NBER WORKING PAPER SERIES

ENVIRONMENTAL JUSTICE AND COASIAN BARGAINING:
THE ROLE OF RACE AND INCOME IN LEASE NEGOTIATIONS FOR SHALE GAS

Christopher Timmins
Ashley Vissing

Working Paper 29487
http://www.nber.org/papers/w29487

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2021

We thank seminar participants the PERC, NBER, UCLA, Colorado School of Mines, the National University of Singapore, and Pennsylvania State University for their helpful comments, and we thank Stephen Ryan and Erik Gilje, as well, for helpful discussion and suggestions. All remaining errors and omissions are our own. We also thank the personnel at the Texas Railroad Commission, Drillinginfo, and the Tarrant County Appraiser Office for guidance in collecting the well and housing/parcel data. Timmins' work on this research was supported by a National Science Foundation Grant #1559481. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Environmental Justice and Coasian Bargaining: The Role of Race and Income in Lease Negotiations for Shale Gas
Christopher Timmins and Ashley Vissing
NBER Working Paper No. 29487
November 2021
JEL No. K32,Q40,Q51,Q58

ABSTRACT

Using a unique combination of datasets and estimation techniques, we test whether private lease negotiations to extract oil and natural gas exhibit features of Coasian efficiency. We demonstrate that measures of wealth (including income, house square footage, and land acreage), typically determinants of willingness to pay for environmental quality, do affect bargaining outcomes. However, race, ethnicity, and language also play important roles after conditioning upon these variables, suggesting an environmental injustice and a breakdown of efficient Coasian bargaining. We further demonstrate that failure to negotiate protections in leases leads to increased risk of future drilling violations, which are not offset by local ordinance restrictions.

Christopher Timmins
Department of Economics
Duke University
209 Social Sciences Building
P.O. Box 90097
Durham, NC 27708-0097
and NBER
christopher.timmins@duke.edu

Ashley Vissing
Cornerstone Research
avissing@gmail.com

1 Introduction

Changes in oil and natural gas extraction technology have brought households closer than ever to drilling operations. Whereas drilling activity was once confined to rural regions, shale wells, facilitated by horizontal drilling techniques, now impact suburban households owning sub-acre lots. When the rights to extractable minerals are privately owned, firms must negotiate individual leases with each affected household. Consequently, landowners have an opportunity to protect their families and property from negative drilling externalities not already regulated by the state. These protections can include limits to noise, traffic, road usage, site restoration, and restrictions on chemical compounds used for fracking, pre-drilling soil and water testing, among others. Leases play an important role in the *de facto* regulatory process, and while it may be difficult to contend that the lease negotiation process satisfies all of the conditions for efficient Coasian bargaining, it is often argued that privately negotiated protections in lease agreements can substitute for a more bureaucratic and burdensome regulatory regime. However, contracts are individually negotiated and not all leases are the same – in particular, leases may include more or fewer landowner protections. To the extent that these differences in bargaining outcomes reflect differences in willingess to pay to avoid the externalities associated with drilling, they can still signal an efficient Coasian bargaining process (Coase, 1960). To the extent that they correspond to other household characteristics (e.g., race and language), differences in lease outcomes will suggest that negotiations are inefficient. Moreover, given the drilling externalities involved, they may be also an important source of environmental injustice.

Because each lease is individually negotiated between large leasing and drilling firms and landowners, we frame the negotiations as outcomes of Coasian bargaining. Coase (1960) suggests that under a strict set of assumptions, including (i) minimal transactions costs, (ii) well-defined property rights, and (iii) agents that are symmetrically well-informed, an efficient bargain would be struck under which the external cost of the last unit of pollution to victims would equal its benefit to the polluter. Transfers ensure that no one is made worse off as a result, and depending upon bargaining power, the bargaining solution may represent an improvement for all parties. This result is compelling and has long been used to justify a laissez faire approach to the government's management of polluting activities. In some contexts, the assumptions underlying Coase's theorem are largely satisfied. However, a problem arises in many environmental justice contexts when information is less accessible to certain groups. This may be the case, for example, if language creates a barrier to effective negotiation. More generally, racial discrimination may make it difficult for certain groups to bargain effectively. Indeed, there have been many instances in the past where race (and especially linguistic isolation) have been used to influence negotiations about the location of pollution sources.[1]

Our assembled data uniquely allows us to study the outcomes of privately negotiated contracts and link these outcomes to household characteristics. We find consistent evidence that minority households receive worse contracts along nearly all dimensions, and the effects are compounded by linguistic barriers to bargaining, suggesting an important information asymmetry. Further, we explore how these effects interact with other observable and potentially correlated determinants of bargaining outcomes by estimating a non-parametric, Chaid-algorithm moment forest model. That model shows that minorities' negative outcomes are stronger for linguistically isolated, poorer, and less educated minority populations.

Our empirical models allow us to directly test whether private negotiations lead to equitable and efficient allocations of legal rights and responsibilities as suggested by the Coasian framework. We conclude that leasing in the oil and natural gas industry is neither equitable nor efficient. Making matters worse, we find that, to the extent that landowners can negotiate for protections, oil and natural gas contracts will successfully substitute for local regulations overseeing drilling activities. The negotiated outcomes therefore affect both negotiating landowners and their neighbors whereby less restrictive leases can lead to damage and other long term consequences borne by households residing near wells.

To answer these questions, we combine data describing household characteristics, including racial and ethnic composition and annual income levels, with property characteristics and contracting outcomes, including specific clauses agreed upon by drilling firms and landowners. We then estimate a Seemingly Unrelated Regression (SUR) model of contract composition, which tests whether certain household and property characteristics are more or less likely to lead to certain contract outcomes. Controlling for income, evidence from the SUR model suggests that minority households receive fewer protections, and the more linguistically isolated the household, the fewer protections. Second, we show that these results are robust to using alternative methods for coding race and ethnicity. Specifically, we employ the Bayesian Improved Surname Geocoding (BISG) algorithm, which uses surnames and neighborhood composition to predict a household's likely racial and ethnic composition.[2] Our conclusions hold up to using this alternative measure of race and ethnicity. Third, in order to characterize how different household characteristics interact, we estimate non-parametric, heterogeneous effects of being a minority household using a moment forest model adapted from computer science (Asher et al., 2016). We apply the model to the SUR framework

[1]See, for example, the case of Kettleman City described in Cole and Foster (2001).

[2]The BISG is a model used by the utilized by the Consumer Financial Protection Bureau to monitor fair lending practices in the absence of data describing customers' reported race or ethnicity. This measure may be more relevant for firms negotiating leases by phone or, in general, where firms deciding with whom to negotiate and what to offer before knocking on doors.

whereby we simultaneously estimate a set of equations. With the technique, we are able to measure how effects on minorities vary along multiple interacting dimensions including linguistic isolation, median household income, and levels of educational attainment. With these additional dimensions of heterogeneity, we conclude that lease outcomes are worse for minority households that are not only poorer and less educated, but also linguistically isolated. Finally, we test whether these differences in bargaining outcomes are salient for long-run outcomes. In particular, are more restrictive contracts associated with fewer violations? We find that this is indeed the case. This confluence of results suggests potential environmental injustice in the market for privately negotiated natural gas leases.

