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, Revised January 2025

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, Revised January 2025
JEL No. D23, Q32, Q33, R11

ABSTRACT

Oil and gas lease negotiations provide mineral owners the opportunity to negotiate for both compensation and clauses that may protect their health and the enjoyment of their properties. We use optical character recognition to assemble the most comprehensive dataset to date on royalty rates and clauses in nearly 60,000 leases signed in Pennsylvania's Marcellus Shale. We leverage our data to produce three descriptive findings. First, we find a positive relationship between royalty rates and the prevalence of protective clauses. Second, we find that as development of the shale play progressed over time, royalty rates rose and leases became more likely to contain several protective clauses. Third, we find that royalty rates and the presence of protective clauses bear a weak relationship with the geologic productivity of nearby wells, explained by few firms competing in geographically segregated leasing markets. Some leases simultaneously containing higher royalty rates and more protective clauses suggests that there is a bargaining surplus in leasing markets, but the division of this surplus does not depend on the productivity of the mineral estate. Instead, it may reflect differing preferences, as well as differing negotiating skills, legal resources, and access to information. By documenting 43 clauses found in shale leases and their prevalence—more than double the number identified in past research—we provide critical information that can help mineral owners overcome information asymmetries and increase transparency and equity in leasing markets.

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

Pramod Manohar
University of Rochester
Department of Economics
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 led to income and employment gains in the United States (Hausman and Kellogg, 2015), but has also led to social and ecological disruptions in communities that host drilling (Mason et al., 2015; Black et al., 2021). 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 the contractual lease agreement that provides the oil and gas firm rights to develop their estate. Primary lease 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. They are compensated with one-time bonus payments and with royalties that are calculated as a percentage of the value of oil or natural gas production. Leases also contain auxiliary clauses that shape the terms of development. For example, one auxiliary clause may require the firm to place a fence around the well site, and another may require them to test nearby groundwater before and after drilling and provide replacement water if the tests show that it has been contaminated.

Auxiliary clauses present an opportunity for mineral owners to protect their health and safety, and influence how their properties are developed. But there has been relatively little research systematically documenting the types of clauses found in shale leases and their prevalence, and mineral owners may have little information about the universe of available clauses when negotiating new leases. The exceptions are Vissing (2015) and Timmins and Vissing (2022) who compile auxiliary clause data in one county in Texas to find that lower income, non-white, and non-English speaking individuals sign less protective shale leases. We build on these studies by creating the most comprehensive dataset on auxiliary clauses to date. Our novel dataset covers 58,559 leases signed between 2001 and 2016 that span Pennsylvania's Marcellus Shale. We read through 300 randomly selected leases to saturate a list of 43 unique auxiliary clauses, more than double the number identified in past research. We use optical character recognition (OCR), a type of computer program that converts images of text documents into machine-readable data, to identify the clauses within each lease.

Our data allow us to answer two descriptive research questions. Our first question asks: what auxiliary clauses have been signed in the Marcellus Shale and what are their prevalence? We find that 99 percent of the leases in our sample have at least one auxiliary clause, but the share of leases that contain any one of the 43 unique clause types ranges from less than one percent to 96 percent. Of the most prevalent clauses are those that require oil and gas firms to compensate landowners for damages (82 percent of leases), place wells a certain distance from a structure or water well (42 percent), and agree to a property development plan prior to starting development (37 percent). Other clauses that may be important are less prevalent, such as those that allow extraction of oil and natural gas from the subsurface but no access to the surface for placing well pads or other infrastructure (9 percent), and those that prohibit wastewater impoundments and compressor stations from being placed on the surface (1 percent).

Our second descriptive research question asks: what accounts for the variability in oil and gas leasing outcomes across individual leases? To answer this question, we connect the auxiliary clauses to data on lease locations and royalty rates. This link allows us to test three hypotheses about the determinants of leasing outcomes. Whether royalty rates and auxiliary clauses are positively or negatively associated is an empirical question. With a limited bargaining surplus in leasing markets, as would be the case with greater competition to sign leases between firms, mineral owners may be unable to add more clauses without reducing royalty rates. Conversely, with a larger bargaining surplus, royalty rates and auxiliary clauses may be positively associated if some mineral owners have the skills, knowledge, or financial resources to negotiate for leases that are more favorable all-around. We take this second explanation as our first hypothesis, because our data and past studies suggest that there are few oil and gas firms competing to sign leases in geographically segregated markets (Brown et al., 2016). Our empirical approach relies on year, municipality, and firm fixed effects, and finds a positive association between royalty rates and 27 of the 43 clause types that benefit mineral owners. Not a single protective clause type is negatively associated with royalty rates.

Our second hypothesis is that leases will grow stronger over time, as measured by increases in royalty rates and increases in the prevalence of protective clauses. 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 (Arnould and Grabowski, 1981; Mannering and Winston, 1995). Another explanation is that firms become less environmentally intensive over time, and therefore willing to include more protective clauses. Conditional on firm and municipality fixed effects, we find that the typical lease signed at the end of the study period in 2016 contains four more unique clauses than one signed near the beginning of the study period from 2001 to 2005. As widespread development of the Marcellus Shale progressed, ten protective clause types became more prevalent. For example, leases signed later in the study period are more likely to contain clauses that require oil and gas firms to test their water prior to drilling and provide replacement water in the event of contamination. Leases signed later in the study period are also more likely to prohibit surface use. Only two protective clause types become less prevalent, both with clear institutional explanations.[1] Royalty rates also increase over time, with the average rate remaining constant at the state-mandated minimum of 12.5 percent from 2001 to 2007, and increasing to a maximum of 16.8 percent for leases signed in 2010—one year after the commercial potential of the play was realized.

In addition to varying over time, royalty rates may vary across space if oil and gas firms offer higher rates in areas with greater geologic productivity. Our third hypothesis is that owners of more productive mineral resources will be able to negotiate for higher royalty rates and greater clauses. To test this hypothesis, we use oil and gas production reports to estimate geologic productivity at each lease as the average first year production of all wells within two and three kilometers (km) of the lease. Our lease-level data allows us to include fixed effects for year, municipality, and oil and gas firms. This approach controls for time-varying

[1] In Section 6.2 we show that a clause that prohibits placing pipelines on the property trends downward, likely due to the development of several high-volume pipelines and associated collection lines later in the development of the Marcellus. Setback requirements are clauses that specify a distance from the well and a home or water source. Setback requirements that are greater than the state-mandated minimum setbacks become virtually absent from leases after a legislative change in 2012 that extended the setback from 200 to 500 feet.

market conditions that may correlate with both production and royalty rates. We further isolate exogenous variation in production using the average vertical drilled depth of nearby wells as an instrumental variable. We find that a doubling of the average first year production of nearby wells leads to only a one to two percent increase in the royalty rate. For the typical lease in our sample with a 14.5 percent royalty rate, this means the doubling leads to at most a .3 percentage point increase—from a 14.5 percent rate to a 14.8 percent rate.

Our findings contribute to two bodies of empirical literature and have important practical implications. First, we contribute to several studies of the economic impacts of royalty income from shale drilling (Brown et al., 2019; Harleman and Weber, 2017; Ikonnikova et al., 2015; Herrnstadt et al., 2024). Most directly, we find a weak relationship between geologic productivity and royalty rates consistent with Brown et al. (2016). By using first year production of wells in close proximity to a lease, we concur with their estimates of the relationship between average royalty rates and average expected ultimate recovery at the county level. Our localized approach rules out the possibility of limited pass-through into royalty rates at the county level masking relationships between royalty rates and geologic productivity in many highly productive "sweet spots" within counties. We also extend their analysis by examining whether geological productivity affects negotiations for auxiliary clauses, and find that the quality of the mineral estate bears no relationship with the quality of auxiliary clauses.[2] Low pass-through into rates may be explained by firms exercising oligopsony power in geographically segregated leasing markets and uncertainty on the part of mineral owners about the quality of their mineral estate: the typical census tract and year has an average and median of only three firms competing to sign leases, and the typical census block group has only two firms. This means that mineral owners typically do not see a wide array of lease offers providing them with information on competitive royalty rates and available clauses.

Practically, we contribute by providing mineral owners and policymakers with the most comprehensive information to date on the universe of available clauses, thereby reducing information asymmetries across firms and mineral owners in an oligopsonistic market. To overcome these asymmetries, our comprehensive list of auxiliary clauses could inform the design of several low-cost state policies proposed in past research, including non-mandatory leasing templates, leasing guides and checklists, and digitized web records of lease agreements. In light of our other findings, such policies may help create more equitable leasing outcomes. Positive associations between clauses and royalty rates (within year, municipality, and firm) suggest that there is a bargaining surplus in leasing markets. Our finding of limited pass-through suggests that the division of this surplus does not depend on geologic productivity. It may instead depend on mineral owners' skills, knowledge, and financial and legal resources to negotiate for leases that are more favorable all-around.[3] In supplementary analysis, we find correlational evidence that individuals in census tracts that are whiter, more educated, and higher median income sign leases with more legal protections and with slightly higher royalty rates.

[2]With Brown et al. (2016), we find limited pass-through of geologic productivity into royalty rates. But pass-through of geologic productivity into total compensation necessarily occurs through the quantity effect.

[3]A second explanation is that there are different "types" of mineral owners within municipalities that differentially value the marginal cost of damages from drilling. We interpret our evidence in light of this explanation in Section 7 and conclude that it still points towards the utility of state policies to overcome information asymmetries.

Policies to address information asymmetries may be particularly beneficial for individuals belonging to disadvantaged groups if they have less knowledge or financial resources to hire legal counsel. The policies may have their greatest practical value for the leasing of private surface and subsurface lands to support renewable energy production and storage. While the supply of unleased domestic oil and gas acreage has declined in recent years, leasing of vacant land and farmland for wind and solar installations is anticipated to increase in the coming years (Jacquet, 2015; Winikoff and Parker, 2024; Spangler et al., 2024; U.S. Department of Energy, 2021; Lawrence Berkeley National Laboratory, 2024). Leasing of subsurface formations to extract brines containing lithium, a key input in batteries for electric vehicles and renewable energy storage, may also be poised to expand (Tscherning and Chapman, 2021; Vera et al., 2023; Farahbakhsh et al., 2024). Policies that supply landowners with information regarding available compensation and protections early on in a geographically concentrated wave of leasing could be important tools to create more equitable leasing outcomes across demographic groups.

We also contribute to a second literature on the elasticity of industrial revenues with respect to costly environmental regulation. We find little evidence that oil and gas firms respond to costlier *private* regulations by decreasing royalties to protect their revenues, as royalties and many clauses are positively related. Other studies have found that the oil and gas industry does respond to *public* regulations, suggesting that firms expect compliance with public regulations to be costlier than with auxiliary clauses. Boomhower (2019) finds that increases in insurance requirements to prevent well abandonment through bankruptcy reduce the development of marginal wells in Texas. Black et al. (2018) find that the introduction of a fee paid by oil and gas firms to the Pennsylvania government reduced leasing activity, and that firms passed through half of the fee to lessors that signed after its introduction. Using data similar to our own, future research could compare the elasticity of production to measures of public versus private regulation to confirm which is more efficient at engendering a given level of environmental protection. In a broader context, studies of the major United States environmental regulations have found that they have caused large reductions in industrial production and capital stock (Greenstone, 2002), and social costs in the form of permanent job losses (Bartik, 2015).

2 Production and Leasing in Pennsylvania's Marcellus Shale

The Marcellus Shale 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). Leasing in Pennsylvania's Marcellus began to pick up in 2005 and peaked in 2010 (Figure 1). Drilling followed quickly behind leasing, with widespread commercial drilling picking up after 2007. Drilling reached its peak in 2011, and by that year over 50,000 leases covering nearly 11,000 square miles had 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 firm. 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.

3 Determinants of Lease Negotiation Outcomes

In this section, we form three hypotheses about the determinants of oil and gas leasing outcomes. First, we consider a simple utility maximization framework to understand whether mineral owners will make trade-offs between negotiating for lease clauses and royalties. Second, we explore general literature on how firms and individuals learn about risk, and a narrower qualitative literature that documents communities learning about the localized impacts of shale development to explore whether mineral owners will adopt more protective auxiliary clauses over time. Third, we explore recent findings in the nonrenewable resource literature to consider whether owners of more productive mineral resources will negotiate for higher royalty rates and more protective clauses.

3.1 Negotiating for Payments and Auxiliary Clauses

In the case of the oil and gas industry in Pennsylvania, most mineral resources are owned by private individuals. Although government regulations attempt to protect private individuals from some negative externalities of the oil and gas industry (i.e., regulations on how wells are constructed are intended to prevent groundwater contamination), they do not address them all. 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 auxiliary clauses that they include and the payments that mineral owners receive. Consider Figure 2, where curve $S1$ is the indifference curve for the mineral owner, and curve $B1$ is the iso-profit curve for the oil and gas firm. In panel (a), the buyer (in this case the oil and gas firm) offers the seller (the mineral owner) a lease that includes $P1$ in payments and clauses that prevent $C1$ dollars of damages. If the seller (mineral owner) finds this amount acceptable, they will sign the lease. The seller may alternatively make a counter-offer. In panel (b), we rotate the seller's indifference curve to $S2$, corresponding to a seller with a greater preference for clauses (relative to curve $S1$). We keep the buyer's iso-profit curve $B1$ the same as in panel (a). The value of clauses increases to $C2$. If panel (a) represented a perfectly competitive environment with no bargaining surplus, in which the marginal benefit to the firm of causing environmental damage equals the marginal cost of the damage, the firm would be unwilling to add more clauses without reducing payments, which would fall from $P1$ to $P2$. This illustrates that in a perfectly competitive market with no bargaining surplus, mineral owners would make trade-offs between clauses and payments.

In practice, such an environment is unrealistic because there are a limited number buyers (oil and gas firms) competing to sign oil and gas leases in geographically segregated leasing markets (as noted by Brown et al. (2016)). This suggests that there is a bargaining surplus to be divided between the mineral owner and the firm. Panel (c) of Figure 2 illustrates a case where the mineral owner has greater negotiating skills, information about the array of clauses available to them, or information about the relative value of their minerals. Such a buyer may attempt to simultaneously negotiate for greater payments and more protective clauses by shifting their indifference curve northeast to $S3$. If there is a bargaining surplus and the buyer expects that they will still maximize profits versus leasing elsewhere, they will accept the offer at the combination $(P3, C3)$. This more realistic case motivates our first hypothesis:

Hypothesis 1: *Within a given year and municipality and across leases signed by the same oil and gas firm, royalty rates and the presence of auxiliary clauses that are protective of the environment and human health will be positively associated.*

A positive association could indicate that certain mineral owners have greater ability to negotiate, or greater information about the impacts of the shale industry, the array of clauses they could potentially include, or the net benefits of drilling experienced by the firm. A non-mutually exclusive explanation for the association is that there are different "types" of mineral owners that differentially value the marginal cost of damages from drilling. We interpret our evidence in light of both explanations in Section 7.

