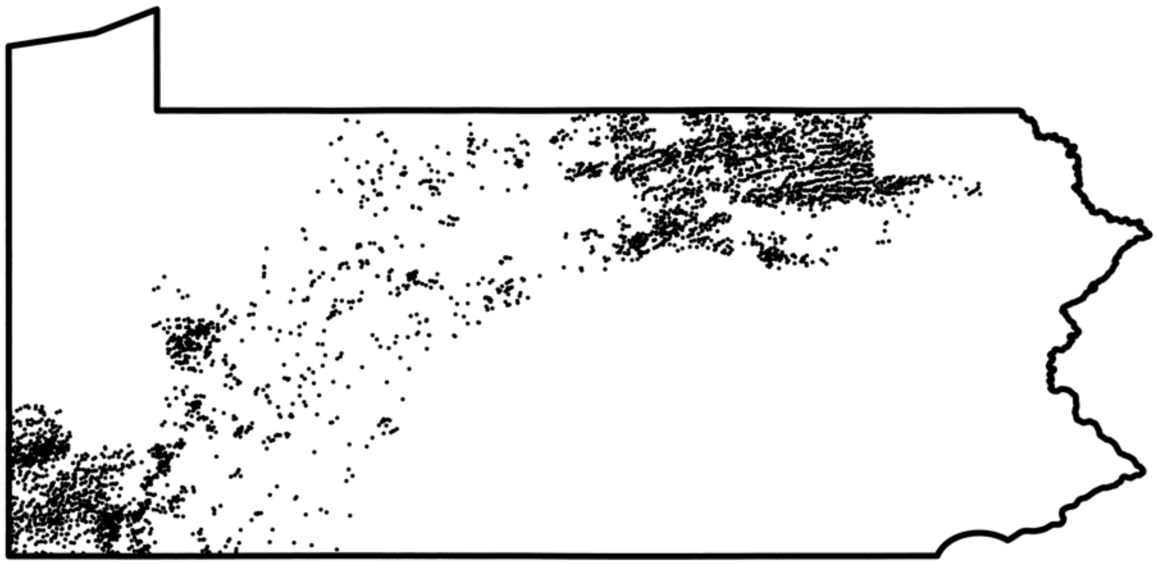
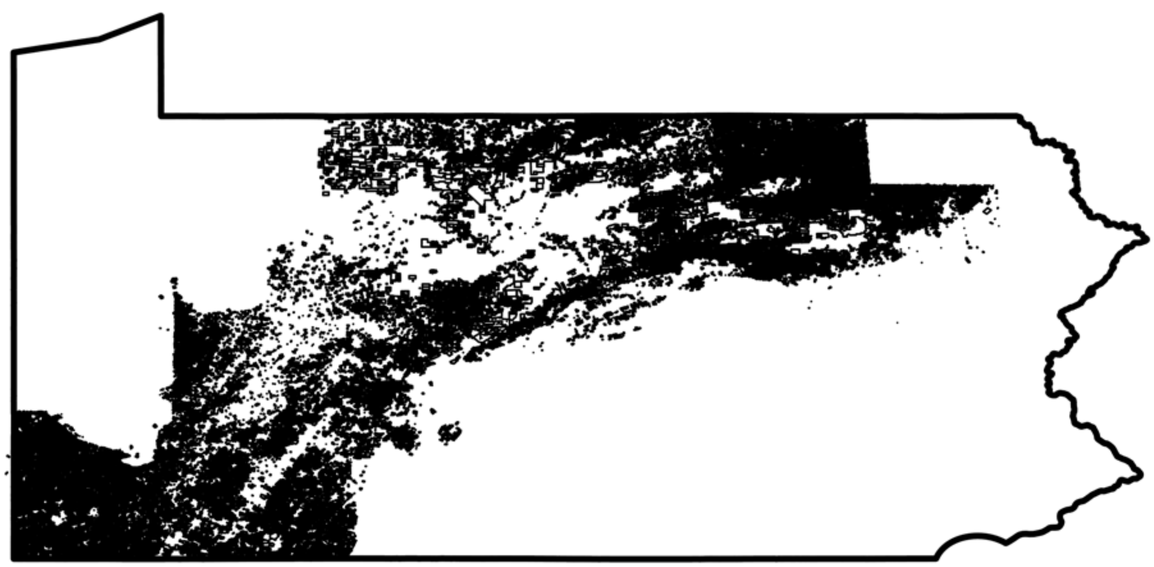