This paper touches on a number of relevant literatures. In particular, we explore whether the leasing process is consistent with efficient Coasian bargaining, or whether there are other factors (in particular, factors that could be associated more directly with race and language) that also play a role. Our work is therefore related to the literatures on the Coase Theorem and environmental justice. Environmental justice has a rich interdisciplinary history; for a discussion from an economics perspective, see Banzhaf et al. (2019) and Banzhaf et al. (2018). For an overview of the Coase Theorem we refer the reader to one of several reviews (Schweizer, 1988; Medema, 1994; Zelder, 1998; Medema, 2014). Jenkins et al. (2004) and Hamilton (1995) explore the efficiency of the Coasian bargaining process. Finally, because we focus on the effect of lease negotiations on exposure to the nuisances associated with shale gas extraction, the literature on the health impacts of that exposure is also relevant (Colborn et al., 2011; McKenzie et al., 2012; Bamberger and Oswald, 2012; Hill, 2013). Other papers have focused specifically on exposures through drinking water (Vengosh et al., 2014; Hill and Ma, 2017a,b).

This paper proceeds as follows. Section 2 describes the relevant institutional and legal frameworks surrounding shale gas development in Texas. The specifics of the legal structure surrounding lease clauses, lease negotiation, and municipal regulation play a particularly important role in our conclusions. Section 3 describes our data, which are drawn from information about the outcomes of lease negotiations and proprietary data on housing transactions combined with publicly available data on the race and income of homeowners. Section 4 motivates our empirical approach with a brief discussion of the Coase Theorem and relates it to environmental justice. Section 5 describes our empirical modeling strategy and summarizes results. Section 6 concludes with a discussion of our results with a particular focus on policy implications.

2 Legal and Institutional Detail

The following sections describe the technological developments driving the natural gas industry's increased productivity and the institutional details behind the regulation of the industry. Over the past 20 years, combined horizontal drilling and large-scale hydraulic fracturing technology have evolved in combination to allow access to natural gas contained in tight-shale formations spread over a large area while requiring fewer well pads. Further, these technological developments have increased activity in urban plays, literally bringing drilling into suburban households' backyards. Regulation guiding industry practices, however, has been largely crafted for activity in less densely populated areas – the more common setting for natural resource extraction. We describe the technological innovations and regulatory structure relevant to our story of Coasian bargaining and environmental justice.

2.1 Hydraulic Fracturing and Horizontal Drilling

The process of hydraulic fracturing enables firms to extract natural gas from tight shale formations by artificially stimulating the strata. This increases the flow of natural gas within the shale, resulting in its eventual release and collection at the wellhead. Horizontal drilling techniques allow firms to drill wells accessing minerals located within a large radius surrounding the wellhead. Fewer drill sites are therefore required to reach a larger subsurface area and better access is provided for broad resource deposits. Horizontal drilling allows firms to extract large quantities of natural gas from a smaller surface footprint, facilitating extraction from areas of higher population density. As these unconventional drilling techniques have been adopted by the industry, individuals in (sub)urban areas have subsequently found themselves to be parties to negotiations with operators over mineral rights leases.

2.2 Texas Railroad Commission and State-Level Regulation

The Texas Railroad Commission (TRC) oversees the majority of the oil and natural gas industry in the state of Texas, which includes the approval of permits to drill wells.[3] However, prior to permit approval, firms must first amass a large and sufficient mineral estate acreage that is spaced far enough away from existing well infrastructure to be approved and permitted by the TRC.[4] Mineral estate acreage is obtained by signing leases with mineral rights owners. Households signing leases with natural gas firms or "landmen" are tasked with weighing the trade-offs between future income derived from natural gas extraction from their mineral estate with that of the potential known and unknown risks of living near an active well. Once a well is permitted, the operator typically has two years to begin drilling the well before the permit expires.

The TRC's jurisdiction regulating the industry extends to the drilling and production phases; however, the TRC does not regulate noise, traffic, or well pad appearance, nor does it require air pollution testing. By law, operators have access to surface water to be used

[3]The Texas Railroad Commission has jurisdiction over the "exploration, production, and transportation of oil and gas prior to refining or end use," and the TRC executes its jurisdiction by enforcing rules written in the Texas Administrative Code, Chapter 3.

[4]Texas Administrative Code, Chapter 3, Rules 3.37 & 3.38.

to treat the well, and chemical disclosure is restricted to only the non-proprietary chemicals used to fracture a well.[5] In general, the dis-amenities experienced by households from nearby shale gas activities are unregulated by state and federal entities and are addressed in the private leases signed between them and firms.

More specifically, federal and state regulators generally do not have direct jurisdiction over the private contracts drawn between landowners and parties interested in leasing land for exploration and production of oil and natural gas. Higher-level regulation is limited to royalty payments (stipulating when they are to be paid), the required information that must be provided (and that which can be requested) by firms, notification upon re-assignment of leased rights, and determining the consequences of delinquent payments. In addition, the TRC has jurisdiction over enforcing and undertaking remediation from undue negligence on the part of firms and broadly enforcing the protection of ground and surface water from contamination caused by the industry. However, the TRC's jurisdiction over the leases signed between households and firms, and subsequently, the protection of households while a well is drilled and after production ends is limited, and a well-informed household may negotiate more comprehensive contracts with leasing firms to protect their interests beyond the minimal coverage of the law.

2.3 Municipal Regulation

During the period we study, cities in Texas exercised "home rule," which allows municipalities to pass and enforce local ordinances that restrict drilling activity within their jurisdictions. These local ordinances further restrict the activities of firms by requiring, for example, larger set-back distances, additional permits and fees, well construction restrictions, and additional environmental tests.

Local municipalities often employ land-use policy to restrict oil and gas development within their jurisdictions. Municipal governments are able to enact local ordinances that stipulate types and locations of land use and permissible damage for the purposes of protecting public health and welfare. Local ordinances are rendered preempted (or essentially invalid) if state-level legislation has passed that limits local power directly (expressed preemption); the state rules already occupy a field (implied preemption); or if they conflict with existing state laws (Urban Lawyer, 545-546). The last of these usually restricts local zoning ordinances that loosen state rules, but it can uphold those that strengthen them.

Tarrant County is unique because it is largely comprised of municipalities with ordinances restricting drilling behavior. These "home rule" ordinances were passed as early as

[5]Texas Administrative Code, Chapter 3, Rule 3.29. All chemicals and quantities used are reported via the Chemical Disclosure Registry and publicly disclosed unless a component of the composition is considered a trade secret, and the trade secret component is omitted from public disclosure.

1996 (Lakeside) and as late as 2011. However, the majority of drilling in the Barnett Shale and other mineral-rich regions of Texas does not occur within municipal boundaries, and as a consequence, landowners are more protected through stringent leasing standards. Further, Texas state legislators passed HB 40 in 2014 in response a Denton-county ban on hydraulic fracturing techniques that significantly limits the jurisdiction of "home rule," or local ordinance regulation of drilling behavior, thereby increasing the importance of individually negotiating protective leases.