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 learned of their life-saving abilities through media coverage and word of mouth. This is analogous to Figure 2 panel (b)—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. Residents may also update their expectations about the profitability of shale development by learning about royalties and bonuses received by neighboring landowners. Outside of the spread of information 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:

Hypothesis 2: *Within a given municipality and across leases signed by the same oil and gas firm, leases will grow stronger over time as measured by increases in the prevalence of protective clauses and in 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 capture more rents than their counterparts in less profitable areas, in the form of higher royalty payments and more protective lease clauses. Figure 2 panel (c) illustrates that as drilling becomes more profitable, oil and gas firms should be willing to sign leases with higher royalty payments and more protective auxiliary clauses. This theoretical expectation motivates our third hypothesis:

Hypothesis 3: *Within a given year and municipality and across leases signed by the same oil and gas firm, owners of more productive mineral resources will negotiate higher royalty rates and more protective 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 one and two 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 inframarginal 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 and a 3 km radius of the lease. Our approach accounts for significant variation in geologic productivity within counties, and that observing no relationship between production and royalties at the county level may mask stronger relationships in highly-productive "sweet spots" within counties. Our data on auxiliary clauses also allow us to extend their analysis by examining whether geological productivity passes through as protective clauses.

4 Data

4.1 Leases

Our analysis leverages 185,968 lease documents which were filed and scanned at county courthouses in Pennsylvania. We purchased these data in 2017 as scanned PDF and image documents from the private data provider DrillingInfo. To identify the universe of clause types within the leases, we read through lease documents to saturate a list of clause types and variations of the language used by each drilling firm to describe them. Altogether, we identified a list of 43 unique clause types, 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 prohibiting an oil and gas firm from injecting 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 another. We selected at random and read through 275 leases before saturating a list of regular expressions associated with the 43 clause types (see Appendix A). Reading an additional 25 for a total of 300 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 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. This created a dataset containing the 172,616 documents as rows, with 43 binary variables indicating whether each unique clause type was present.

DrillingInfo also provides machine-readable data on lease polygons (the geospatial boundaries of each mineral lease), which contain 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, as some counties are not available in the polygon data.[4] Of the matched documents, we kept those that meet three inclusion criteria. First, we kept 63,082 that are the "lease" document type, and dropped lease amendments, extensions, and memos, which do not contain complete information on auxiliary clauses and royalty rates. Second, we dropped 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 dropped 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 after this year. We are left with 58,559 leases, which represents our full sample.

The leasing data allow us to answer our first descriptive research question, which aims to understand which auxiliary clauses have been signed in the Marcellus Shale. Of the 58,559 leases, 99 percent of these leases contain at least one auxiliary clause (1). Each of the auxiliary clauses is grouped into one of six clusters—surface protection, externalities, water protection, legal protection, favorable to producer, and optional. The first five follow and build upon Vissing's (2015) typology. The surface protection cluster includes clauses to minimize disturbances at the surface of the drilling site. This cluster includes a requirement to bury pipelines or prohibits placing pipelines on the surface, and requirements that oil and gas firms compensate mineral owners for any damage they cause to land, trees, or crops. Clauses in the externalities cluster protect mineral owners against noise pollution, traffic congestion, and solid waste disposal. Clauses in the water protection cluster are intended to prevent water contamination or replace drinking water in the case of contamination. The legal protection cluster includes clauses to indemnify mineral owners and protections on mineral owner compensation. The favorable to producer cluster includes clauses that benefit the drilling firm, such as free access to surface or groundwater on the property. We add an "optional" cluster to the Vissing typology to capture additional clauses such as those providing free household gas to mineral owners, restrictions on transferring the lease to a smaller firm, and restrictions on placing certain complementary infrastructure on the leased property.

[4]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.

4.2 Oil and Gas Well Data

Data on the characteristics, location, and production of unconventional oil and gas wells come from the Unconventional Natural Gas Well Geodatabase published by the (Carnegie Museum of Natural History, 2022). We use data on wells drilled between 2008 to 2021 to estimate the average first year production of wells within a 2 km and a 3 km radius of a lease. We use oil and gas well production report data to estimate a decline curve for each well and calculate the integral of oil and gas production (in trillions of British thermal units (Btu)) over the first 365 days of the well's lifespan. Estimated first-year production serves as a proxy for expected ultimate recovery. First year production is highly correlated with expected ultimate recovery, because daily natural gas production peaks immediately after drilling and gradually declines without further stimulation (Ikonnikova et al., 2015). Data from the US Energy Information Administration (2021) suggests that over 40 percent of a typical Marcellus well's ultimate recovery comes in its first year. Appendix B describes in detail the steps that we take to spatially connect wells to leases, and for each lease estimate average first year production and average vertical depth across all wells within 2 km and 3 km.

Table 2 presents descriptive statistics for the complete sample of 58,559 leases and a subset of 8,118 leases, which have at least ten wells within 2 km. This subset serves as our main sample for estimating pass-through. It shows that the typical lease is near wells that produce an estimated 1.6 trillion Btu in their first year. This is similar to past estimates from decline curve production models that show the typical Marcellus well producing 1.5 trillion Btu in its first year (Harleman, 2021; US Energy Information Administration, 2021). Our estimate is slightly higher because of a greater concentration of wells in places with more productive geology.

The 8,118 leases in the pass-through sample are near wells with an average vertical depth of 12,509 feet. Data on the total vertical depth of wells comes from the Pennsylvania Department of Conservation and Natural Resources (2022). We describe how we assign each lease an average vertical depth in Appendix section B. Figure 3 displays maps of wells and leases: panel (a) shows the 7,500 wells with complete vertical depth and first-year production data, while panel (b) shows the 58,559 leases in our full analytical sample

5 Empirical Approach

To answer our second, empirical research question that examines the determinants of oil and gas leasing outcomes, we adopt an empirical approach in three parts. First, to test whether a trade-off exists between protective clauses and royalty rates, we estimate the following model with ordinary least squares (OLS):

$$RoyaltyRate_{lmt} = \beta_0 + \beta_1 Clause_{lmt} + \tau_t + \lambda_m + \alpha + \varepsilon_{lmt} \tag{1}$$

Our outcome variable is $RoyaltyRate_{lmt}$, which is the royalty rate contained in lease l which was signed in municipality m and year t. Our explanatory variable of interest in the first specification is $Clause_{lmt}$, 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. This and all subsequent models are estimated with robust standard errors, two-way clustered by year and municipality.

In our preferred specifications we include year (τ_t), oil and gas firm (α), and municipality (λ_m) fixed effects to account for correlations between clauses and royalty rates that are not reflective of trade-offs within a given leasing negotiation. The year fixed effects account for temporal trends in royalty rates across the entire state. Oil and gas firm fixed effects account for firms' proprietary leasing templates, which lead to uniformity in the clauses across leases signed by a particular firm. Municipality fixed effects account for waves of leasing in a particular jurisdiction. Royalty rates and clauses may be correlated within a municipality due to factors such as firms' perceptions of geologic productivity, shared knowledge or concerns among landowners, simultaneous signing of many leases at locations like firehouses or community centers, or firms offering uniform leases in a given area to facilitate horizontal drilling across multiple estates.

Next, we examine how clauses evolve over time by plotting average royalty rates, the proportion of leases signed in year t containing a given auxiliary clause, and the total count of unique protective lease clauses included in each lease. Simple time plots are not an unbiased reflection of mineral owners negotiating over time if different temporal waves of leasing are driven by certain oil and gas firms or occur in certain municipalities. Therefore, we estimate the following model with firm and municipality fixed effects:

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

We also estimate equation 2 using binary variables for the presence of each auxiliary clause ($Clause_{lmt}$) as the outcome variable, as well as the total count of unique protective lease clauses included in the lease. We plot our estimates of β_τ, which represent the difference in the royalty rate (or likelihood of clause adoption) in year t relative to 2001 as the omitted year.

Our third model to examine the determinants of oil and gas leases builds off the approach for measuring the pass-through of geologic productivity into royalty rates from 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 on given lease, the time-varying price of natural gas, the time-varying market return on capital, and fixed expenditures to drill the lease. 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 expected ultimate recovery of the typical well in the county. We estimate a more localized relationship between royalty rates and geologic productivity:

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

Where $ln(FirstYearProduction_{lmt})$ is the estimated amount of oil and gas produced in the first year in trillion Btu. We calculate this amount using two approaches: a simple average and an Epanechnikov kernel weighted average. We apply each approach twice: once across all wells within a 2 km radius and once across all wells within a 3 km radius. In estimating equation 3, we include only leases with at least 10 wells drilled within the 2 km or 3 km radius.

As in the other two models, the lease-level data allow us to include fixed effects for year, municipality, and oil and gas firm. They effectively control for the time-varying market conditions that Brown et al. (2016) control for by estimating averaging annual interest rates for each county. We also estimate the model by replacing the outcome with binary variables indicating the presence of each auxiliary clause ($Clause_{lmt}$). This allows us to explore whether owners of more productive oil and gas resources can negotiate more favorable auxiliary clauses.

In a separate specification of model 3, we estimate the first year production variable using only wells within 3 km that were drilled *before* the lease was signed. This specification allows us to test whether the pass-through of productivity into royalty rates is stronger when mineral owners or firms learn about the productivity or royalty rates offered in a given geographic area.

Since $FirstYearProduction_{lmt}$ is estimated for each well and averaged across all wells within 2 km, it is subject to measurement error. To account for measurement error that would result in a downward bias in β_1, we utilize the average vertical depth of all wells within the radius as an instrument for average first year production. Vertical 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

Tables 3 through 5 display estimated models in the form of equation 1. The coefficients labeled "clause" are interpreted as the association between the clause cluster (in the first column of each panel) or individual clause type (in the subsequent columns) and the lease's royalty rate, expressed as a whole number (i.e., the state minimum rate of 12.5 percent is coded as 12.5). For example, the presence of any one of the eight clauses in the surface protection cluster is associated with a .21 percentage point average increase in royalty rates (Table 3, panel (a), column 1).

Two overarching findings arise from the three tables. First, none of the clusters or clause types favoring mineral owners show statistically significant negative correlations with royalty rates. Second, there is a positive and significant relationship between royalty rates and 27 of the 43 clauses that favor mineral owners. The results range from roughly a .2 percentage point higher royalty rate on average for a lease with a clause, to a 2.2 percentage point higher royalty rate. The strongest individual associations include a 2.18 percentage point increase in the royalty rate for leases that require a performance bond, which is a cash or surety bond that is released to the mineral owner if the firm fails to fulfill the terms of the lease (Table 5 panel (a), column 6). Presence of a clause allowing the lessor to terminate the lease is also strongly associated with a 1.52 percentage point higher rate on average (Table 5 panel (a), column 7). Strong positive associations also exist between royalty rates and clauses that prevent disposal of solid waste (+1.19 percentage points), toxic waste (+0.96), hazardous materials (+0.86), and the storage of natural gas (+0.67). Only two clauses favoring mineral owners take a negative sign, including a prohibition on surface use (Table 3 panel (a), column 2). This suggests that firms are willing to provide lower royalties to mineral owners that prohibit drilling, placing pipelines, or building roads on their surface.

Table 5 panel (b) displays the clauses favoring oil and gas producers. It shows the inverse of the other panels: negative and statistically significant associations with royalty rates. There significant negative associations with a clause allowing the firm to utilize the subsurface for the passage of wellbores or other uses, and a clause allowing the firm to deduct certain production costs prior to calculating royalty payments.

6.2 Trends of Royalty Rates and Clauses Over Time

First, we examine how the number of unique protective clauses included in the typical lease evolves over time. For this analysis, we exclude the five clause types that are favorable to producers. The left side of Figure 4 shows that the number of protective clauses in the typical lease increases from 7.4 in 2001 to a maximum of 14.2 in 2015. The right side of Figure 4 estimates equation 2 with the total count of protective clauses in the lease as the dependent variable. Conditional on municipality and oil and gas operator fixed effects, the typical lease signed in 2016 contains about four more unique clauses than one signed in 2001.

Next, we estimate equation 2 for each of the 43 unique clause types. Twelve of them display a clear upward trend over our study period, and only two display a clear downward trend. Appendix A lists the clause types that trend upward and downward. Figures 5 through 7 present twelve models that summarize our results. The left side displays simple bar graphs of the percentage of leases signed in a given year containing the clause. On the right side, the solid lines display the coefficients estimated by model 2, which are estimated with firm and municipality fixed effects. Dotted lines represent 95 percent confidence intervals.

Figure 5 panel (a) shows that less than one percent of leases signed in 2001 required baseline water quality testing prior to drilling. By the final two years of the study period, nearly half of all leases contained this requirement. Conditional on firm and municipality leasing characteristics, leases signed in 2016 are 36 percentage points more likely to contain the water testing clause than those signed in 2001. There are similar upward trends for clauses that require the firm to replace contaminated drinking water (panel (b)), require the firm to take precautions to protect ground and surface water (panel (c)), and prohibit the firm from injecting wastewater into the subsurface (panel (d)).

Figure 6 panel (a) shows that surface use prohibitions are virtually absent from leases signed before 2008, but leases signed in the last year of the study period are 37 percentage points more likely to contain surface use prohibitions than in the first year. Panel (b) displays results for a clause that requires the firm to come to a mutually agreeable development plan before drilling, one of the 29 clause types without a clear linear trend over the study period. That its prevalence begins to decline in the middle of the study period may be because surface use prohibitions begin to crowd out the need for a surface development plan.