Figure 2: Lease Negotiations Between Operators and Mineral Owners

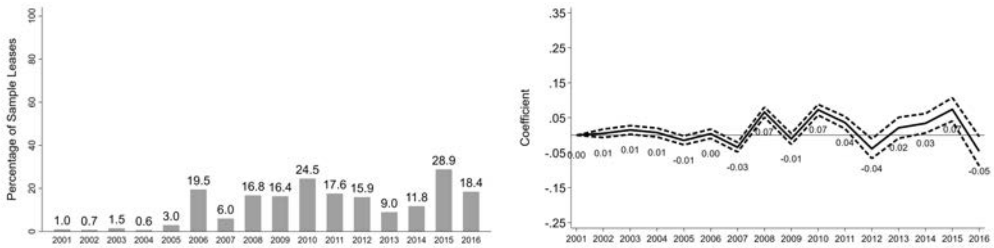


(a) Map of Wells

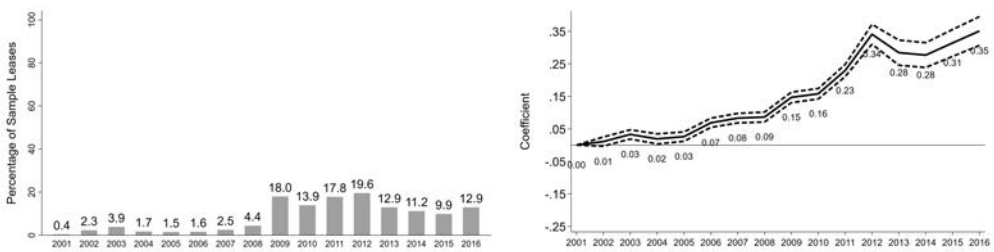


(b) Map of Leases

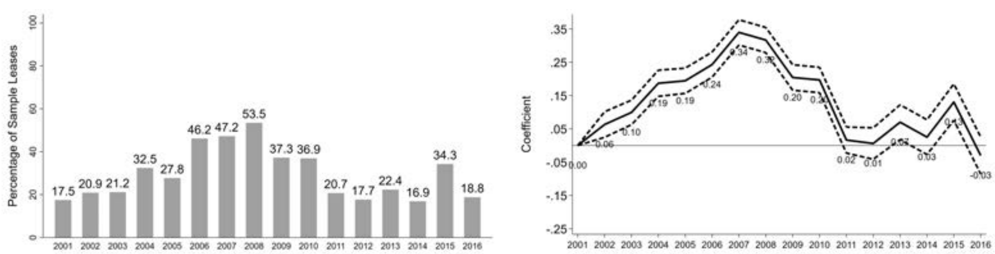
Figure 3: Map of Sample Wells and Leases



(a) No Gas Storage



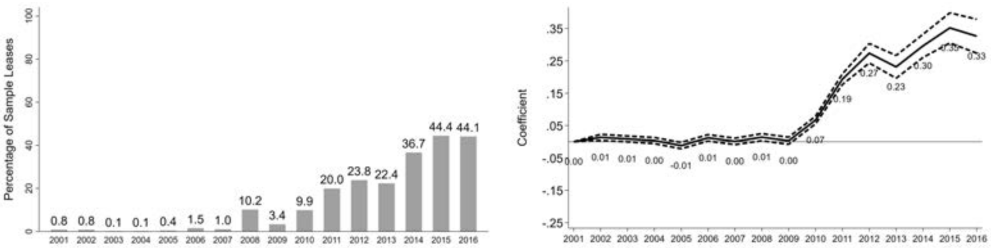
(b) No Surface Use



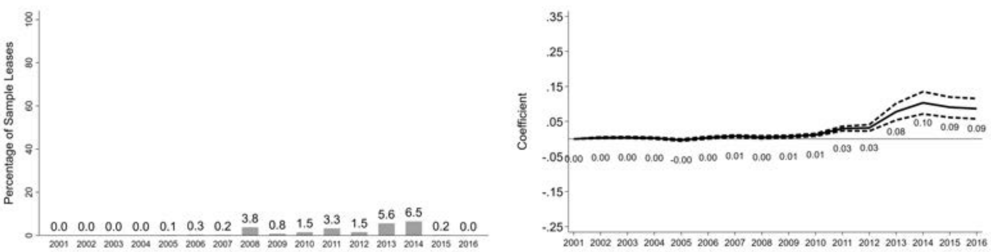
(c) Development Plan

Note: On the left are simple bar graphs of the percentage of leases signed in a given year that contained the clause. On the right, the solid line is the estimated OLS coefficient from equation 2, and the dotted lines are 95% confidence intervals, calculated using robust standard errors. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001.