2.4 Lease Components

Signing a comprehensive leasing document is important for households protecting their rights while they are royalty interest holders. In particular, lease terms can compensate for the absence of state or municipal regulations. Leasing agreements contain a set of primary clauses common to all leases drawn in the industry and sets of auxiliary clauses that are negotiable between lessors (mineral rights owners) and lessees (exploration and oil and natural gas production firms). Primary clauses include a careful description of the minerals leased to the lessee; information about the royalty payments owed to the lessor once the well begins to produce in paying quantities; the duration of the lease, or primary term; and opportunities for extension once the primary term has expired.

Auxiliary clauses are written into the agreement to protect one or both of the parties, but may not be included in all leases. Negotiators may draft surface damage clauses ensuring that the operator restores the surface to a condition agreed upon by both negotiators once production is complete; environmental clauses requiring producers to perform regular environmental quality tests on the surface and ground water or soil samples; and pooling restrictions ensuring that the leased land value is not diluted in terms of royalty payments by being grouped into an unnecessarily large acreage, to name a few.

2.5 Split Estate

One might typically assume that the lease signatory is the household, or surface-rights owner; however, the state of Texas allows the mineral estate to be split (or "severed") from the surface estate. The individual signing a lease with a natural gas firm may not, therefore, be the individual living in the house positioned on the surface estate. As early as 1953, Texas courts declared that landowners may reserve mineral rights and the oil and gas contained as in the case Benge v. Scharbauer [259 S.W.2d 166 (Tex. 1953)], thereby enabling the mineral estate to be severed from the surface estate (Merrill and Merrill (2013)).[6] In the event of severance, the mineral estate dominates in terms of exploration and extraction, and

[6] A grant or reservation of minerals by the fee owner affects a horizontal severance and the creation of two separate and distinct estates: an estate in the surface and an estate in the minerals [Acker v. Guinn, 464 S.W.2d 348, 352 (Tex. 1971)] (Fields (2006)).

the mineral lessee assumes the same rights owed to the mineral estate owner since the leasing document is perceived as a temporary transference of ownership.[7] Colloquially, the owner of the mineral estate may lease the minerals to third parties for exploration, but the law only requires that the lessee (i) notify surface owners of the "intent to explore and drill;" (ii) have access to as much land as is necessary to explore and drill; (iii) be able to remove trees and fences to make way for well and equipment; (iv) be able to take up to one acre of land for the well pad; and (v) be able to erect pipelines to transport the natural gas off the property (Rahm (2011)).[8,9]

As an independent entity, the mineral estate may exercise its rights without consulting the surface estate owners. Subsequently, a firm leasing the mineral rights for purposes of oil and gas exploration and extraction need only negotiate with the mineral estate owners whether they also own the surface estate or not. The owners of the mineral estate may exercise their rights without consulting the owners of the surface estate, and they are only required to inform the surface estate when drilling is imminent on their property due to legislation passed in 2007 (Maxwell, 2008).[10] Additionally, the mineral estate may use as much surface water from the leased land as is reasonably necessary to carry out operations, given that the use is not wasteful, and it may inject wastewater into sub-surface formations.[11,12] Moreover, the mineral estate does not accept responsibility for the full restoration of the property,[13] nor is it required to pay surface damages as long as the damage is not unreasonable.

We develop an algorithm, described in Appendix A, to identify split estates. This is useful in our application, as it allows us to consider only properties where we know that the surface rights holder negotiated the agreement to lease the mineral rights. This is important, as it is the surface rights holder for whom we can identify race and income (see below).[14]

---

[7] If the minerals are not reserved as of the house purchase date, the mineral estate automatically goes to the buyer along with the surface conveyance (Fambrough (1987)).

[8] Under the dominance of the mineral estate, there are five interests including the right to develop the mineral estate, or ingress and egress, which includes exploration; to lease; and to receive bonus payments, delay rentals and royalty payments (Vanham and Riley (2011)).

[9] There are three exceptions to the dominant mineral estate including excessive use of land in exploration and operation activities to access the minerals, unnecessarily injuring the surface, and not accommodating the existing surface use, the latter more formally entitled the Accommodation Doctrine (Letter of the Law, 1997).

[10] Texas Natural Resource Code, 91.703(a): Not later than the 15th business day after the date the commission issues an oil or gas well operator a permit to drill a new oil or gas well or to reenter a plugged and abandoned oil or gas well, the operator shall give written notice of the issuance of the permit to the surface owner of the tract of land on which the well is located or is proposed to be located.

[11] Warren Petroleum Corp. v. Martin, 271 S.W.2d 410 (Tex. 1954).

[12] Unless specified in the deed, the water rights fall to the surface owner but they are accessible with reasonable use by the mineral estate (Vanham and Riley (2011)).

[13] Warren Petroleum Corp. v. Monzingo, 304 S.W.2d 362 (Tex. 1957).

[14] Fractional interests, whereby groups of individuals may own parts of what appears to be one piece of property, are another potential issue related to mineral rights. It is our understanding that an oil and natural

2.6 Rights of Surface Owners

Texas has not passed a law protecting the surface estate, or a surface damage act, as has been passed in other states with prominent oil and natural gas industries (including New Mexico, Oklahoma, North and South Dakotas, and Montana). As mentioned above, surface owners are not owed any remuneration for the opportunity cost of the piece of their property that is lost during the drilling period, nor must they be paid for reasonable damages to the land caused by drilling. If there is any perceived misuse of the land by mineral rights owners, surface owners are responsible for proving unreasonable conduct, which does not include surface damage or inconvenience. Surface owners are marginally protected by the Accommodation Doctrine, which protects existing surface owner uses.[15]

In lieu of state regulation, and in some cases, local ordinances, lessors can negotiate a surface damage clauses into the leasing agreement that protect the surface during production and ensure remediation after production ends. Otherwise, the surface estate may only claim and prove unnecessary damage through litigation. However, since the state of Texas does not require operators to negotiate with the surface estate independent of the mineral estate, a severed estate with a lease may not naturally include a surface damage clause. Property owners holding deeds to both the surface and sub-surface can negotiate surface damage clauses that might include complete restoration of the surface upon completion of the well and perhaps assign a property damage fee; however, among mineral estates, there is little incentive for a mineral estate owner to negotiate a surface damage clause with a potential operator. In the latter case, among severed estates, surface owners have little legal protection.

3 Data & Variable Construction

We use a unique combination of lease, well activity, housing, and demographic datasets. This section details the sources of data and describes how certain variables are constructed. The data originate from several sources and are compiled for analysis by a series of novel matching techniques, primarily based on text strings.

---

gas lease would be signed by each fractional mineral interest holder under these circumstances, unless there is an alternative legal set-up granting one person the right to sign on behalf of all, for example. We may not observe fractional interest holders in our data. We match leases to specific properties by using the name information that we have from leasing and housing records, using the listed mineral interest holder names in the leasing data to link to the housing data. That said, it is also our understanding that Tarrant County may not have as many fractional interests as more rural parts of Texas where there is a legacy of drilling and earning money from royalties and bonuses. We are not able to formally assess whether fractional interests could bias our results.