Figure 6 panels (c) and (d) display the only two clause types that trend downward, and both have clear institutional explanations. Panel (c) is a clause that prohibits placing pipelines on the property. The downward conditional trend could be explained by the development of several high-volume pipelines and associated collection lines during the 2010s (State Impact Pennsylvania, 2018). For instance, the Mariner East pipelines were completed in 2014 to carry 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 may have become less likely to offer pipeline prohibitions. Panel (d) shows a weak downward trend in setback requirements that are greater than the state-wide minimum setbacks from residential buildings and water wells. This clause became virtually absent from leases after a legislative change in 2012 that extended the setback from 200 to 500 feet.[5]

Panels (a) and (b) of Figure 7 show that oil and gas firms may also be learning and updating the terms they offer over time, as two clauses that benefit the firm trend upward. The first requires certain types of legal conflicts to be settled in arbitration. The second is a force majeure clause, which holds the firm harmless for losses, financial burdens, or contractual obligations if an unexpected external event occurs.

Figure 7 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. Rates likely peaked in 2010 due to relatively high natural gas prices through 2010 that made drilling more profitable and leases more desirable (US Energy Information Administration, 2022). After 2010, royalty rates appear to fall with future expectations of natural gas prices as the shale gas boom led to elevated supply. Panel (d) shows that firms are 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.[6]

6.3 Pass-through of Geology into Royalty Rates 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 vertical drilled depth (our instrumental variable). As discussed in 5, we follow Brown et al. (2016) by using vertical depth as an instrument for production. The first stage results are in Table 6. The first two columns use estimates of first year production calculated using a simple average among leases within 2 km and 3 km, with the 3 km radius taking averages only of wells drilled before the lease was signed. The third and fourth columns use an Epanechnikov kernel weighted average to calculate first year production across wells within 2 km and 3 km. All columns display positive and statistically

[5]To construct the setback clause binary variable, we generate an indicator that equals one if the setback in the lease is greater than the statewide minimum setback, and zero otherwise. To account for the 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.

[6]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 firms in 2015 and 2016 (Cocklin, 2012). Companies 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.

significant coefficients on average total vertical depth, 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 7. OLS estimates are in the odd columns, and two-stage least squares (2SLS) estimates are in the even columns. In all columns, we observe a weak positive relationship between first year production and royalty rates. Notably, estimated first year production from wells within 2 km bears a more consistent relationship with royalty rates than additionally including wells from 2 to 3 km. This suggests firms make decisions to offer slightly higher royalty rates using hyper-local production estimates.

We use the largest and smallest 2SLS 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 1.12 to 1.68 percent ($= 0.70 \times 1.6\%$, $0.70 \times 2.4\%$) 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.

One limitation of our approach is that we do not observe bonus payments amounts in our data. Brown et al. (2016) estimate an elasticity of bonus payments with respect to royalty rates of +0.53, meaning that bonus payments and royalties are complements rather than substitutes. Given our finding of a 1 to 2 percent increase in royalty rates for a doubling of first year production, bonus payments would increase by only 0.5 to 1 percent. Even at a very high bonus payment of $10,000 per acre, a doubling of first year production would mean an increase of per acre $100 in bonuses captured by the mineral owner, which is still a small increase in the total value of production if $9 million for the typical well in the Marcellus (Harleman and Weber, 2017, p.286-287).

In Appendix C we extend our pass-through analysis to examine whether mineral owners are able to leverage greater geological productivity to negotiate for more protective clauses. We find that the quality of the mineral resource underlying the leased acreage is not positively associated with the quality of auxiliary clauses included in lease. This lends further support for the conclusion that mineral owners are not trading off royalty rates for more protective clauses. Limited pass-through of geological productivity into royalty rates does not appear to be offset by its pass-through in auxiliary clauses.

6.4 Supplemental Analysis and Sensitivity Tests

In Appendix D, we explore whether census tract-level measures of race, education, income are associated with royalty rates and the prevalence of protective clauses. The strongest associations indicate that leases signed where there are more white and college educated residents are more likely to include legal protections, and those signed where there are more white and high income residents have higher royalty rates. We also find different patterns of leasing between southwestern Pennsylvania counties with a longer history of drilling relative to northeastern counties. In Appendix E, we control for the number of mineral acres included in each lease, which may proxy for individual measures of wealth, but find that it bears little relationship with lease terms and does not affect our main results.

In our preceding analysis, we focus on learning among mineral owners (the suppliers of leased acreage) to explain variation in clauses and royalty rates across time and space. We use firm fixed effects to estimate trends in leasing that are independent of idiosyncratic firm characteristics such as leasing templates. In Appendix F we explore two potential demand-side drivers of leasing terms: the number of firms competing to sign leases in a given area and year, and whether the firm that signs the lease eventually drills a well on the estate. We find that both variables show weak and inconsistent relationship with clause prevalence and royalty rates.

7 Discussion and Policy Implications

In this section we summarize the results of our three hypothesis tests, and discuss their explanations and practical implications. Our first hypothesis predicted that royalty rates and the presence of protective auxiliary clauses would be positively associated, conditional on year, municipality, and firm fixed effects. Our evidence supports this hypothesis, as mineral owners simultaneously negotiate for higher royalty payments and more protective clauses. There is a positive and significant relationship between royalty rates and 27 of the 43 clauses that favor mineral owners. Notably, not a single clause type that benefits mineral owners is negatively associated with royalty rates.

Our second hypothesis predicted that mineral owners would sign leases with more protective clauses and higher royalty rates over time as the shale play develops. Our evidence supports the second hypothesis, as the typical lease signed at the end of the study period contains four more unique clauses than one signed at the start of the study period, and royalty rates and ten clauses that benefit mineral owners trend upward over the study period. Only two trend downward, and both have clear institutional explanations. Four clauses intended to protect ground and surface water trend upward, which is notable because concerns about water quality were dominant in the public discourse around fracking. The increase in surface use prohibitions is also notable, as it provides the most overarching security against localized environmental impacts. These upward trends are likely due to a combination of affected residents learning about specific industry practices and health concerns through word of mouth or the media, and firms becoming less environmentally intensive and willing to include more protective clauses. 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, making them more willing to sign surface use prohibitions.

Our third hypothesis predicted that owners of more productive mineral resources will be able to negotiate higher royalty rates and more protective clauses, conditional on year, municipality, and firm fixed effects. We find support for this hypothesis, but the effect is very small. A doubling of the average first year production of nearby wells leads to only a one to two percent increase in the royalty rate. For the typical lease in our sample with a 14.5 percent royalty rate, this means the doubling leads to at most a .3 percentage point increase—from a 14.5 percent rate to a 14.8 percent rate. This 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 as predicted by Brown et al. (2016). Their explanation for this result is that oil and gas firms are monopsonists in geographically segregated leasing markets, and uncertainty about productivity along with non-renegotiable rates lock in mineral owners to initial leases even if the resource proves to be very productive. In partial support of this explanation, we find evidence for oligopsonistic leasing markets (as opposed to monopsonistic): in the typical census tract and year there are only three firms signing leases, and in the typical census block group and year there are only two firms signing leases. Recent theoretical and empirical studies of oligopsonistic markets suggest that they benefit from lower input prices relative to more competitive markets (Rubens, 2023; Panagiotou and Stavrakoudis, 2017; Macedoni and Tyazhelnikov, 2024).

Together, our results indicate that oil and gas firms simultaneously make concessions by raising royalty rates and approving auxiliary clauses, but the concessions are not contingent upon the productivity of the mineral resource. This suggests that there is a bargaining surplus in leasing markets, and some mineral owners negotiate for leases that are more favorable all-around. One explanation for these more favorable leases is that some mineral owners have greater skills, knowledge, or financial and legal resources to negotiate leases. A second, non-mutually exclusive, explanation is that there are different "types" of mineral owners within municipalities that differentially value the marginal cost of damages from drilling. If this were the case, the observed disparity in lease quality could still mean that all mineral owners are signing leases that make them better off, resulting in an efficient "Coasean bargain" (Coase, 1960).

Coase (1960) shows that an efficient allocation of an externality between two parties can be achieved via private negotiations, under the assumptions that the property rights are well defined, there are zero transaction costs to negotiating, and the parties are symmetrically informed about the costs and benefits of the externality and potential agreements. The assumption of uniform information symmetry is unlikely to hold in our case: with few firms competing in limited geographic areas, mineral owners will not receive a wide array of lease offers to inform them of competitive royalty rates and available clauses. Moreover, in Appendix F we show that the degree of competitiveness among firms in local leasing markets bears a weak relationship with lease quality. Informational disadvantages may be particularly pronounced in certain communities, as in Appendix D we show that individuals in less white, less educated, and lower income census tracts sign leases with fewer legal protections and slightly lower royalty rates. Especially early in the development of a play before the learning we document can occur, less informed mineral owners (or those less capable of hiring legal counsel) may sign relatively weak leases that lock in for decades but fail to provide benefits that outweigh unforeseen damages. Such a case of asymmetric information can create *procedural environmental injustices*, or situations where certain groups have inequitable opportunities to participate in decisions about whether, where, and how an environmentally intensive industry can operate in their community (Skinner-Thompson, 2022). While some economic conceptions of environmental injustice focus exclusively on differences across racial and ethnic groups (Timmins and Vissing, 2022), broader policy definitions are inclusive of differences in income or disability (U.S. White House, 1994; EPA, 2024).

In practice, it is likely that variation across leases is driven both by differential willingness to accept drilling across "types" of mineral owners and by asymmetric information between firms and certain mineral owners. Regardless of the relative explanatory weight of these mechanisms, three low-cost state policies could be important tools to mitigate information asymmetries and create more equitable leasing outcomes for all mineral owners, regardless of their "type." These policies could be particularly important early in the development of new oil and gas plays, before learning among mineral owners occurs. They could also be important for future leasing of private surface and subsurface lands to support renewable energy production and storage, with leasing for wind and solar farms (Spangler et al., 2024; U.S. Department of Energy, 2021; Lawrence Berkeley National Laboratory, 2024) and direct lithium extraction (Tscherning and Chapman, 2021; Vera et al., 2023; Farahbakhsh et al., 2024) poised to increase domestically in the coming years. First, states could digitize the terms included in prior leases and make them available to landowners either online or at in-person access points at county courthouses (Vissing, 2015). Second, McFarland (2022) recommends that states provide landowners with leasing guides that include a checklist of clauses that can be incorporated in the lease. For future subsurface leasing, our comprehensive list of 43 auxiliary clauses could be used to inform the design of such guides. Third, state-sponsored leasing templates could provide a starting point for landowners without access to legal resources. Templates could include a set of auxiliary clauses that provide greater than minimum protections against externalities created by the industry. All three of these policy options promote transparency, and recent research has found that fostering more transparent leasing markets may lead to better outcomes. Covert and Sweeney (2023) leverage the quasi-random assignment of leasing rules in Texas to find that government organized auctions of leases for the development of state-owned minerals vastly outperform private negotiations conducted between firms and surface owners, with auctions leading to greater revenues for mineral owners due to a combination of higher bonus payments, higher royalty rates, and greater production. They also find that surface owners with more experience in negotiating with oil and gas firms, and those that own more land, perform best in term of securing higher levels of compensation. This aligns with our interpretation that more highly-skilled and knowledgeable mineral owners secure better leases.

8 Conclusion

We study the determinants of oil and gas leasing outcomes in Pennsylvania's Marcellus Shale and identify three key findings. First, royalty rates are positively associated with the prevalence of protective clauses. Second, royalty rates and the inclusion of environmentally protective clauses become more prevalent as the shale play develops over time. Third, a hyper-localized measure of geologic productivity shows only a weak positive relationship with royalty rates and the likelihood of protective clauses being included in leases, consistent with past findings at coarser geographic levels. Our results point to the role of limited competition and asymmetric information in geographically segregated leasing markets. Practically, our study addresses these asymmetries by documenting the most comprehensive set of auxiliary clauses to date. This information can guide the development of low-cost state policies—such as standardized leasing templates, lease guides, and digitized records of lease agreements—to increase transparency and promote equity 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.

Bartik, T. J. (2015). The social value of job loss and its effect on the costs of us environmental regulations. *Review of Environmental Economics and Policy*.

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.

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

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

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*.

Covert, T. R. and R. L. Sweeney (2023, March). Relinquishing riches: Auctions versus informal negotiations in texas oil and gas leasing. *American Economic Review 113*(3), 628–63.

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

EPA, U. (2024). Learn about environmental justice: Definitions. Accessed December 27, 2024.

Farahbakhsh, J., F. Arshadi, Z. Mofidi, M. Mohseni-Dargah, C. Kök, M. Assefi, A. Soozanipour, M. Zargar, M. Asadnia, Y. Boroumand, V. Presser, and A. Razmjou (2024). Direct lithium extraction: A new paradigm for lithium production and resource utilization. *Desalination 575*, 117249.

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.

Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures. *Journal of political economy 110*(6), 1175–1219.

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.

Herrnstadt, E., R. Kellogg, and E. Lewis (2024). Drilling deadlines and oil and gas development. *Econometrica 92*(1), 29–60.

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.

Jacquet, J. B. (2015). The rise of private participatio in the planning of energy projects in the rural united states. *Society and Natural Resources 28*(3), 231–245.

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.

Lawrence Berkeley National Laboratory (2024). Land-based wind market report. https://emp.lbl.gov/sites/default/files/2024-08/Land-Based%20Wind%20Market%20Report_2024%20Edition.pdf.

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

Macedoni, L. and V. Tyazhelnikov (2024). Oligopoly and oligopsony in international trade. *Canadian Journal of Economics/Revue canadienne d'économique 57*(2), 401–429.

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.

Panagiotou, D. and A. Stavrakoudis (2017). A stochastic production frontier estimator of the degree of oligopsony power in the us cattle industry. *Journal of Industry, Competition and Trade 17*, 121–133.

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*.

Rubens, M. (2023, September). Market structure, oligopsony power, and productivity. *American Economic Review 113*(9), 2382–2410.

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.

Skinner-Thompson, J. (2022). Procedural environmental justice. *Washington Law Review 97*, 399.

Spangler, K., E. A. Smithwick, S. Buechler, and J. Baka (2024). Just energy imaginaries? examining realities of solar development on pennsylvania's farmland. *Energy Research and Social Science 108*, 103394.

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.

Tscherning, R. and B. Chapman (2021). Navigating the emerging lithium rush: lithium extraction from brines for clean-tech battery storage technologies. *Journal of Energy and Natural Resources Law 39*(1), 13–42.

U.S. Department of Energy (2021). Solar futures study. `https://www.energy.gov/sites/default/files/2021-09/Solar%20Futures%20Study.pdf`.

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`.