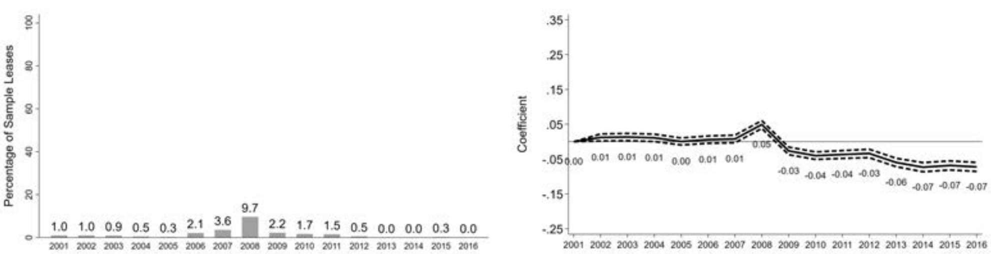
Figure 4: The Prevalence of Select Clauses Over Time



(a) Water Testing



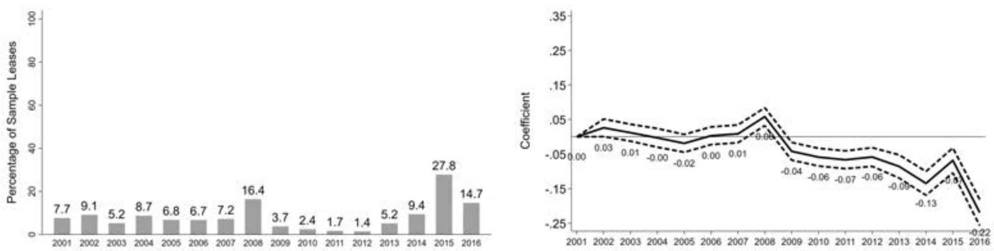
(b) Water Replacement



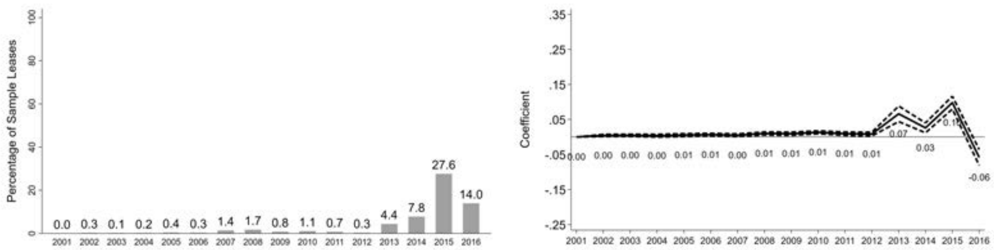
(c) Greater than Minimum Setback

Note: On the left are simple bar graphs of the percentage of leases signed in a given year that contained the clause. On the right, the solid line is the estimated OLS coefficient from equation 2, and the dotted lines are 95% confidence intervals, calculated using robust standard errors. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001.

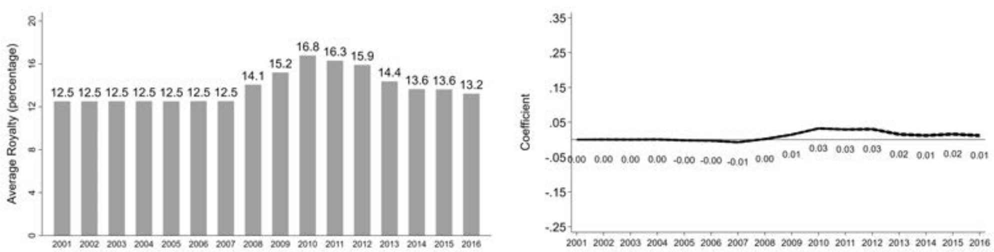
Figure 5: The Prevalence of Select Clauses Over Time (cont.)