[15] Accommodation Doctrine: [W]here there is an existing use by the surface owner which would otherwise be precluded or impaired, and where under the established practices in the industry there are alternatives available to the [mineral owner] whereby the minerals can be recovered, the rules of reasonable usage of the surface may require the adoption of an alternative by the [mineral owner]. (Tarrant County Water Control & Improvement Dist. No. 1 v. Haupt, Inc., 854 S.W.2d 909, 911 (Tex. 1993)) (Merrill and Merrill (2013)).

Of primary importance are the well and lease data with which we describe the precise amount of well exposure and the types of leases signed by individual households. We have also collected relevant housing transaction and appraisal data from the Tarrant County Appraiser District Office and Corelogic, a national real estate data aggregator. Demographic information is collected from publicly available Home Mortgage Disclosure Act (HMDA) data. Data from DrillingInfo, the Texas Railroad Commission, and the Texas Commission on Environmental Quality are also used in the analysis.

3.1 Property and Household Characteristics Data

The housing data is the nexus for several of our data sources. Leases are merged to specific properties found in the housing data by address using various string matching methods, and household characteristics are merged to specific properties using loan characteristic information. The combined data span roughly 2003 to 2014, encompassing most of the leasing and drilling activity in Tarrant County, Texas. Finally, since we know the precise property location, we map properties to the relevant U.S. Census Tracts, nearest wells, city ordinances, neighborhood violations, water sources, and cities. We control for property characteristics like the parcel size, the living quarter size, and whether the house likely accesses groundwater. Table 1a summarizes the property characteristics relevant to the empirical analysis.

HMDA. Loan characteristics from the Corelogic housing data facilitate matching specific properties to individuals in the HMDA data, which allows us to characterize the race, ethnicity, and income of the household members who sign leasing contracts in our data.[16] HMDA legislation originally passed in 1975 (and was amended in 1989) to ensure fair and adequate lending across demographic groups and to limit discriminatory lending practices. Consequently, for each housing purchase that includes a mortgage, financial institutions must publicly disclose the income level, race, ethnicity, and gender, along with terms of the loan. Table 1a summarizes the demographic attributes in the HMDA data that are matched to our landowners.

Because our analysis highlights how minority communities are disproportionately located closer to natural gas wells, we summarize our primary model variables by race and ethnicity for Hispanic, Black, and white households in Table 2. Average income is between \$8,000 and \$25,000 greater for white households than Black and Hispanic, respectively. Black and Hispanic households are also about 5% more likely to live within two kilometers of a well.

We also find that there is a fair amount of racial segregation in our data. Table 3 reports the ANOVA F-statistics and adjusted R-squared values for the individual measures

[16] Loan characteristics from HMDA are matched to specific properties in the housing data based on the property-level U.S. Census Tracts, loan amounts, transaction year, and lender identities.

Table 1: Summary Statistics

(a) Household characteristics, HMDA & BISG Samples

HMDA Sample	Mean	(St.Dev.)	*BISG Sample*	Mean	(St.Dev.)
Household Income	67,254	(45,261)			
Land area (sqft.)	9,626	(6,103)		9,745	(6,440)
Living area (sqft.)	1,895	(739)		1,881	(760)
% Groundwater properties	60.8%	(48.8)		57.9%	(49.4)
% Split estate properties	3.1%	(17.2)		4.3%	(20.3)
% Houses within 1 km. of a well	47.1%	(49.9)		47.4%	(49.9)
% Houses within 2 km. of a well	91.5%	(27.9)		91.0%	(28.4)
% Hispanic household	28.6%	(45.2)	% Pred. Hispanic	29.3%	(41.1)
% Black household	10.9%	(31.2)	% Pred. Black	15.7%	(23.4)
% Asian household	4.3%	(20.2)	% Pred. Asian	5.0%	(19.1)
% White household	73.1%	(44.4)	% Pred. White	48.1%	(37.8)

(b) U.S. Census Tract Summaries

HMDA Sample	Mean	(St.Dev.)	*BISG Sample*	Mean	(St.Dev.)
Median HH Income	68,990	(25,426)		67,463	(26,253)
% Spanish Speaking	14.7%	(14.4)		15.6%	(15.7)
% Owner Occupied	69.7%	(18.2)		69.1%	(18.1)
% College Grad.	20.7%	(10.4)		20.0%	(10.7)
% Population > 60 yrs.	7.6%	(4.9)		7.7%	(4.9)
Observations	26,551			67,121	

Notes: (i) Panel 1a summarizes the individual household characteristics for the HMDA and BISG samples for which we observe lease attributes; (ii) Panel 7 summarizes the measures of lease quality for the HMDA and BISG samples; (iii) Panel 1b summarizes the U.S. Census Tract-level characteristics, and there are 60 unique tracts covered by the HMDA and 175 tracts covered by the BISG samples; (iv) note that the Hispanic describes ethnicity, while Black, Asian and White describe racial categories, so the sum of all four demographic shares will not necessarily be less than 100 percent.

of race and ethnicity, and they confirm that there is more variation in racial and ethnic characteristics across U.S. Census tracts than within for individuals who are Black, Hispanic, and white.

U.S. Census Tract. Additionally, Table 1b reports U.S. Census Tract-level variation in the fraction of the population for whom Spanish is a primary language (*% Spanish Speaking*), median household income, and the fractions of the population who are college graduates and/or older than 60 years of age. We map each property to a U.S. Census Tract using geographic information system (GIS) software.[17] We use tract-level variables to control for

[17]U.S. Census Tracts are usually contiguous, small delineations of land formed by the U.S. Census Bureau based on decennial census participation with an optimal population size of 4,000 residents. `https://www.census.gov/geo/reference/gtc/gtc_ct.html`

Table 2: Summary Statistics by Race & Ethnicity

	Hispanic		Black		White	
	Mean	(St.Dev.)	Mean	(St.Dev.)	Mean	(St.Dev.)
(a) House Attributes						
Income	46,651	(28,111)	62,252	(36,843)	71,474	(47,586)
Land area (sqft.)	8,763	(4,933)	8,548	(4,771)	9,992	(6,479)
Living area (sqft.)	1,552	(579)	2,009	(772)	1,927	(736)
Groundwater	47.6	(50.0)	54.5	(49.8)	63.3	(48.2)
% Split estates	2.7	(16.3)	2.3	(15.1)	3.3	(17.9)
% Houses within 2 km.	95.9	(19.9)	96.0	(19.6)	90.0	(30.0)
(b) U.S. Census Tract Attributes						
Median HH Income	56,429	(21,331)	69,778	(25,291)	70,856	(25,569)
% Spanish Speaking	23.7	(20.4)	11.8	(8.3)	13.6	(13.3)
% Owner Occupied	67.8	(15.5)	71.4	(19.0)	69.8	(18.3)
% College Grad.	14.7	(9.5)	21.2	(9.5)	21.6	(10.3)
% Population > 60 yrs.	7.8	(4.5)	6.4	(4.6)	7.8	(5.0)
Observations	7,581		2,897		19,388	

Notes: (i) Panel 2a summarizes the individual household characteristics for households part of the HMDA sample; (ii) Panel 2b summarizes the U.s. Census Tract-level characteristics for households part of the HMDA sample; (iii) white households report higher income and have fewer natural gas wells located within one and two kilometers compared to Hispanic and Black households.