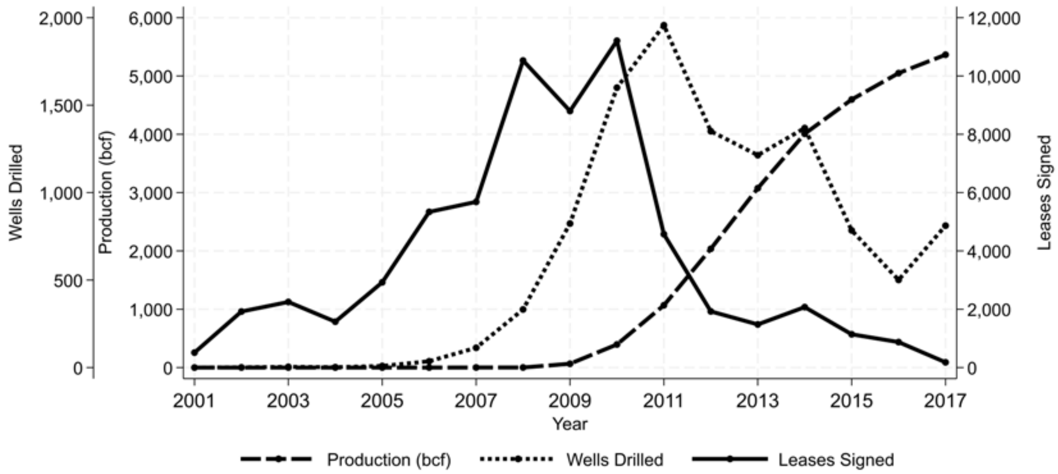
US Geological Survey (2022). Usgs releases new assessment of gas resources in the marcellus shale. `https://pubs.usgs.gov/fs/2011/3092/pdf/fs2011-3092.pdf`.

U.S. White House (1994). Executive order no. 12898, federal actions to address environmental justice in minority populations and low-income populations. Published in 59 Fed. Reg. 7629, February 16, 1994.

Vera, M. L., W. R. Torres, C. I. Galli, A. Chagnes, and V. Flexer (2023). Environmental impact of direct lithium extraction from brines. *Nature Reviews Earth and Environment 4*(3), 149–165.

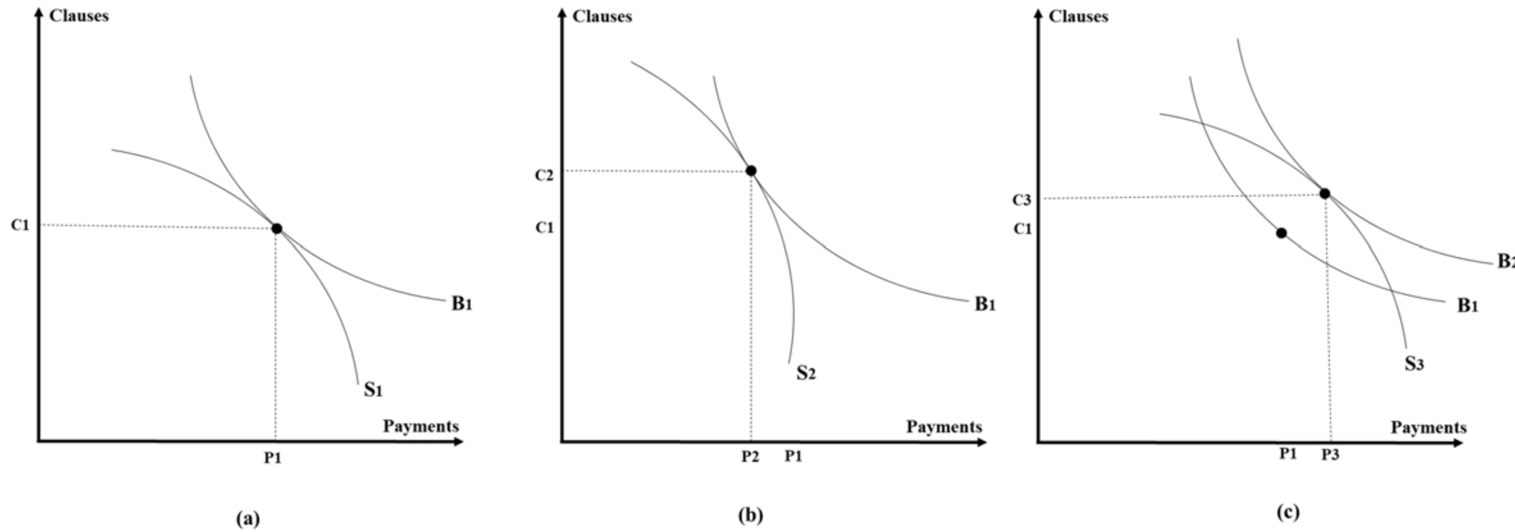
Vissing, A. (2015). Private contracts as regulation: A study of private lease negotiations using the texas natural gas industry. *Agricultural and Resource Economics Review 44*(2), 120–137.

Winikoff, J. B. and D. P. Parker (2024). Farm size, spatial externalities, and wind energy development. *American Journal of Agricultural Economics 106*(4), 1518–1543.



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), and data on leasing are from the private data provider DrillingInfo.

Figure 1: Unconventional Natural Gas Drilling and Production in Pennsylvania's Marcellus



Notes: Panel (a): $S1$ is the indifference curve for the seller (mineral owner) and $B1$ is the iso-profit curve for the buyer (oil and gas firm). The negotiated lease results in $P1$ in payments and clauses that prevent $C1$ dollars of damages.

Panel (b): The indifference curve ($S2$) is rotated to correspond to a mineral owner with a greater preference for clauses (relative to curve $S1$), and the oil and gas firm's iso-profit curve stays the same at $B1$. The value of clauses increases to $C2$, and payments fall to $P2$.

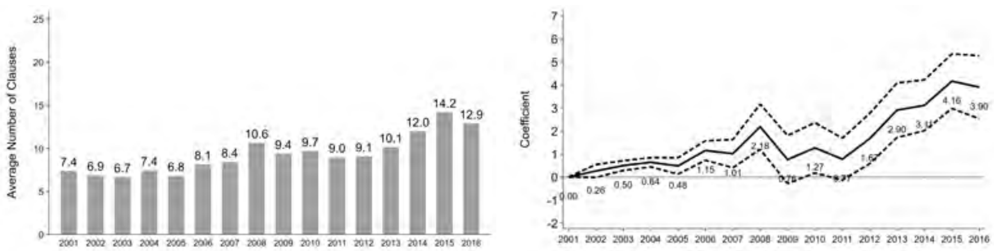
Panel (c): The indifference curve ($S3$) is shifted to the northeast to illustrate a mineral owner with greater negotiating skills or information. If the oil and gas firm maximizes expected profits at $B2$ versus leasing elsewhere, they will accept the offer at the combination $(P3, C3)$.

Figure 2: Lease Negotiations Between Companies and Mineral Owners

(a) Map of Wells

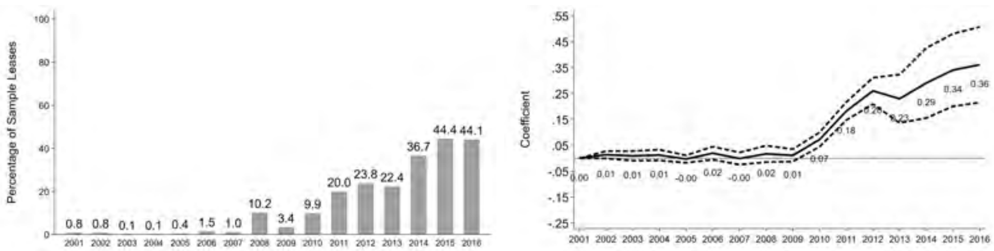
(b) Map of Leases

Figure 3: Map of Sample Wells and Leases

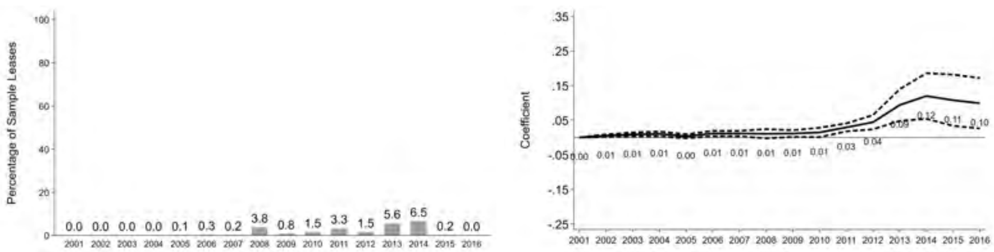


Note: On the left is a simple bar graphs of the average number of clauses signed in the typical lease in a given year. On the right, the solid line is the estimated OLS coefficient from equation 2, and the dotted lines are 95% confidence intervals, calculated using robust standard errors that are two-way clustered by year and municipality. The coefficients represents the difference in the average number of clauses signed in the typical lease in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

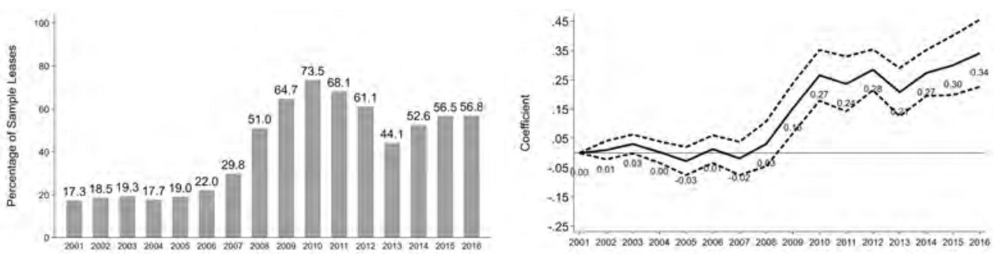
Figure 4: Average Number of Auxiliary Clauses Over Time



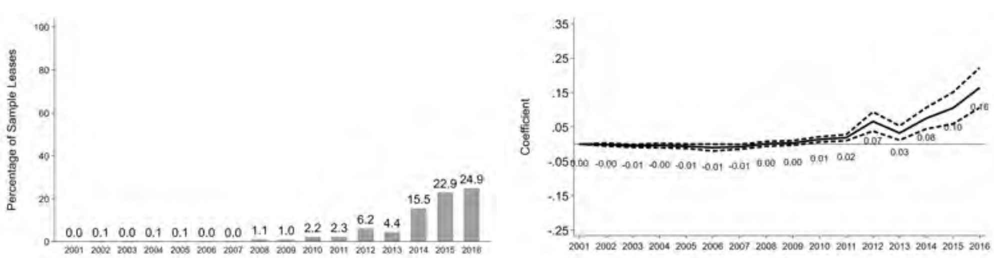
(a) Water Testing



(b) Water Replacement



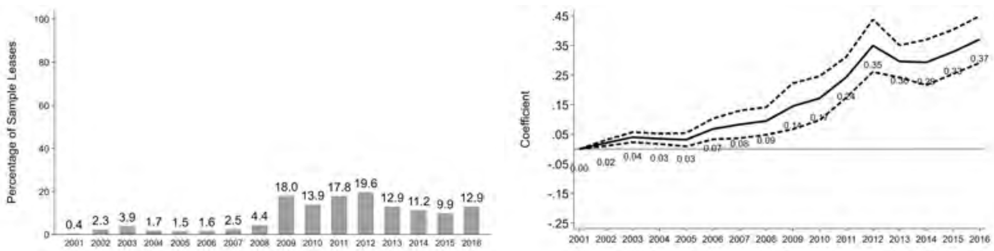
(c) Water Protection



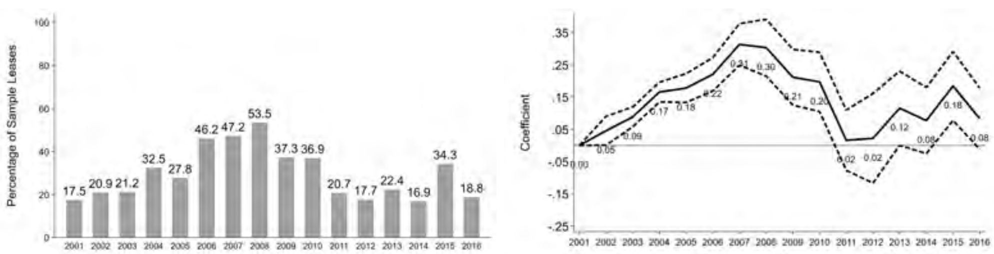
(d) No Wastewater Injection

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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

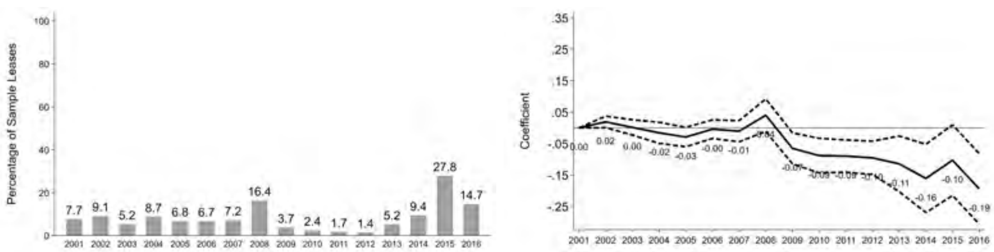
Figure 5: The Prevalence of Select Clauses Over Time