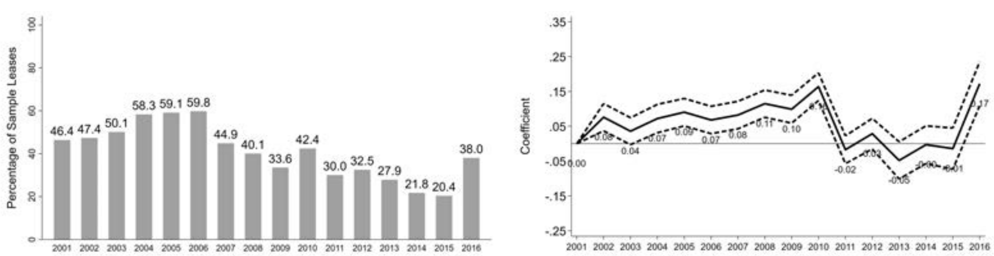
(a) No Pipelines



(b) No Compressors, Tanks, Waste Storage



(c) Royalties



(d) No Royalty Deductions

Note: On the left are simple bar graphs of the percentage of leases signed in a given year that contained the clause. On the right, the solid line is the estimated OLS coefficient from equation 2, and the dotted lines are 95% confidence intervals, calculated using robust standard errors. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001.

Figure 6: The Prevalence of Select Clauses Over Time (cont. 2)

Table 1: Leases and the Prevalence of Auxiliary Clauses

	With Clause	**Without Clause**	**Percent With Clause**
Surface Protection Cluster	**50,642**	**7,917**	**86**
No Surface Use	5,310	53,249	9
Minimize Disturbance	5,868	52,691	10
Damage Compensation	47,967	10,592	82
Development Plan	21,777	36,782	37
Road Restriction	326	58,233	1
No Pipelines	4,345	54,214	7
Bury Pipelines	15,346	43,213	26
Well Spacing	2,719	55,840	5
Externalities Cluster	**32,042**	**26,517**	**55**
Environmental Protection	428	58,131	1
Noise Restriction	469	58,090	1
Working Hours	4	58,555	0
Development Notification	2,337	56,222	4
Setback	24,655	33,904	42
Traffic	128	58,431	0
Fences and Gates	16,219	42,340	28
No Solid Waste Disposal	191	58,368	0
Water Protection Cluster	**31,521**	**27,038**	**54**
Water Protection	28,727	29,832	49
Replace Drinking Water	923	57,636	2
Water Test	5,175	53,384	9
Surface Casing	294	58,265	1
No Surface Water	2,986	55,573	5
No Gas Storage	8,849	49,710	15
No Waste Injection	1,341	57,218	2
Optional Clause Cluster	**26,983**	**31,576**	**46**
Addendum	16,640	41,919	28
Free Gas	15,210	43,349	26
No Storage Tank	3	58,556	0
No Impoundment	425	58,134	1
No Compressor Station	761	57,798	1
No Hazardous Material	335	58,224	1
No Toxic Waste	208	58,351	0
No Small Firms	559	58,000	1
Legal Protection Cluster	**57,506**	**1,053**	**98**
Insurance, Indemnity	21,576	36,983	37
Offset Well	370	58,189	1
Reporting	55,982	2,577	96
Delay Payment	52,910	5,649	90
Performance Bond	30	58,529	0
Lessor Termination	21	58,538	0
No Royalty Deduction	24,618	33,941	42
Favorable to Producer Cluster	**48,118**	**10,441**	**82**
Subsurface Easement	3	58,556	0
Free Water Access	13,165	45,394	22
Royalty Deduction	31,636	26,923	54
Arbitration	27,521	31,038	47
Force Majeure	34,894	23,665	60
Pugh	2,562	55,997	4
Any Auxiliary Clause	**57,830**	**729**	**99**

Table 2: Descriptive Statistics of Royalty Rates and Average First Year Production

	Mean	SD	Min.	p25	p50	p75	Max	N
Royalty Rate (Full Sample)	0.14	0.03	0.13	0.13	0.13	0.15	0.21	58,559.00
Royalty Rate (Pass through)	0.14	0.02	0.13	0.13	0.13	0.16	0.20	8,445.00
Average FYP (trillion btu)	1.63	0.87	0.04	0.88	1.64	2.25	3.76	8,445.00
Average Depth (feet)	12,627.78	1,868.78	6,604.00	11,382.75	12,587.24	13,862.75	22,062.00	8,445.00

Note: The first row displays the average royalty rate for our full analytical sample of 58,559 leases that are used to estimate models 1 and 2. The second row display the average royalty rate of the 8,445 leases used to estimate model 3, with full details on how this sample was created in the online appendix. The bottom two rows are descriptive statistics of lease-level averages of first year production (FYP) and total drilled depth of all wells within a 2 km radius.