Table 3: Race and Ethnicity ANOVA-Measured Variation by U.S. Census Tract

	F-stat	Adj. R-Squared
Black	12.12	0.11
Asian	3.70	0.03
Hispanic	28.34	0.24
White	11.72	0.11
Observations	26,523	

Notes: (i) Table 3 summarizes the variation in race and ethnicity across versus within U.S. Census tracts for the HMDA sample; (ii) the F-statistic and adjusted R-squared are results from the ANOVA test and confirm that there is more variation across U.S. Census tracts compared to variation within.

variation in neighborhood characteristics throughout our analyses.[18]

[18] Prior versions of the paper controlled for wellpad fixed effects, and the models estimated with them produced similarly interpreted results. Wellpad fixed effects are constructed by matching each property with the nearest drilled well and horizontal lateral, and each well is matched to a wellpad by assigning all wells and laterals drilled within 100 meters to a single wellpad. The horizontal drilling technology facilitates drilling wells in close proximity to one another with laterals that extend in many directions from that central location. This allows operators to confine intensive drilling activity to smaller, wellpad spaces while extracting oil and natural gas from a much larger subsurface mineral acreage. A single wellpad can be comprised of up to 3,000 individual, sub-acre plots of land in a region like Tarrant County, Texas. On average, there are 145

Drilling Exposure. We also map each property to (a) the nearest well before each home/property purchase and (b) the nearest well ever drilled, which allows us to measure well exposure for each individual property. Well permitting and production information comes from two sources: the Texas Railroad Commission and DrillingInfo, which is a proprietary aggregator of drilling activity information. In addition to drill timing, exposure is calculated for well activity within one and two kilometers of each property. Exposure reported in Table 1a summarizes drilling that ever occurs within one and two kilometers of each property. Landowners in the data are 91% likely to live within two kilometers of an hydraulically fractured natural gas well by the end of the data (2014).

Table 4 captures how drilling activity has occurred nearer to households in recent years. On average, households live within two kilometers of 2.2 wells by 2013, which increased from only 0.24 wells as of 2003. Similarly, households are located, on average, 1.3 kilometers away from the nearest well in 2013 compared to a distance of 10.4 kilometers in 2003.

Table 4: Well Exposure (within 2000 meters)

	Well count within 2000 meters			Minimum distance (kilometers)		
	Mean	N	Std. Dev.	Mean	N	Std. Dev.
2003	0.24	726049	1.22	10.4	727668	5.2
2004	0.38	724328	1.48	9.7	727668	5.5
2005	0.56	721845	1.71	7.2	727668	5.0
2006	0.81	718515	2.00	5.4	727668	4.2
2007	1.03	716709	2.20	3.8	727668	3.1
2008	1.24	713039	2.35	2.8	727668	2.1
2009	1.40	710400	2.40	2.5	727668	1.8
2010	1.65	709020	2.47	2.1	727668	1.5
2011	1.96	706499	2.55	1.6	727668	1.2
2012	2.15	704173	2.56	1.4	727668	1.0
2013	2.20	702881	2.57	1.3	727668	0.9

Notes: (i) The first three columns describe the mean count of wells drilled within 2000km of the household across the years in the sample; (ii) the last three columns describe the mean distance to the nearest well; (iii) the large sample sizes reflect observing the well exposure for each property in the data for each month and year spanning Jan. 2003 to Dec. 2013.

3.2 Lease Data

The lease characteristics and legal language contained therein are primary and unique sources of data used to describe the outcome of the bargaining process conducted between two parties – the lessee (who is typically a well operator or third party "landman") and the lessor (the owner of the mineral rights, who is also the owner of the surface rights in the case of a full estate).[19] Signing the lease conveys the interests of the mineral estate from the

properties per wellpad.

[19]Full estates are also referred to as fee-simple, or whole, estates.

lessor to the lessee, and we use the names of buyers, sellers, and owners (if different), along with detailed addresses, in the housing data to match each property to a natural gas mineral lease.[20]

Signing a comprehensive lease document is important for households protecting their rights while they are royalty interest holders to a drilled well. Lease terms can compensate for the absence of state or municipal regulations that restrict firm behavior and protect households from harmful drilling externalities. Each contract is comprised of (a) primary terms common to leases, (b) a bonus payment paid when the lease is signed, and (c) combinations of auxiliary clauses that are negotiable between lessors and lessees. Primary clauses include a careful description of the minerals leased to the lessee, royalty payments owed to the lessor once the well begins production, the years allowed for firms to drill before losing the leased rights (i.e., primary term), and opportunities for extension once the primary term has expired. Auxiliary clauses govern other aspects of the drilling process such noise, fracturing chemicals, easements, and well setbacks from homes and waterways.

We collected data describing the primary and auxiliary terms of these privately negotiated lease contracts, in addition to a small sample of bonus payments. In particular, our data accessed through Drillinginfo describes the primary clauses (i.e., royalty rate and lease term) of most natural gas leases, and bonuses for a few leases, negotiated in Tarrant County, Texas between 2000 and 2013. Auxiliary clauses were extracted from pdf copies of leases collected from the "Drilling Down Series" published by the *New York Times*[21] from 2011 into 2012 and from the Tarrant County Clerk website.[22] We web-scraped these data, converted the images to text files, and then mined the text files for words and phrases indicating the existence of specific clauses using an algorithm written in Python. This process is described in Appendix A, and a list of these clauses and descriptions is included in Appendix B.

Table 5 summarizes primary clauses (royalty rate and primary term length in months), bonuses, and a set of different auxiliary clauses that we use throughout our analyses. The mean royalty rate is roughly 0.23 and the most frequently included clauses are *Environmental Restrictions*, *Noise Restrictions*, and *Subsurface Easement*. The sample of bonuses is comprised of initial and extension bonuses, which are highly correlated in the sample for which we observe both.[23]

Table 5 also reports the F-statistic and Adjusted R-squared for the ANOVA test for variation within U.S. Census Tract. All leases have large F-statistics that suggest variation

[20]See the Appendix for more details on the matching process.

[21]`http://www.nytimes.com/interactive/us/DRILLING_DOWN_SERIES.html`

[22]`https://tcrecordsonline.com/?linklocation=supermenu&linkname=Real%20Estate%20Search`.

[23]We do not report the correlation between initial and extension bonuses in a table; however, in the sample of leases reporting both values, the correlation is roughly 0.9.