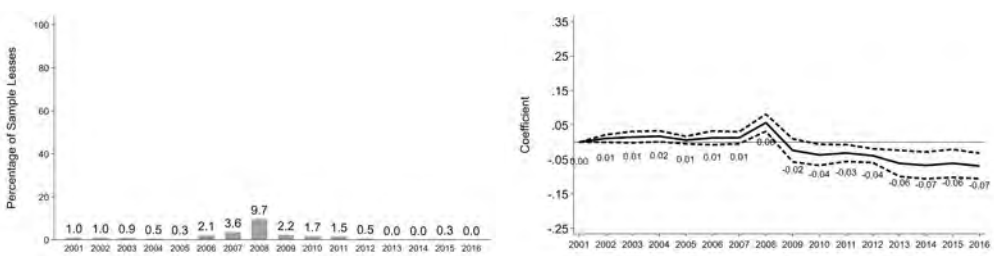
(a) No Surface Use



(b) Development Plan



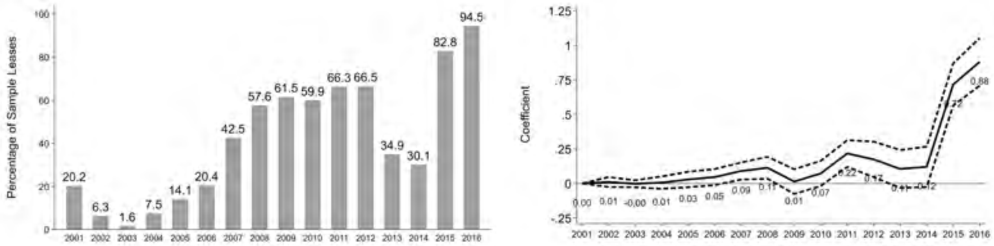
(c) No Pipelines



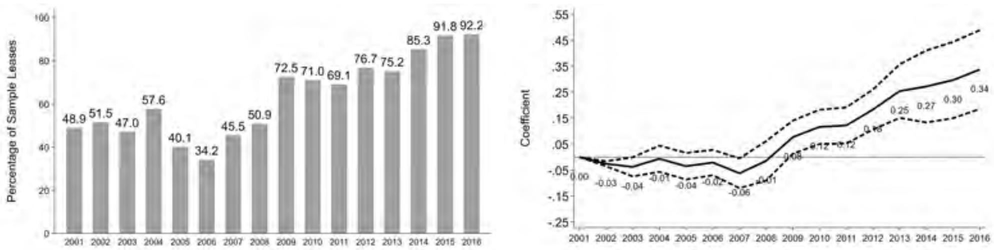
(d) Greater than Minimum Setback

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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

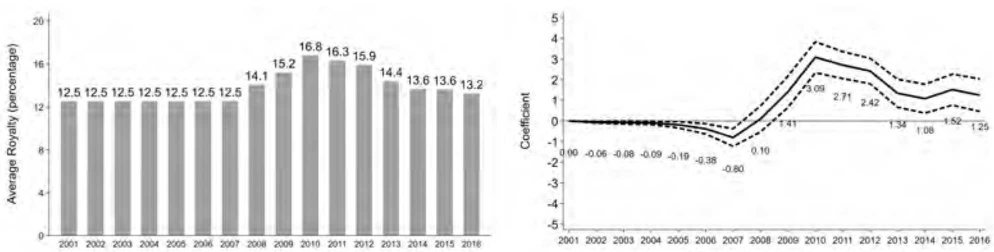
Figure 6: The Prevalence of Select Clauses Over Time (cont.)



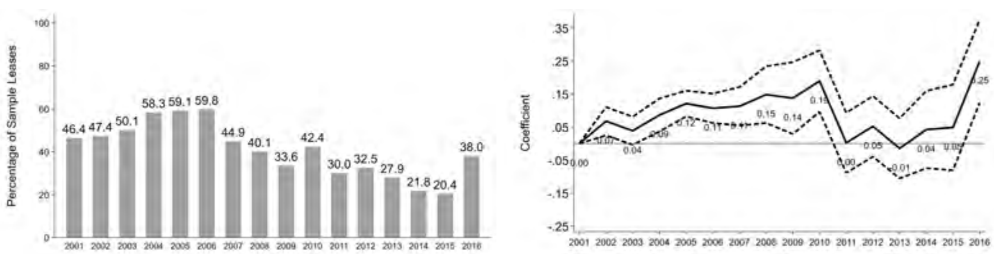
(a) Arbitration



(b) Force Majeure



(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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

Figure 7: 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	**16,314**	**42,245**	**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	**57,517**	**1,042**	**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
Pugh	2,562	55,997	4
Favorable to Producer Cluster	**47,815**	**10,744**	**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
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)	14.37	2.53	12.50	12.50	12.50	15.00	21.00	58,559.00
Royalty Rate (Pass Through)	14.47	2.45	12.50	12.50	12.50	16.00	20.50	8,118.00
Average FYP (Trillion BTU)	1.59	0.85	0.04	0.86	1.59	2.16	3.76	8,118.00
Average Depth (Feet)	12,508.66	1,764.44	6,604.00	11,364.19	12,517.00	13,759.74	20,403.14	8,118.00

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 6,839 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 Specific Clauses

(a) Surface Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.21*	-0.15	0.40**	0.16**	0.30	-0.37	0.28	0.17	0.14
	(0.09)	(0.10)	(0.12)	(0.05)	(0.15)	(0.34)	(0.17)	(0.10)	(0.20)
Intercept	14.25***	14.38***	14.33***	14.24***	14.36***	14.37***	14.35***	14.33***	14.36***
	(0.05)	(0.01)	(0.01)	(0.03)	(0.00)	(0.00)	(0.01)	(0.02)	(0.01)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Externalities Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.03	0.62*	0.16**	0.10	0.27*	-0.03	0.05	0.27**	1.19*
	(0.07)	(0.27)	(0.04)	(0.25)	(0.12)	(0.13)	(0.47)	(0.09)	(0.51)
Intercept	14.35***	14.36***	14.37***	14.37***	14.27***	14.37***	14.37***	14.29***	14.37***
	(0.03)	(0.00)	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.02)	(0.00)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Surface Protection Cluster, (2) No Surface Use, (3) Minimize Disturbance, (4) Damage Compensation, (5) Development Notification, (6) Road Restriction, (7) No Pipelines, (8) Bury Pipelines, (9) Well Spacing. The Surface Protection Cluster (1) is not inclusive of Damage Compensation (4) due to little variation in the cluster with its inclusion.

Panel (b): (1) Externalities Cluster, (2) Environmental Protection, (3) Noise Restriction, (4) Working Hours, (5) Development Plan, (6) Setback, (7) Traffic, (8) Fences and Gates, (9) No Solid Waste Disposal

Table 4: Relationship Between Royalty Rates and Specific Clauses

(a) Water Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Clause	0.31**	0.25*	0.42*	0.46***	-0.43	0.29	0.67***	0.52**
	(0.08)	(0.09)	(0.14)	(0.07)	(0.21)	(0.14)	(0.13)	(0.13)
Intercept	14.20***	14.25***	14.36***	14.33***	14.37***	14.35***	14.27***	14.36***
	(0.04)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.67	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Optional Clause Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Clause	0.37***	0.34**	0.32*	0.51***	0.38**	0.86**	0.96***	0.44*
	(0.09)	(0.09)	(0.13)	(0.10)	(0.11)	(0.26)	(0.23)	(0.19)
Intercept	14.27***	14.28***	14.37***	14.36***	14.37***	14.36***	14.37***	14.37***
	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Water Protection Cluster, (2) Water Protection, (3) Replace Drinking Water, (4) Water Test, (5) Surface Casing, (6) No Surface Water, (7) No Gas Storage, (8) No Waste Injection

Panel (b): (1) Optional Clause Cluster, (2) Free Gas, (3) No Storage Tank, (4) No Compressor Station, (5) No Impoundment, (6) No Hazardous Material, (7) No Toxic Waste, (8) No Small Firms

Table 5: Relationship Between Royalty Rates and Specific Clauses

(a) Legal Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.54***	0.50***	0.51**	0.09	0.03	2.18*	1.52***	0.50***	0.66***
	(0.09)	(0.09)	(0.16)	(0.07)	(0.08)	(0.85)	(0.17)	(0.11)	(0.08)
Intercept	14.05***	14.19***	14.37***	14.28***	14.34***	14.37***	14.37***	14.16***	14.34***
	(0.05)	(0.03)	(0.00)	(0.06)	(0.06)	(0.00)	(0.00)	(0.04)	(0.00)
R-squared	0.67	0.67	0.66	0.66	0.66	0.66	0.66	0.67	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Favorable to Producer Cluster

	(1)	(2)	(3)	(4)	(5)	(6)
Clause	-0.07	-1.05**	0.06	-0.14*	-0.11	0.05
	(0.04)	(0.34)	(0.10)	(0.06)	(0.08)	(0.06)
Intercept	14.43***	14.37***	14.36***	14.45***	14.42***	14.34***
	(0.02)	(0.00)	(0.02)	(0.02)	(0.03)	(0.02)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Legal Protection Cluster, (2) Insurance, Indemnity, (3) Offset Well, (4) Reporting, (5) Delay Payment, (6) Performance Bond, (7) Lessor Termination, (8) No Royalty Deduction, (9) Pugh.

Panel (b): (1) Favorable to Producer Cluster, (2) Subsurface Easement, (3) Free Water Access, (4) Royalty Deduction, (5) Arbitration, (6) Force Majeure

Table 6: First Year Production and Depth Instrument: First Stage

	(1)	(2)	(3)	(4)
Average Total Vertical Depth	0.0001***	0.0002**	0.0001***	0.0001***
	(0.0000)	(0.0001)	(0.0000)	(0.0000)
F	34	17	32	28
R-squared	0.86	0.90	0.84	0.88
N	8,095	1,273	8,095	9,480
Spatial Approach	Buffer	Buffer	Kernel	Kernel
Sample Within	2km	3km	2km	3km
Wells Before Lease	N	Y	N	N

*Note:** $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. This is the first stage for equation 3. All columns include year, oil and gas operator, and municipality fixed effects.

Table 7: 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)	0.004	0.024*	0.020	0.016	0.003	0.024*	0.005	0.021
	(0.005)	(0.010)	(0.009)	(0.008)	(0.005)	(0.011)	(0.007)	(0.013)
R-squared	0.56	0.56	0.56	0.57	0.56	0.56	0.53	0.53
N	8,095	8,118	1,273	1,289	8,095	8,118	9,480	9,499
Spatial Approach	Buffer	Buffer	Buffer	Buffer	Kernel	Kernel	Kernel	Kernel
Sample Within	2km	2km	3km	3km	2km	2km	3km	3km
Wells Before Lease	N	N	Y	Y	N	N	N	N

*Note:** $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are 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 whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5). All columns include year, oil and gas operator, and municipality fixed effects.

Appendix For Online Publication

A Auxiliary Clauses

Altogether we identified 43 clauses in six clusters, which we list here. The clauses were captured with 360 unique regular expressions that we identified by reading leases. Along with this appendix, we have compiled a database that documents each of the 360 regular expressions. 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. The list below also indicates the 12 clause types that trend upward and the 2 that trend downward from our model 2 results, described in Section 6.2 of the main text.

Surface Protection Cluster: clauses that minimize disturbances at the surface of the drilling site.

1. **No Surface Use**: prohibits the firm from accessing or drilling wells on the mineral owner's surface (upward trend).
2. **Minimize Disturbance**: requires the firm to minimize disturbance in terms of soil or surface erosion.
3. **Damage Compensation**: requires the firm to either pay for or repair damage to the surface, including filling pits, reseeding, and restoring soil.
4. **Development Plan**: requires the firm to mutually agree to a development plan with the mineral owner prior to drilling or accessing the surface.
5. **Road Restriction**: prohibits the firm from constructing an access road (upward trend).
6. **No Pipelines**: prohibits the firm from placing pipelines on the property (downward trend).
7. **Bury Pipelines**: requires the firm to bury all pipelines placed on the property.
8. **Well Spacing**: requires the firm to space wells greater than a specified minimum distance apart.

Externalities Cluster: clauses that protect mineral owners against negative externalities.

1. **Environmental Protection**: indicates that the firm will take precautions to protect the environment.
2. **Noise Restriction**: requires the firm to limit noise to specified decibels at all times or during specified hours.
3. **Working Hours**: prohibits the firm from working on the property or engaging in certain activities like drilling during certain hours.
4. **Development Notification**: requires the firm to notify the mineral owner a specified number of days before beginning specified development activities (upward trend).
5. **Setback**: prohibits the firm from drilling a well within a specified distance of a structure, groundwater source, or surface water (downward trend).
6. **Traffic**: restricts the path, number, or speed of vehicles on or near the property.
7. **Fences and Gates**: requires the firm to build a fence or gate around the well pad or other infrastructure.

8. **No Solid Waste Disposal**: prohibits the firm from disposing of solid waste on the property.

Water Protection Cluster: clauses meant to prevent water contamination.

1. **Water Protection**: indicates that the firm will take precautions to protect groundwater or surface water supplies (upward trend).
2. **Replace Drinking Water**: requires the firm to replace drinking water if it is found to contaminate a drinking water source (upward trend).
3. **Water Test**: requires the firm to test groundwater or surface water supplies at specified times (upward trend).
4. **Surface Casing**: requires the firm to use surface casing over specified depths to protect groundwater supplies.
5. **No Surface Water**: prohibits the firm from using surface water supplies for production activities.
6. **No Gas Storage**: prohibits the firm from storing natural gas on the property or in the leased subsurface.
7. **No Waste Injection**: prohibits the firm from injecting wastewater on the property or into the leased subsurface (upward trend).

Optional Clause Cluster: additional options that are typically included in an addendum to an oil and gas firm's standard leasing document.

1. **Free Gas**: requires the firm to provide a specified amount of free natural gas to the mineral owner.
2. **No Storage Tank**: prohibits the firm from placing a hydrocarbon or wastewater storage tank on the property.
3. **No Impoundment**: prohibits the firm from placing a wastewater impoundment on the property.
4. **No Compressor Station**: prohibits the firm from placing a compressor station on the property.
5. **No Hazardous Material**: prohibits the firm from storing any hazardous materials on the property (upward trend).
6. **No Toxic Waste**: prohibits the firm from storing any toxic waste on the property (upward trend).
7. **No Small Firms**: prohibits the firm from transferring the lease to a firm below a specified size in terms of capital.

Legal Protection Cluster: clauses that indemnify mineral owners or affect how mineral owners are compensated.

1. **Insurance, Indemnity**: holds the mineral owner harmless for losses or financial burdens associated with the oil and gas production on their property (upward trend).
2. **Offset Well**: restricts the firm from using the property to drill an exploratory, offset well that may not generate significant royalty revenues.

3. **Reporting**: requires the firm to provide drilling, production, and other important reports to the mineral owner.

4. **Delay Payment**: requires the firm to pay a delay rental to the mineral owner prior to drilling if the firm wants to keep the lease in effect.

5. **Performance Bond**: requires the firm to set aside a cash or surety bond that is released to the mineral owner if the firm fails to fulfill the terms of the lease.

6. **Lessor Termination**: allows the lessor to terminate the lease at specified times, such as prior to the end of the primary term.

7. **No Royalty Deduction**: prohibits the firm from deducting specified production costs or other costs from revenues prior to calculating royalty payments.

8. **Pugh**: prohibits the firm from holding the lease past the primary term by producing from part of a pooled unit that is not the mineral owner's estate.

Favorable to Producer Cluster: clauses that benefit the drilling firm.