Table 3: Relationship Between Royalty Rates and Clause Clusters

	Surface Protection		Externalities		Water Quality		Optional		Legal Protection		Favor Producer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Cluster	0.003***	0.002***	0.001***	0.000	0.004***	0.004***	-0.003***	-0.004***	-0.000	-0.002***	0.004***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.141***	0.142***	0.143***	0.144***	0.142***	0.142***	0.145***	0.145***	0.144***	0.145***	0.140***	0.144***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.55	0.61	0.55	0.61	0.56	0.61	0.55	0.61	0.55	0.61	0.56	0.61
N	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396
County FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. These coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a proportion (i.e., the state minimum rate of 12.5 percent is coded as .125).

Table 4: Relationship Between Royalty Rates and Specific Clauses

(a)

	Gas Storage		No Surface Use		Development Plan		Water Test		No Royalty Deduction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Clause	0.007***	0.009***	-0.002***	-0.004***	0.003***	0.004***	0.006***	0.005***	0.003***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.143***	0.142***	0.144***	0.144***	0.143***	0.142***	0.143***	0.143***	0.142***	0.141***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.56	0.62	0.55	0.61	0.55	0.61	0.56	0.61	0.56	0.61
N	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396
County FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	N	Y	N	Y	N	Y	N	Y

(b)

	Water Replacement		Setback		No Pipes		Addt.l Infra.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Clause	0.004***	0.003***	0.002***	0.000	0.004***	0.003***	0.009***	0.008***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Intercept	0.144***	0.144***	0.144***	0.144***	0.143***	0.143***	0.143***	0.144***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.55	0.61	0.55	0.61	0.55	0.61	0.55	0.61
N	58,559	58,396	58,559	58,396	58,559	58,396	58,559	58,396
County FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	N	Y	N	Y	N	Y

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. These coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a proportion (i.e., the state minimum rate of 12.5 percent is coded as .125).

Table 5: First Year Production and Depth Instrument: First Stage

	(1)	(2)	(3)	(4)
Average Total Depth	0.0003***	0.0003***	0.0002***	0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)
F	2,925	1,277	393	220
R-squared	0.54	0.67	0.76	0.89
N	8,445	8,445	8,445	1,367
County FE	N	Y	Y	Y
Year FE	N	Y	Y	Y
Firm FE	N	N	Y	Y
Sample Within	2km	2km	2km	3km
Wells Before Lease	N	N	N	Y

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. This is the first stage for equation 3.

Table 6: First Year Production and Royalty Rates

	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Average FYP)	0.0479***	0.0490***	-0.0023	0.0072**	0.0056*	0.0224***	0.0219***	0.0142
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.007)	(0.009)
R-squared	0.05	0.05	0.51	0.51	0.54	0.54	0.47	0.47
N	8,445	8,445	8,445	8,445	8,440	8,445	1,364	1,367
County FE	N	N	Y	Y	Y	Y	Y	Y
Year FE	N	N	Y	Y	Y	Y	Y	Y
Firm FE	N	N	N	N	Y	Y	Y	Y
Sample Within	2km	2km	2km	2km	2km	2km	3km	3km
Wells Before Lease	N	N	N	N	N	N	Y	Y

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. The odd columns report OLS estimates for equation 3, and the even columns report 2SLS estimates for equation 3. Royalty rates are the dependent variable and are in the dataset as a proportion (i.e., the state minimum rate of 12.5 percent is coded as .125).

Table 7: First Year Production and Specific Clauses

(a)

	Gas Storage		No Surface Use		Development Plan		Water Test		No Royalty Deduction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
ln(Average FYP)	0.012	0.025^{*}	-0.036^{***}	-0.038^{***}	0.037^{***}	0.033^{*}	0.003	0.001	0.052^{***}	0.048^{**}
	(0.008)	(0.012)	(0.006)	(0.010)	(0.009)	(0.016)	(0.005)	(0.008)	(0.010)	(0.017)
R-squared	0.23	0.23	0.13	0.13	0.29	0.29	0.10	0.10	0.25	0.25
N	8,440	8,445	8,440	8,445	8,440	8,445	8,440	8,445	8,440	8,445
County FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

(b)

	Water Replacement		Setback		No Pipes		Addt.l Infra.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
ln(Average FYP)	-0.003	-0.004	0.002	0.011^{*}	0.008	0.007	0.000	0.002
	(0.002)	(0.003)	(0.003)	(0.005)	(0.006)	(0.009)	(0.002)	(0.003)
R-squared	0.03	0.03	0.13	0.13	0.12	0.12	0.03	0.03
N	8,440	8,445	8,440	8,445	8,440	8,445	8,440	8,445
County FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. The odd columns report OLS estimates for equation 3, and the even columns report 2SLS estimates for equation 3.