Table 5: Lease ANOVA-Measured Variation by U.S. Census Tract

	Obs.	Mean	(St.Dev.)	F-stat	Adj. R-Squared
Royalty	25,565	0.232	(0.023)	37.18	0.30
Term Length (months)	26,358	44.82	(12.11)	43.08	0.33
Bonus	3,462	7,676	(8,166)	26.53	0.65
Bonus per acre	3,462	1,924	(2,699)	14.21	0.49
% of leases containing the following clauses ...					
Environment		19.53	(39.64)	40.90	0.31
Noise		22.82	(41.97)	35.08	0.28
Traffic		11.20	(31.41)	59.59	0.40
Surface Restoration		6.39	(24.46)	34.65	0.28
Set-back		12.79	(33.40)	70.08	0.44
Vertical Pugh		9.04	(28.68)	63.18	0.42
Post-production Cost		17.17	(37.71)	47.16	0.35
Lessee Defends		10.8	(31.11)	63.32	0.42
Subsurface Easement		44.79	(49.73)	16.13	0.15
Free water		16.90	(37.47)	13.34	0.12
Observations	26,523				

Notes: (i) Table 5 summarizes the variation in lease clauses across versus within U.S. Census tracts for the HMDA sample; (ii) the mean and standard deviation summarize the the lease characteristics for the primary data sample; (iii) the F-statistic and adjusted R-squared are results from the ANOVA test and confirm that there is more variation across U.S. Census tracts compared to variation within.

in leases clauses across U.S. Census tracts is generally greater than variation within tracts.

Finally, we understand that some neighborhoods may negotiate the terms of their leases as a group, which may lead to greater disparity in lease outcomes across groups of neighbors. In particular, neighborhood associations may negotiate on behalf of a group of neighbors, and they may be assisted by legal counsel. If neighborhoods are segregated (because of discrimination or sorting behavior),[24] this could lead to correlations between race and lease outcomes, leading to greater disparity in lease outcomes among groups that are unable to negotiate better terms compared to others. While we do not control for this factor directly in our models, we do control for unobservable neighborhood characteristics by estimating

[24]Note that F-statistics reported in Table 3 confirm that there is segregation across U.S. Census Tracts in Tarrant county.

U.S. Census Tract fixed effects.[25,26]

3.3 Bayesian Improved Surname Geocoding (BISG)

HMDA is our primary source of information describing a landowner's race and ethnicity. While we match a significant fraction of individuals in HMDA to our lease data (via the housing data), we are limited to individuals that, before signing a lease, borrow to purchase their homes. Further, HMDA is an administrative dataset that identifies individuals' race and ethnicity based on mortgage lender responses. Consequently, we employ a parallel analysis that uses a second form of race and ethnicity identification, namely prediction based on Bayesian Improved Surname Geocoding (BISG). In particular, we adapt source code used by the Consumer Financial Protection Bureau to assess fair lending practices across non-mortgage products in which lenders do not collect demographic data.[27] The algorithm proxies race and ethnicity utilizing surnames and geography that are matched to a U.S. Census surname and geography list (zip code, tract, or blockgroup). The BISG algorithm uses a U.S. Census database to probabilistically assign a *surname* to a particular racial or ethnic category. The surname-based probability is updated via Bayes Rule using the likelihood individuals of a particular race or ethnicity reside in that region (zip code, tract, or blockgroup).

Table 1a summarizes the mean predicted probability that an individual is Hispanic, Black, Asian or White. Table 6 reports the correlation between *BISG predicted* race and ethnicity and the HMDA categorizations. The BISG Hispanic predictions best align with HMDA.

Finally, Table 7 summarizes the lease clauses for the two separate samples - the HMDA and BISG samples. In large part, the two samples have very similar means and standard deviations.

---

[25] In a related point, we note that some lease clauses may have spillover effects - e.g., if one household negotiates a noise restriction into their lease, their neighbors likely benefit from the restriction whether it is part of their contract or not. With spillovers, the correlations between race and lease clauses that we measure are likely an understatement of their true impacts to the extent that race is correlated within neighborhood. It suggest that in white neighborhoods where individuals have negotiated good leases, households are likely to benefit from their neighbors' spillovers, and in Black or Hispanic neighborhoods where fewer beneficial clauses are negotiated, neighbors benefit less as well.

[26] Fractional interests, whereby groups of individuals may own parts of what appears to be one piece of property, are another potential issue related to mineral rights. Unless there is an alternative legal set-up granting one person the right to sign on behalf of all, it is our understanding that an oil and natural gas lease would be signed by each fractional mineral interest holder under these circumstances. We match leases to specific properties by using the name information that we have from leasing and housing records along with the listed mineral interest holder names in the leasing data to link to the housing data, but may not observe fractional interest holders. That said, it is also our understanding that Tarrant County may not have as many fractional interests as more rural parts of Texas where there is a legacy of drilling and earning money from royalties and bonuses. We are not able to formally assess whether fractional interests could bias our results.

[27] `https://www.consumerfinance.gov/data-research/research-reports/using-publicly-available-information-to-proxy-for-unidentified-race-and-ethnicity/`

Table 6: HMDA & BISG Correlation

	Hispanic	Black	Asian	White
Pred. Hispanic	**0.74**	-0.16	-0.08	-0.16
Pred. Black	-0.30	**0.46**	-0.05	-0.17
Pred. Asian	-0.08	-0.02	**0.54**	-0.19
Pred. White	-0.59	-0.09	-0.14	**0.37**
Observations	24,286			

Notes: (i) Table 6 reports the correlation across observed HMDA race/ethnicity and BISG predicted (household-level) race/ethnicity variables using the HMDA sample; (ii) Hispanic, in particular, has a high rate of predictability using lessor (homeowner) surnames and home zip codes.

Table 7: Summary Statistics, Lease Attributes

	HMDA Sample		*BISG Sample*	
	Mean	(St.Dev.)	Mean	(St.Dev.)
Royalty	23.2	(2.3)	23.2	(2.3)
Term Length (months)	45.0	(12)	45.1	(12)
Bonus	7,672	(8,171)	7,515	(8,039)
Bonus per acre	1,923	(2,701)	1,900	(2,720)
% of leases containing the following clauses ...				
Environment	19.5	(40.0)	19.0	(39.3)
Noise	22.8	(42.0)	22.4	(41.7)
Traffic	11.2	(31.4)	11.0	(31.3)
Surface Restoration	6.4	(24.5)	6.4	(24.4)
Set-back	12.8	(33.4)	12.6	(33.1)
Vertical Pugh	9.0	(28.7)	9.0	(28.6)
Post-production Cost	17.2	(37.7)	17.0	(37.5)
Lessee Defends	10.8	(31.1)	10.8	(31.1)
Subsurface Easement	44.8	(49.7)	44.5	(50.0)
Free water	16.9	(37.5)	15.9	(36.5)
Observations	26,551		67,121	

Notes: Table 7 summarizes the measures of lease quality across the HMDA and BISG samples.