1. **Subsurface Easement**: allows the firm to utilize the subsurface for the passage of wellbores that may not originate or produce from the lessor's property, or allows for other subsurface uses like the construction of drainage systems.

2. **Free Water Access**: allows the firm to freely use surface water supplies for production activities.

3. **Royalty Deduction**: allows the firm to deduct specified production costs or other costs from revenues prior to calculating royalty payments.

4. **Arbitration**: forces specified types of legal conflicts between the firm and mineral owner to be settled in arbitration (upward trend).

5. **Force Majeure**: holds the firm harmless for losses, financial burdens, or contractual obligations if an unexpected external event were to occur (upward trend).

B Data and Geoprocessing Steps

This appendix describes the individual steps that we took to attach the $ln(FirstYearProduction_{lmt})$ independent variable and the average vertical depth variable to each lease in our sample, both of which are required to estimate model 3 described in Section 5.

B.1 Sample of Unconventional Oil and Gas Wells

The Unconventional Natural Gas Well Geodatabase published by the Carnegie Museum of Natural History (2022) compiles the various oil and gas well reports published by the Pennsylvania Department of Environmental Protection (DEP). We downloaded the data as of the second quarter of 2021, and begin with the shapefile dataset, which includes variables across multiple DEP reports, including permit, spud, production, waste, and compliance report. We exported the attribute table in ArcGIS as a .dbf file and imported it into Stata.

The file contains 22,597 rows, each corresponding to a unique DEP permit number. We dropped 19 wells that are classified as observation wells, coalbed methane wells, or storage wells, which are outside of the scope of our study. We also dropped 9,524 wells that were permitted but never drilled. This leaves us with data on the locations and attributes of 13,054 unconventional oil and gas wells, uniquely defined by their permit numbers.

B.2 Estimation of First Year Production For Each Well

The Unconventional Natural Gas Well Geodatabase published by the Carnegie Museum of Natural History (2022) also provides a .csv file of 17,099 unique wells drilled between 1974 to 2021. Most of the wells are observed multiple times, with a unique row in the data corresponding to an individual production report submitted by the oil and gas company to the DEP. Each row includes a unique identifier for the well, the duration in days of production, and the quantity of natural gas and oil produced within the production period. Over the period 1974 to 2021, companies submitted production reports at least annually, and after 2015 the companies submitted reports as frequently as monthly.

We began with 895,383 reports and dropped 209,390 reports that were (1) missing the number of reported production days, (2) missing the quantity of natural gas produced in MCF, (3) associated with wells that have no production report within their first year of production, (3) associated with wells that were plugged, abandoned, proposed but never materialized, and not drilled (as reported by the company), (4) associated with wells that are in periods of regulatory inactive status but report positive production, because by definition inactive wells cannot report production. (5) associated with wells that are not categorized as "unconventional," (6) reports for which the production quantity was not for an individual well but was averaged over a group of nearby wells, and (7) reports associated with wells drilled before widespread Marcellus production began in 2008.

With the remaining 685,993 reports, we converted the production variables (volume of oil and natural gas) to millions of British thermal units (mmBtu), and then summed the oil and gas production together to obtain a combined production figure. Next, we created a variable capturing the cumulative number of days that the well had been producing from its first day in production through to the last day of the report. We also created a variable that captures the total cumulative lifetime production in mmBtu for each well from its first day in production through to the last day of the report. We ran a separate regression through the origin for each individual well of total cumulative production on a linear term, a squared term, and a cubed term for the total number of days that the well had been producing through to the end of the reporting period. To obtain an estimate of aggregate first year production for each well, we multiplied the estimated coefficients on the linear, squared, and cubed days by 365 days, and added the three products together. We estimated first year production measures this way because for vast majority of wells we do not observe 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).

Altogether, we were able to project first year production for 10,825 wells that meet our seven sample requirements and have more than three reports (which is needed for the regression to run). The first row of Table B1 displays descriptive statistics for our first year production estimates for the 10,825 wells. We compared the estimates to those of Harleman (2021) (who uses a nearly identical methodology) and data from the US Energy Information Administration (2021). These sources indicate that in their first year, the typical Marcellus well produces around 1.5 trillion Btu in its first year (see Harleman (2021), Online Appendix Page 9), with a maximum of between 4 and 5 trillion Btu (see Harleman (2021), Online Appendix Page 20). Based on these benchmarks, the first row of Table B1 suggests that our model vastly overestimates production for some wells, which is largely due to measurement error or misreporting of production in the DEP production reports. Therefore, we drop wells for which our model estimates a first year production greater than the 90th percentile of 4.20 trillion Btu, and proceed with 9,743 wells.

The second row of Table B1 displays the distribution of average FYP for the 9,743 wells. These wells have a mean first year production of 1.45 trillion Btu, a median of 1.18 trillion Btu, and a maximum of 4.20 trillion Btu. These estimates are similar to the estimates found in Harleman (2021) and from the US Energy Information Administration (2021).

Table B1: Descriptive Statistics of Average First Year Production (Well-Level Dataset)

	Mean	SD	Min.	p25	p50	p75	Max	N
Average FYP (trillion Btu)	1.97	3.32	0.00	0.68	1.33	2.57	168.52	10,825.00
Average FYP (trillion Btu)	1.45	1.05	0.00	0.62	1.18	2.12	4.20	9,743.00

*Note:*The first row displays the distribution of average FYP all 10,621 wells for which we are able to project FYP. The second row displays the distribution of average FYP after removing the top 10 percent of observations.

B.3 Depth Instrument

Data on the total depth of each well, our instrument variable, comes from the Pennsylvania Department of Conservation and Natural Resources (2022). The DCNR provides data on depth as part of their EDWIN database, which is a subscription service. We downloaded data on 8,073 wells and their depth for the Marcellus shale play, as well as the Geneseo, Utica, and Rhinestreet plays (which lie above or beneath the Marcellus). EDWIN contains a variable "DeepestProducingDepth," which is the deepest vertical depth at which the oil and gas driller produces oil or gas. We merged the 9,743 wells with estimated first year production with wells with non-missing vertical depth data, and we are left with 7,500 wells with complete data for both variables in model 3 as described in Section 5.

B.4 Creation of a Lease and Unconventional Well Near Table

We imported all 13,054 unconventional oil and gas well locations back into ArcGIS and created a "near table" using the Generate Near Table tool. The near table calculated the distance between each lease polygon boundary and each well, for all lease and well pairs that are within 3 km of each other. The near table contains 982,091 unique lease-well pairs within 3 km and 469,080 unique lease-well pairs within 2 km.

Next, we merged well records for 7,500 wells with complete estimated first year production and depth data with the near table. When keeping lease-well pairs using only wells with complete data, we retained 597,928 unique lease-well pairs within 3 km and 202,731 unique lease-well pairs within 2 km.

We collapsed the lease-well dataset into a lease-level dataset, taking lease-level averages and Epanechnikov kernel-weighted averages of all wells within 2 km (or 3 km) radius for the total depth and average first year production variables. We only retain leases that have at least 10 wells within the radius, to limit the effect of outlying depth or first year production estimates biasing the results. For the within 2 km analysis, this yields a sample of 8,118 leases. For the within 3 km analysis, the sample is 9,499 leases. Finally, for the within 3 km that attaches wells to leases if the well was drilled before the lease was signed, the sample is 1,289 leases.

C Pass-through of Geology into Clauses

As an extension of our pass-through analysis, in Table C1 through Table C3 we examine whether mineral owners are able to leverage greater geological productivity to negotiate for more protective clauses. The tables display 2SLS estimates from models in the form of equation 3. The tables display only one positive and statistically significant coefficient for clauses that benefit mineral owners—the clause that requires firms to replace contaminated drinking water (Table C1 panel (a), column 3). The coefficient indicates that a doubling of first year production leads to a $.014 \times .7 = 0.0098$ or a roughly 1 percentage point increase in the likelihood of the contaminated drinking water clause being present in the lease. All other clause types that benefit mineral owners have an insignificant relationship or even a negative relationship with first year production. Conversely, the favorable to producer clause cluster has a positive relationship with first year production.

The estimates show that the quality of the mineral resource underlying the leased acreage is not positively associated with the quality of auxiliary clauses included in the lease. This lends further support for the conclusion that mineral owners are not trading higher royalty rates for more protective clauses. Limited pass-through of geological productivity into royalty rates does not appear to be offset by its pass-through in auxiliary clauses.

Table C1: First Year Production and Clauses

(a) Surface Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln(Average FYP)	0.0225	-0.0116	-0.0050	-0.0053	0.0158	-0.0066	-0.0078	0.0217	-0.0435**
	(0.0464)	(0.0251)	(0.0487)	(0.0322)	(0.0265)	(0.0099)	(0.0242)	(0.0434)	(0.0150)
R-squared	0.29	0.19	0.17	0.33	0.14	0.14	0.17	0.42	0.46
N	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118

(b) Externalities Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Average FYP)	-0.1079**	0.0118	-0.0124*	0.0103	0.0158	0.0009	-0.0159	-0.0007
	(0.0374)	(0.0083)	(0.0049)	(0.0673)	(0.0180)	(0.0036)	(0.0565)	(0.0050)
R-squared	0.29	0.09	0.06	0.35	0.18	0.05	0.29	0.13
N	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. The table reports 2SLS estimates for equation 3. All columns include year, oil and gas operator, and municipality fixed effects, and use the 2km buffer approach.

Panel (a): (1) Surface Protection Cluster, (2) No Surface Use, (3) Minimize Disturbance, (4) Damage Compensation, (5) Development Notification, (6) Road Restriction, (7) No Pipelines, (8) Bury Pipelines, (9) Well Spacing. The Surface Protection Cluster (1) is not inclusive of Damage Compensation (4) due to little variation in the cluster with its inclusion.

Panel (b): (1) Externalities Cluster, (2) Environmental Protection, (3) Noise Restriction, (4) Development Plan, (5) Setback, (6) Traffic, (7) Fences and Gates, (8) No Solid Waste Disposal

Table C2: First Year Production and Clauses

(a) Water Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Average FYP)	0.0132	-0.0077	0.0141*	-0.0384*	-0.0187	-0.0163	0.0783	-0.0143
	(0.0587)	(0.0428)	(0.0068)	(0.0184)	(0.0117)	(0.0318)	(0.0419)	(0.0201)
R-squared	0.34	0.36	0.09	0.12	0.22	0.19	0.27	0.14
N	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118

(b) Optional Clause Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Average FYP)	0.0323	0.0367	-0.0013	0.0090	-0.0078	0.0023	-0.0023
	(0.0481)	(0.0461)	(0.0007)	(0.0058)	(0.0045)	(0.0038)	(0.0080)
R-squared	0.41	0.41	0.29	0.03	0.14	0.14	0.13
N	8,118	8,118	8,118	8,118	8,118	8,118	8,118

Note: $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. The table reports 2SLS estimates for equation 3. All columns include year, oil and gas operator, and municipality fixed effects, and use the 2km buffer approach.

Panel (a): (1) Water Protection Cluster, (2) Water Protection, (3) Replace Drinking Water, (4) Water Test, (5) Surface Casing, (6) No Surface Water, (7) No Gas Storage, (8) No Waste Injection

Panel (b): (1) Optional Clause Cluster, (2) Free Gas, (3) No Storage Tank, (4) No Compressor Station, (5) No Impoundment, (6) No Toxic Waste, (7) No Small Firms

Table C3: First Year Production and Clauses

(a) Legal Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln(Average FYP)	-0.1194***	-0.0987**	-0.0017	-0.0010	0.0105	0.0009	0.0014	-0.0415	0.0352
	(0.0356)	(0.0362)	(0.0093)	(0.0107)	(0.0214)	(0.0022)	(0.0012)	(0.0422)	(0.0496)
R-squared	0.27	0.24	0.20	0.26	0.35	0.42	0.02	0.32	0.09
N	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118	8,118

(b) Favorable to Producer Cluster

	(1)	(2)	(3)	(4)	(5)
ln(Average FYP)	0.0707*	-0.0147	0.0956***	0.0520	0.0282
	(0.0341)	(0.0457)	(0.0280)	(0.0350)	(0.0393)
R-squared	0.45	0.48	0.38	0.54	0.50
N	8,118	8,118	8,118	8,118	8,118

Note: $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. The table reports 2SLS estimates for equation 3. All columns include year, oil and gas operator, and municipality fixed effects, and use the 2km buffer approach.

Panel (a): (1) Legal Protection Cluster, (2) Insurance, Indemnity, (3) Offset Well, (4) Reporting, (5) Delay Payment, (6) Performance Bond, (7) Lessor Termination, (8) No Royalty Deduction, (9) Pugh.

Panel (b): (1) Favorable to Producer Cluster, (2) Free Water Access, (3) Royalty Deduction, (4) Arbitration, (5) Force Majeure

D Analysis of Community Demographics and Lease Terms

D.1 Community Demographics: Methods

In section 6.1, we find a positive association between royalty rates and clauses, which suggests that certain mineral owners have greater ability to negotiate, or greater information about the impacts of the shale industry. In this Appendix we explore whether community-level measures of race, education, income, and region are associated with royalty rates and the prevalence of clause clusters.

To attach the Census data to our lease polygons, we spatially join the centroids of the leases to the census tract polygon it falls in. We obtained census tract polygons for the year 2010 from the US Census TIGER/Line Shapefile web interface, and data on demographic characteristics for 2010 census tracts from the `the Urban Institute (2021)`. Altogether, we were able to join nearly all of our leases (58,236 in total) to their census tract data. Table D1 shows that communities with leasing are largely white, with a mean and median share of non-white residents of only three percent, and a standard deviation of four percentage points. It also shows that the mean and median share of college-educated individuals (with an associate's degree or higher) is around 15 percent, with a standard deviation of seven percentage points. The average of tract-level Median HHI in real 2021 US dollars is $54,000 in our sample, with a standard deviation of $12,250.

To understand the associations between these three demographic variables and lease terms, we estimate variations of the following model via ordinary least squares:

$$RoyaltyRate_{lmt} = \beta_0 + \beta_1 ShareNonWhite_{lmt} + \beta_2 ShareEdu_{lmt} + \beta_2 MedianHHI_{lmt} + \tau_t + \lambda_m + \alpha + \varepsilon_{lmt} \quad \text{(D1)}$$

We also replace the outcome with a binary variable that indicates the presence of at least one clause in a cluster ($Clause_{lmt}$). We control for year fixed effects (τ_t), municipality fixed effects (λ_m) and firm fixed effects (α), so that our estimates represent associations between the census tract variables that are not due changes in leasing patterns over time, nor idiosyncratic leasing templates that vary across firms and municipalities. In our reported results, we multiply our coefficients by one standard deviation for each of the three independent variables (.04 for $ShareNonWhite_{lmt}$, .07 for $ShareEdu_{lmt}$, and 12.25 for $MedianHHI_{lmt}$ which is in thousands of dollars). In doing so, β_1 through β_3 can be interpreted as the change in the royalty rate (or change in the likelihood of the presence of a clause cluster) for a one standard deviation change in the associated demographic variable.