3.4 Local Ordinances

Since the shale boom, several municipalities across Tarrant County (e.g., Fort Worth, Arlington, etc...) passed local ordinances that govern the oil and natural gas industry within their jurisdictions, strengthening the permitting and drilling regulations. Using state-level regulations as a baseline, we define several regulatory categories that describe how the local ordinances deviate from the state-level rules and increase the strength of regulations across municipalities. Among them, municipalities may require additional permitting requirements (on top of those permits approved by the Texas Railroad Commission)[28], increase the technical standards for wells and well sites, increase "setbacks," or the distance between the proposed well and residential buildings, and require additional tests capturing soil, water, and air quality before any well is drilled in order to establish a baseline level of pollution. Additionally, categories limit externalities like traffic, noise (i.e., work and fracturing hours restrictions).

To code the ordinance data, we read each city's set of zoning and municipal ordinances and define the set of possible ordinance categories, including noise, work hours, pit construction, saltwater disposal, fracturing restrictions, and environmental testing. Within each category, we establish an index based on the total number of restrictions within that category or based on the stringency of the specific regulation. For example, noise restrictions specify acceptable decibel levels at varying times of the day or by task (e.g., drilling, completion, production). We then normalize the specified level with respect to the minimum and maximum levels and weight each city's noise restriction using the total number of possible restrictions in any given locality.[29,30] When there are not varying levels in an ordinance category, we use the fraction of total restrictions feasible in a given category. For example, pit restrictions may (1) be limited to hold certain materials, (2) require a separate permit, or (3) disallow certain pits such as pits to hold drilling fluid, completion fluid, mining water, or non-commercial fluid.

These data are then merged to the housing data by year implemented and city name such that we can estimate a model accounting for the mean value of a passed ordinance, which is based on the techniques we used to quantify the stringency of local ordinances described in the prior paragraph. The ordinance measures used in the analysis are summarized in Table 8a. Most commonly, municipalities restrict fracturing hours and noise levels through local ordinances. Municipalities also frequently restrict how operators dispose of saltwater and drill near flood plains.

3.5 Violations

We analyze the relationship between lease quality and future well violation using complaint data from the Texas Commission on Environmental Quality (TCEQ). The data describe complaints specific to well operations filed with the TCEQ, which includes the filing dates, whether the complaint was related to waste, or air or water quality, resolution date,

[28] Permits may be required for (saltwater) pits, which are used to store waste water resulting from hydraulic fracturing, and disposal or injection wells.

[29] This normalizing technique is similar to the Resources for the Future analysis evaluating the disparity in regulation across states with active oil and natural gas industries (Richardson et al. (2013)).

[30] We considered other strategies for indexing local ordinances, including factor analysis, but concluded that it is more informative to analyze the disaggregated set of rules in order to better understand which rules are more binding for operators.

Table 8: Summary Statistics: Violations & Ordinances

	Mean	St.Dev.	Obs.	Min.	Max.
(a) Ordinance Summary (City)					
Air Restrictions	0.081	0.198	36	0	0.8
Soil Restrictions	0.076	0.161	36	0	0.7
Water Restrictions	0.127	0.192	36	0	0.7
Flood Restrictions	0.255	0.239	36	0	0.8
Noise	0.378	0.296	36	0	0.8
Fracture Hours	0.481	0.411	36	0	1.0
Work Hours	0.110	0.068	36	0	0.2
Pit Restrictions	0.227	0.270	36	0	0.8
Saltwater Restrictions	0.372	0.249	36	0	1.0
(b) Violation Summary (U.S. Census Tract)					
Total	4.20	5.67	105	1	23
Complaint	3.86	5.77	105	0	23
Air	4.04	5.77	105	0	23
Waste	0.15	0.43	105	0	2
Producing Well Count	66.16	144.1	1,707	1	861

Notes: (i) The top panel describes the frequency of a particular local ordinance across cities in Tarrant County; (ii) the lower panel describes the complaint/violation frequency across most U.S. Census Tracts in Tarrant County; (iii) *Producing Well Count* captures the average number of producing wells across U.S. Census Tract and month.

and a complaint location. We mapped each violation location to a U.S. Census Tract and calculated the cumulative violations occurring in a tract before and after each of the household leases were signed. Regressions relate the count of violations occurring in a tract to the average characteristics of previously signed leases in that tract. These data are summarized in Table 8b. The first row reports the average number of violations or complaints of any type reported in each tract. The following rows categorize the violations by complaints and the specific air and waste violations.

3.6 Sample Description

We conduct the Coasian analysis on properties (a) sold during our data period (2003 to 2014) and (b) where landowners sign the leases after the house purchase date. These are properties that we can use information from HMDA to infer the race, ethnicity, and income of the occupants on the date the lease was signed (or the mineral leasing date). In particular, we match 134,818 transactions from Corelogic housing data to lease observations, and of those, we observe auxiliary clauses for 77,433 observations. Among those observations with auxiliary clauses, we match 36,502 to HMDA data that provides household-level

Table 9: Mean Values Across Leased and Not Leased Properties

	Not Leased (1)	Leased (2)	Absolute Diff. (1)-(2)	T-Statistic
Housing Variables				
Living Size (sqft)	2,154	1,974	180	46.22
Land Size (sqft)	9,665	9,837	172	-5.00
% groundwater	58.6	64.3	5.7	-24.13
% well within 2km	75.5	86.2	10.6	-55.05
HMDA Variables				
White	81.1	79.2	1.9	6.90
Hispanic	18.6	25.8	7.2	-24.70
Black	9.9	10.2	0.3	-1.35
Asian	6.3	4.6	1.7	10.85
Income	82,741	76,750	5,991	17.59
BISG Variables				
% predicted White	58.4	51.8	6.6	35.09
% predicted Hispanic	20.4	25.9	5.5	-30.01
% predicted Black	12.2	14.9	2.6	-25.48
% predicted Asian	6.9	5.4	1.5	13.60
Median HH Income	79,114	72,137	6,977	52.18

Notes: (i) Table 9 compares the mean values across properties with and without leases; (ii) leased properties tend to have larger acreage but less living space, and the difference is only about 2% and 8%, respectively; (iii) the leased properties tend to by owned by more Hispanic and Black households, and non-leased properties tend to be higher-income households.

race, ethnicity, and income information.[31] Among those observations matched to HMDA, 28,416 observations report the household-level ethnicity. Finally, of the remaining sample, we match 2000/2010 U.S. Census Tract information to 24,233 observations, which is our primary estimation sample.[32] Using the BISG predictor to describe household-level race and ethnicity, we expand our estimation sample size to 64,202.

Because the samples are subject to merging along specific dimensions, Table 9 reports the differences in mean property values, income, race, and ethnicity across leased and non-leased properties. We find that, in general, non-leased properties tend to have larger living areas, are owned by individuals earning higher incomes, and are less often owned by minority households.

[31]To match the property/lease data to HMDA, the buyer of the home must have signed a lease after the house purchase.

[32]Note that we observe more term lengths than other lease terms. This is because the lease memos tend to report the term lengths, and we observe this data in the DrillingInfo dataset.