In our main specification, we identify the relationship between the three demographic variables and lease terms by leveraging *within* municipality variation, operationalized by the municipality fixed effects. However, we may expect different leasing outcomes *across* municipalities due to different historic trends in oil and gas development across regions. For example, southwestern counties in Pennsylvania experienced a long history of conventional natural gas development, dating back to the early 20th century. We hypothesize that leases signed in the southwestern counties in Pennsylvania (Allegheny, Armstrong, Beaver, Butler, Greene, Fayette, Indiana, Lawrence, Washington, Westmoreland, and Somerset) could have more protective terms and higher royalty rates than those in the north-central and northeastern counties due to historical learning about the oil and gas industry. We explore this by creating a variable $Southwest_{lmt}$ that equals one for leases signed in one of the ten southwestern counties, and in one model substitute it for the municipality fixed effects.

We note that we cannot be certain that the relationships that we observe between demographics and lease terms at the census tract level are causal relationships. Although we control for municipality fixed effects to account for broad geographic and community characteristics that may drive leasing outcomes, we cannot be confident that we observe all of the relevant census tract characteristics (e.g., public goods, spatial proximity of homes, real estate values, etc.) that may affect our outcomes. Moreover, we do not have data on individual mineral owners' race, education status, and income. We cannot rule out that aggregating demographic measures to the census tract level masks relationships between demographics and lease terms that could be identified if researchers observed individuals as the unit of analysis, along with a rich array of demographic controls that describe them. An individual-level analysis of leasing terms and demographics in the Marcellus represents an opportunity for future research.

D.2 Community Demographics: Results

Tables D2 and D3 display our results of our community demographics and lease term analysis. The relationship between a one standard deviation increase in each of the three demographic characteristics and clause clusters are generally weak in magnitude, but in the intuitive direction. Where there are significant relationships, as the share of non-white residents increases by one standard deviation, the likelihood of the lease containing one of the legal protection clauses decreases by 2.6 percentage points (Table D2, column 5). In the same column, as the share of college education increases by one standard deviation, the likelihood of the lease containing one of the legal protection clauses increases by 2.2 percentage points.

Table D2, column 7 reveals that as the share of non-white residents in the tract increases by one standard deviation, royalty rates decrease by .139 percentage points on average, and as median household income in the tract increases by one standard deviation royalty rates increase by .155 percentage points.

Table D3 displays fairly large differences in lease terms across regions. Leases signed in southwestern counties are .21 percentage points less likely to contain surface protections, .23 percentage points less likely to contain legal protections for mineral owners, .13 percentage points more likely to contain clauses that favor producers, and carry a royalty rate that is 1.56 percentage points lower. Taken together, leases in the southwestern counties are less environmentally-protective and more favorable to producers. This contrasts with our hypothesis that mineral owners in counties with a longer history of drilling would learn to sign stronger leases. Instead, it may be the case that more familiarity with the industry leads to greater comfort signing leases quickly without negotiating for as many clauses.

D.3 Community Demographics: Tables

Table D1: Additional Descriptive Statistics

	Mean	SD	Min.	p25	p50	p75	Max	N
Share Non-White	0.03	0.04	-0.03	0.02	0.03	0.04	0.46	58,236.00
Share With College Degree	0.15	0.07	0.04	0.12	0.14	0.17	0.70	58,236.00
Median HHI (000s)	54.64	12.25	13.36	48.43	53.73	58.97	148.29	58,217.00
Unique Firms in Tract and Year	3.41	2.56	1.00	2.00	3.00	4.00	18.00	58,236.00
Leased and Drilled	0.14	0.34	0.00	0.00	0.00	0.00	1.00	58,236.00

*Note:*The table displays descriptive statistics for the entire sample of leases that could be connected to census polygons.

Table D2: Relationship Between Tract Characteristics and Lease Terms

	Surface Protection (1)	Externalities (2)	Water Quality (3)	Optional (4)	Legal Protection (5)	Favor Producer (6)	Royalty (7)
Share Non-White	0.010*	0.003	-0.005	-0.008	-0.026**	-0.004	-0.139*
	(0.001)	(0.000)	(0.001)	(0.001)	(0.006)	(0.000)	(0.215)
Share College Degree	0.009	-0.002	0.013	-0.000	0.022**	0.004*	-0.043
	(0.001)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.024)
Median HHI (000s)	-0.008	-0.016**	0.002	0.008	0.009	-0.002	0.155**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Intercept	0.566***	0.620***	0.507***	0.251***	0.529***	0.822***	13.899***
	(0.008)	(0.000)	(0.029)	(0.009)	(0.020)	(0.006)	(0.107)
R-squared	0.36	0.38	0.49	0.59	0.32	0.52	0.61
N	58,055	58,055	58,055	58,055	58,055	58,055	58,055

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and county are in parentheses. All columns include year, oil and gas operator, and county fixed effects. The coefficients represent OLS estimates from equation D1. In columns 1 through 6 the dependent variable is the presence of at least one clause in a cluster, defined as a binary variable. In column 7, royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5). The first three coefficients in each column are standardized to reflect a one standard deviation change in the independent variable.

Table D3: Relationship Between Region and Lease Terms

	Surface Protection	Externalities	Water Quality	Optional	Legal Protection	Favor Producer	Royalty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Southwest	-0.205***	-0.011	0.009	0.032***	-0.238***	0.129***	-1.555***
	(0.006)	(0.006)	(0.005)	(0.004)	(0.006)	(0.004)	(0.028)
Intercept	0.677***	0.669***	0.536***	0.294***	0.711***	0.777***	14.993***
	(0.010)	(0.010)	(0.007)	(0.007)	(0.010)	(0.007)	(0.043)
R-squared	0.28	0.29	0.44	0.54	0.29	0.41	0.58
N	58,055	58,055	58,055	58,055	58,055	58,055	58,055

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors in parentheses. All columns include year and oil and gas operator fixed effects, and controls for the share of college education, share non-white, and median HHI at the census tract level. The coefficients represent OLS estimates from equation D1. In columns 1 through 6 the dependent variable is the presence of at least one clause in a cluster, defined as a binary variable. In column 7, royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5). The first three coefficients in each column are standardized to reflect a one standard deviation change in the independent variable.

E Sensitivity to Acreage Control

Another explanation for the variation in clauses and royalty rates across time and space is that mineral owners with larger properties will negotiate differently than those with smaller plots. In our baseline specifications, we do not control for the number of mineral acres leased in each contract because acreage could proxy for wealth. Wealth may be part of the causal mechanism linking higher royalty rates to greater clause prevalence, if wealth is associated with the skills or resources to negotiate for more favorable leases. If this were the case, controlling for acreage could mask a relationship between royalty rates and clause prevalence in Tables 3 through 5. Similarly, it could mask the relationship between geologic productivity and lease terms in Table 7.

Nevertheless, here we test the sensitivity of our results to the inclusion of a control for acreage, a variable that is fully observed for our main sample of 58,559 leases and has a median of 15.2 acres and a mean of 62 acres. Below, we reproduce all tables from the main text with the inclusion of a control for hundreds of acres. The results show nearly identical relationships between royalties and clause prevalence and between royalties and geologic productivity as those in the main text. They also show that acreage bears little relationship with nearly all common leasing terms. While at first glance this might seem to suggest that wealthier mineral owners do not sign stronger leases, acreage may not be a suitable proxy to estimate the impact of negotiating skills and resources on lease terms. Over a higher range of acreage, a mineral owner may be willing to accept a weaker lease because the large plot does not contain their primary domicile (e.g., land held for hunting or investment purposes), any surface activity could occur far from structures, or the total payments from the large plot could be very high even with a low royalty rate. In additional results not reported here, including a term for hundreds of acres squared produced similar results.

Table E1: Relationship Between Royalty Rates and Specific Clauses

(a) Surface Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.21^{*}	-0.15	0.40^{**}	0.16^{**}	0.30	-0.37	0.28	0.17	0.13
	(0.09)	(0.10)	(0.12)	(0.05)	(0.15)	(0.35)	(0.17)	(0.10)	(0.20)
Hundred Acres	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.07	0.07
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Externalities Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.03	0.61^{*}	0.15^{**}	0.10	0.27^{*}	-0.03	0.05	0.27^{**}	1.18^{*}
	(0.07)	(0.27)	(0.04)	(0.25)	(0.12)	(0.13)	(0.47)	(0.09)	(0.51)
Hundred Acres	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06
	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Surface Protection Cluster, (2) No Surface Use, (3) Minimize Disturbance, (4) Damage Compensation, (5) Development Notification, (6) Road Restriction, (7) No Pipelines, (8) Bury Pipelines, (9) Well Spacing

Panel (b): (1) Externalities Cluster, (2) Environmental Protection, (3) Noise Restriction, (4) Working Hours, (5) Development Plan, (6) Setback, (7) Traffic, (8) Fences and Gates, (9) No Solid Waste Disposal

Table E2: Relationship Between Royalty Rates and Specific Clauses

(a) Water Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Clause	0.31**	0.25*	0.42*	0.46***	-0.43	0.29	0.67***	0.52**
	(0.08)	(0.09)	(0.14)	(0.07)	(0.21)	(0.14)	(0.13)	(0.13)
Hundred Acres	0.07	0.07	0.08	0.08	0.08	0.08	0.05	0.08
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.67	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Optional Clause Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Clause	0.37***	0.34**	0.32*	0.51***	0.38**	0.85**	0.95***	0.44*
	(0.09)	(0.09)	(0.13)	(0.10)	(0.11)	(0.26)	(0.22)	(0.19)
Hundred Acres	0.07	0.08	0.07	0.08	0.07	0.06	0.06	0.07
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Water Protection Cluster, (2) Water Protection, (3) Replace Drinking Water, (4) Water Test, (5) Surface Casing, (6) No Surface Water, (7) No Gas Storage, (8) No Waste Injection

Panel (b): (1) Optional Clause Cluster, (2) Free Gas, (3) No Storage Tank, (4) No Compressor Station, (5) No Impoundment, (6) No Hazardous Material, (7) No Toxic Waste, (8) No Small Firms

Table E3: Relationship Between Royalty Rates and Specific Clauses

(a) Legal Protection Cluster

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Clause	0.54***	0.50***	0.50**	0.09	0.03	2.10*	1.52***	0.50***	0.66***
	(0.09)	(0.09)	(0.16)	(0.07)	(0.08)	(0.87)	(0.17)	(0.11)	(0.08)
Hundred Acres	0.06	0.06	0.07	0.07	0.07	0.05	0.07	0.07	0.07
	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
R-squared	0.67	0.67	0.66	0.66	0.66	0.66	0.66	0.67	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334	58,334

(b) Favorable to Producer Cluster

	(1)	(2)	(3)	(4)	(5)	(6)
Clause	-0.07	-1.05**	0.06	-0.14*	-0.11	0.05
	(0.04)	(0.34)	(0.10)	(0.06)	(0.08)	(0.06)
Hundred Acres	0.07	0.07	0.07	0.07	0.07	0.07
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
R-squared	0.66	0.66	0.66	0.66	0.66	0.66
N	58,334	58,334	58,334	58,334	58,334	58,334

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. The coefficients represent OLS estimates from equation 1. Royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).

Panel (a): (1) Legal Protection Cluster, (2) Insurance, Indemnity, (3) Offset Well, (4) Reporting, (5) Delay Payment, (6) Performance Bond, (7) Lessor Termination, (8) No Royalty Deduction, (9) Pugh

Panel (b): (1) Favorable to Producer Cluster, (2) Subsurface Easement, (3) Free Water Access, (4) Royalty Deduction, (5) Arbitration, (6) Force Majeure

Table E4: First Year Production and Depth Instrument: First Stage, Acres Control

	(1)	(2)	(3)	(4)
Average Total Vertical Depth	0.0001***	0.0002**	0.0001***	0.0001***
	(0.0000)	(0.0001)	(0.0000)	(0.0000)
Hundred Acres	-0.0012	0.0005	-0.0010	-0.0004
	(0.0010)	(0.0022)	(0.0012)	(0.0003)
F	17	9	16	14
R-squared	0.86	0.90	0.84	0.88
N	8,095	1,273	8,095	9,480
Spatial Approach	Buffer	Buffer	Kernel	Kernel
Sample Within	2km	3km	2km	3km
Wells Before Lease	N	Y	N	N

*Note:** $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. This is the first stage for equation 3. All columns include year, oil and gas operator, and municipality fixed effects.

Table E5: First Year Production and Royalty Rates, Acres Control

	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Average FYP)	0.004	0.024*	0.020	0.016	0.003	0.024*	0.005	0.020
	(0.005)	(0.010)	(0.009)	(0.009)	(0.005)	(0.011)	(0.007)	(0.013)
Hundred Acres	-0.000	-0.000	0.000	0.000	-0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
R-squared	0.56	0.56	0.56	0.57	0.56	0.56	0.53	0.53
N	8,095	8,118	1,273	1,289	8,095	8,118	9,480	9,499
Spatial Approach	Buffer	Buffer	Buffer	Buffer	Kernel	Kernel	Kernel	Kernel
Sample Within	2km	2km	3km	3km	2km	2km	3km	3km
Wells Before Lease	N	N	Y	Y	N	N	N	N

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are 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 whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5). All columns include year, oil and gas operator, and municipality fixed effects.

F Demand-Side Explanations

In most of our analysis, we largely focused on learning among mineral owners (the suppliers of leased acreage) to explain variation in clauses and royalty rates across time and space. We use firm fixed effects to estimate trends in leasing that are independent of idiosyncratic firm characteristics such as leasing templates. In this appendix, we explore two important demand-side drivers of leasing terms.

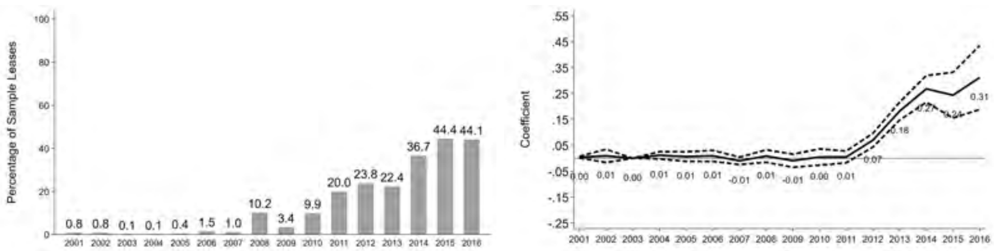
The first driver recognizes that leasing markets can be competitive in a geographic area. With more firms competing to sign leases in a given area and year, mineral owners might stand to benefit from higher royalties and stronger clauses. In Table D1 we display descriptive statistics for the variable “Unique Firms in Tract and Year” which is a count of the number of firms that signed leases in a given census tract and year, which we calculated for the entire sample of leases that could be connected to census polygons. On average, there are 3.41 firms (median of 3 firms) competing to sign leases in a given tract and year. At the more spatially disaggregated census block group, this number is only 2.24 firms on average (median 2 firms).

The second driver recognizes that some firms signing leases do not intend to drill wells themselves, but rather lease many contiguous plots and sell them to firms that intend to drill wells. If this were the case, we expect that some firms would value greater uniformity across leases that they sign, if it is easier to liquidate a uniform block of leases. In Table D1 we display descriptive statistics for the variable “Leased and Drilled.” This variable captures the concept of leasing with the intent to drill. It assigns a value of one to leased acreage where the nearest well within 2 km is drilled by the same firm that signed the lease, and zero if the firms differ or the lease was not drilled during the study period. Table D1 shows that about 14 percent of the sample is leased and drilled by the same firm.

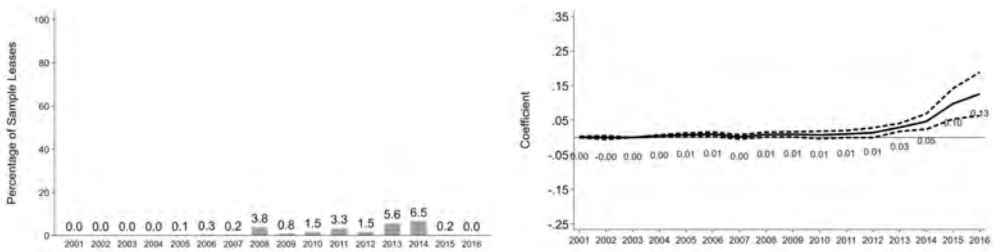
To examine these relationships, we run the same model as in equation D1, replacing the census tract demographic variables with the two demand-side variables. Table F1 displays the results. The tables show that the number of firms leasing in a given census tract and year has a weak and inconsistent relationship with clause prevalence and royalty rates. For instance, an additional firm is associated with about a 2 percentage point greater likelihood of containing a clause in the water quality cluster, and only a 0.06 percentage point higher royalty rate.

A well leased with the intention to drill also appears to bear little relationship with clause clusters and royalty rates. The exception is with the cluster capturing clauses favorable to the producer, with its prevalence increasing by roughly 4 percentage points if the lease is signed by the company that drills it.

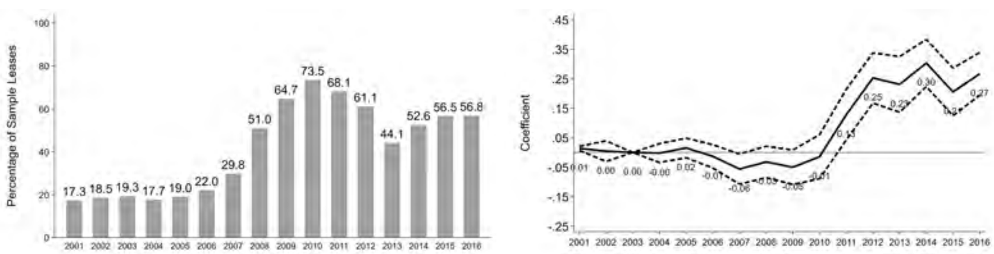
In Figures F1 through F3, we reproduce the time trends estimated by equation 2 with the inclusion of the two demand-side variables as controls. This allows us to understand whether the evolution of lease terms is affected by temporal changes on the demand-side of leasing markets. The figures show point estimates that are nearly identical to those in the main text. The exception is with the arbitration clause, which still increases over time with inclusion of the controls, but not as steeply. This suggests that competition among leasing firms does not significantly modify our finding that clause prevalence and royalty rates increase over the development of the play due to some combination of mineral owners learning or firms becoming less environmentally intensive.



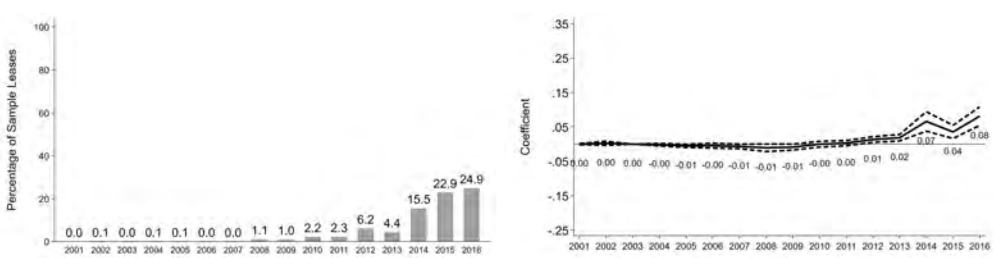
(a) Water Testing



(b) Water Replacement



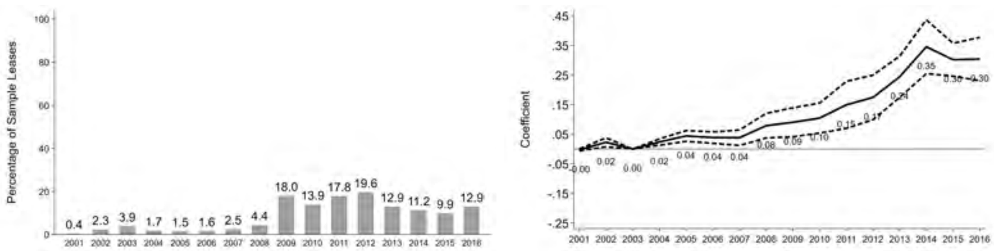
(c) Water Protection



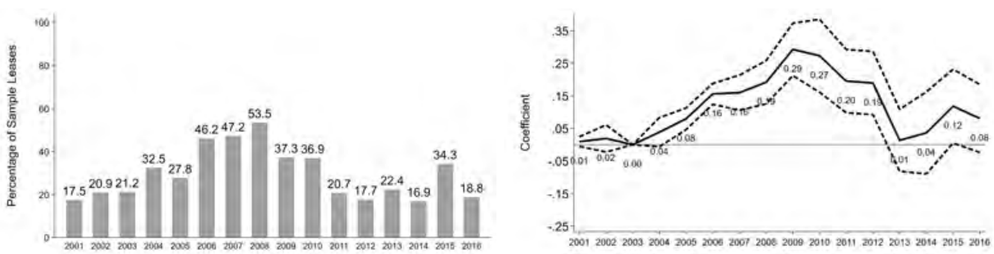
(d) No Wastewater Injection

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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

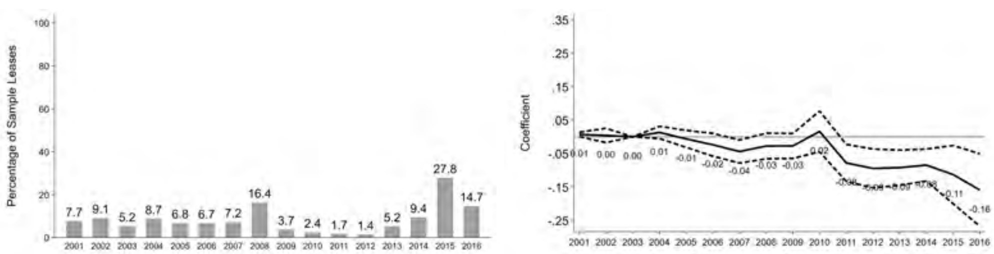
Figure F1: The Prevalence of Select Clauses Over Time, Demand Controls



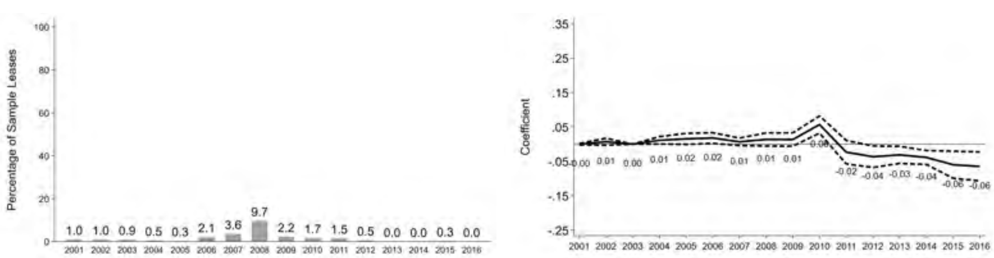
(a) No Surface Use



(b) Development Plan



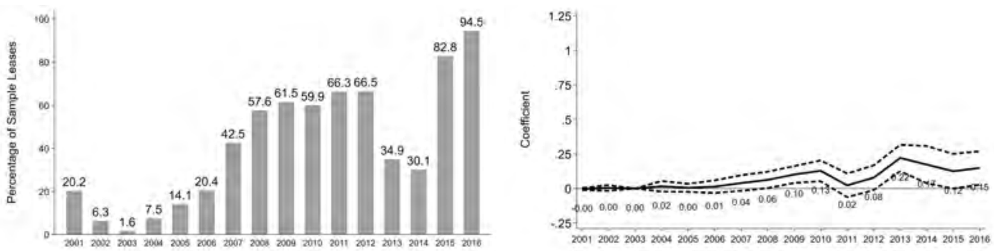
(c) No Pipelines



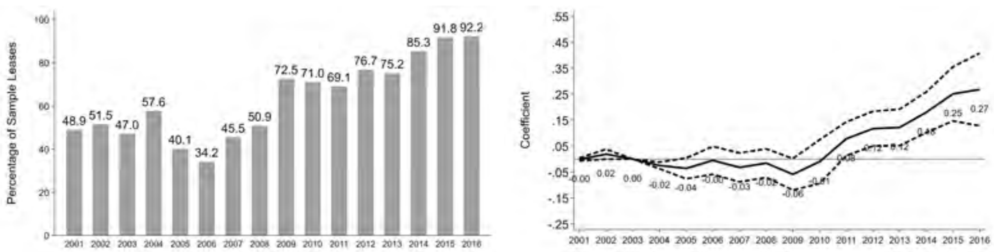
(d) Greater than Minimum Setback

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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

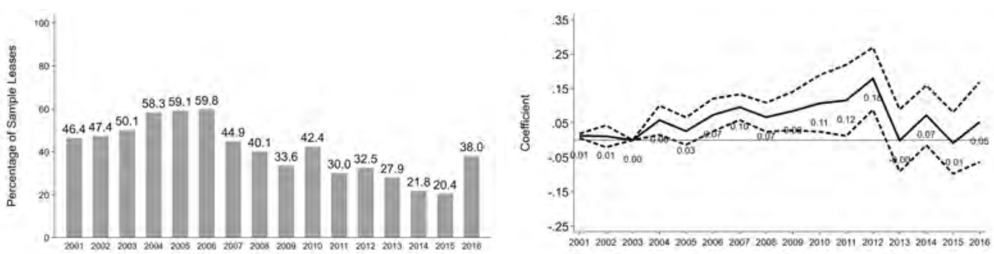
Figure F2: The Prevalence of Select Clauses Over Time, Demand Controls (cont.)



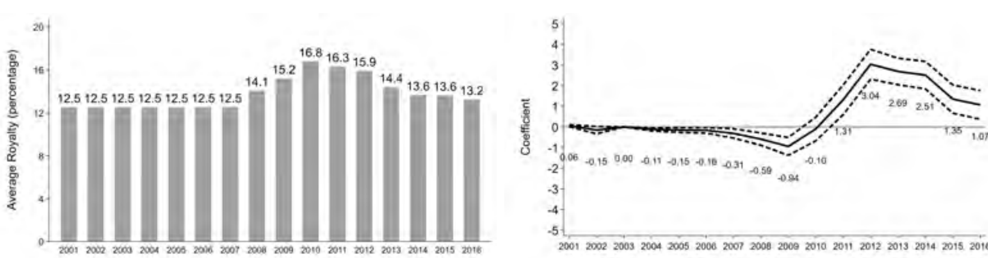
(a) Arbitration



(b) Force Majeure



(c) No Royalty Deductions



(d) Royalties

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 that are two-way clustered by year and municipality. The coefficients represents the difference in the royalty rate or the likelihood of clause adoption in year t, relative to the omitted year, 2001. All models include year, oil and gas operator, and municipality fixed effects.

Figure F3: The Prevalence of Select Clauses Over Time, Demand Controls (cont. 2)

Table F1: Relationship Between Demand-Side Variables and Lease Terms

	Surface Protection (1)	Externalities (2)	Water Quality (3)	Optional (4)	Legal Protection (5)	Favor Producer (6)	Royalty (7)
Unique Firms in Tract	0.002	0.001	0.018***	0.007	0.011**	-0.001	0.060*
	(0.008)	(0.004)	(0.003)	(0.005)	(0.003)	(0.002)	(0.026)
Leased and Drilled	0.040	-0.017	-0.023	-0.005	0.013	0.038*	-0.153
	(0.023)	(0.030)	(0.018)	(0.026)	(0.030)	(0.016)	(0.091)
Intercept	0.547***	0.546***	0.482***	0.256***	0.559***	0.815***	14.190***
	(0.025)	(0.012)	(0.010)	(0.014)	(0.008)	(0.006)	(0.084)
R-squared	0.41	0.42	0.54	0.63	0.38	0.55	0.66
N	58,012	58,012	58,012	58,012	58,012	58,012	58,012

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robust standard errors, two-way clustered by year and municipality are in parentheses. All columns include year, oil and gas operator, and municipality fixed effects. In columns 1 through 6 the dependent variable is the presence of at least one clause in a cluster, defined as a binary variable. In column 7, royalty rates are the dependent variable and are in the dataset as a whole number (e.g., the state minimum rate of 12.5 percent is coded as 12.5).