4 Environmental Justice and Coasian Bargaining

The simplest version of the Coase Theorem states that, in the absence of prohibitive transaction costs and with full information and well-defined property rights, parties will negotiate to an efficient equilibrium in the presence of an externality. The efficient equilibrium will yield the level of the externality that maximizes the joint value to the two parties, while transfers between the parties will reflect bargaining power and are subject to the constraint that both parties will find the negotiated agreement to be superior to simply walking away. We demonstrate such a bargain with a simple example. Consider the following profits (Π_i) from a shale well drilled by operator i and the utility (U_j) received by mineral rights holder j who negotiates a shale lease with that operator. D represents damages from shale gas drilling operations – damages increase with fewer precautions taken by the operator. Relaxing precautions is beneficial to the operator up to a point, while mineral rights holder utility is always decreasing in damages. T represents a transfer from the operator to the mineral rights holder, which can come in the form of bonus or royalty payments.

$$\Pi_i = \alpha_i + \beta_i D - \frac{\gamma_i}{2} D^2 - T \tag{1}$$

$$U_j = \theta_j - \phi_j D + T \tag{2}$$

The efficient bargaining outcome will maximize the joint value:

$$\underset{[D]}{max} \left[\alpha_i + \beta_i D - \frac{\gamma_i}{2} D^2 - T + \theta_j - \phi_j D + T\right] \tag{3}$$

Solving for the optimal value of D^* yields:

$$D^* = \frac{\beta_i - \phi_j}{\gamma_i} \tag{4}$$

Importantly, the optimal level of damages depends upon operator and mineral rights holder attributes that affect profits and utility, but does not depend upon any factors that only affect the abilities of the two parties to bargain over T. T can vary with bargaining power subject to the following constraints and the result is still efficient.

$$\alpha_i + \beta_i D^* - \frac{\gamma_i}{2} D^{*2} - T > 0 \tag{5}$$

$$\theta_j - \phi_j D^* + T > 0 \tag{6}$$

We may therefore indeed expect D^* to vary with characteristics of the operator and the mineral rights holder that affect $(\beta_i, \gamma_i, \phi_j)$. Some operators, for instance, might use production techniques that benefit less from damaging drilling activities. Some mineral rights holders might have a greater willingness to pay to avoid shale drilling damages if, for instance, their income is higher. We control for these possible determinants of D^* in our empirical specifications – using operator fixed effects for the former and household income for the latter. However, after controlling for these factors, we would not expect to find that the race of the mineral rights holder should matter for D^* if bargaining is efficient. Race easily could (and, given the history of race in the US south, likely would) affect bargaining power, but this would only alter T (i.e., bonus or royalty payments) in an efficient Coasian bargaining outcome. A finding that, after conditioning upon income, race also affects D^* would suggest a breakdown of the efficient bargaining process. A finding that race (specifically Hispanic ethnicity), in conjunction with English speaking ability, also affects D^* would suggest that this breakdown is the result of an information asymmetry of the sort assumed away by Coase.

In the following section, we provide empirical estimates which show that race indeed has a direct effect on the outcome of bargaining over D^*, measured by the lease clauses that, if implemented, would reduce damages. We also find that race has a direct effect on royalty and bonus payments, reflecting an impact on T.

5 Model Estimates

Below we present estimates that capture the relationships between household characteristics and the types of landowner concessions negotiated into their leases, or measures of lease quality. Our models test whether minority households sign leases with fewer landowner concessions and whether the disparity is attributable to linguistic isolation, suggesting an informational asymmetry. It is particularly problematic for efficient bargaining if, after controlling for income, education, age, and other household variables, Hispanic ethnicity matters more in neighborhoods where individuals are less likely to speak English well.

Following the lease quality analyses, we measure the efficacy of lease restrictions at deterring bad drilling behavior by measuring their relationship with future violations at the U.S. Census Tract and city levels, respectively. We use these estimates to better understand how lease quality is distributed across different racial and ethnic groups, suggesting environmental injustice in the leasing market, and to motivate policies that protect these landowners through homogenizing the leasing phase that precedes any drilling or production.

5.1 Analysis of Lease Attributes

We first present the results of linear Seemingly Unrelated Regressions in which we model a set of twelve lease attribute outcomes. We then increase the sample size by estimating the model using BISG predicted measures of race and ethnicity, which is described in Section

3, and our results are consistent across the two model types. We then allow for heterogeneous effects across observable neighborhood attributes by implementing a Moment Forest algorithm, as proposed in Asher et al. (2016) to estimate heterogeneous treatment effects.

5.1.1 Seemingly Unrelated Regression

We conduct our analysis of lease clauses treating each as a separate bargaining outcome, while recognizing that clauses are negotiated as part of a single bargaining process. As such, we treat the equations describing the outcomes of clauses i of lease l as a seemingly unrelated system of regressions as in Equation (7). Lease outcomes measured in the analysis are summarized in Table 7.[33] At the individual level, we control for (i) income, race, and ethnicity using household-level information from HMDA, (ii) a dummy variable indicating whether the house likely relies on groundwater, and (iii) measures of house and property sizes (i.e., $X_{i,l}\beta_l$ in Equation 7). We control for other household characteristics that vary spatially, such as education levels, average ages, median/average income, and other unobserved spatial heterogeneity by estimating each model with U.S. Census Tract fixed effects ($w_{i,j}$). We control for unobserved lessor attributes that may affect the lease negotiations by estimating lessor fixed effects ($\alpha_{i,f}$).[34] Finally, we control for unobserved temporal and spatial heterogeneity by estimating each model with lease year fixed effects ($\mu_{i,t}$), city fixed effects ($\delta_{i,j}$), and property coordinates (latitude and longitude).[35]

% *Spanish Speaking* measured at the U.S. Census Tract-level is used to describe linguistic isolation.[36] We expect that linguistic isolation puts a household at a severe disadvantage in mineral rights lease negotiations with experienced lessors, and we interact whether household is Hispanic with whether their neighborhood is predominantly Spanish speaking (i.e., tract-level % *Spanish Speaking*) to test our hypothesis.

$$y_{i,l} = X_{i,l}\beta_l + w_{i,j} + \alpha_{i,f} + \mu_{i,t} + \delta_{i,j} + \epsilon_{i,l}, \forall l = 1, ...L \tag{7}$$

Looking first at household and property characteristics, we find that residents of larger homes, as measured by *Living area (log sqft.)*, obtain leases that are more beneficial for the landowner. Similarly, a greater *Land area (log sqft.)* is also associated with better (and more) concessions. Conversely, *Groundwater* homes have mixed results and the vast majority

[33] We also estimate models for *clause bundles* that describe whether a lease contains a set of legal or surface clauses, and estimates are consistent with our individual lease clause-level findings.

[34] The models are also robust to including well operator fixed effects and omitting all firm-level fixed effects.

[35] City fixed effects allows us to flexibly control for local ordinances passed through home rule in Texas.

[36] Formally, the U.S. Census Bureau defines a linguistically isolated household to be one in which no person over the age of 14 speaks English at least very well. The U.S. Census asks which language is spoken at home when there is not an individual that speaks English "very well." We use the percent that reply Spanish to capture linguistic isolation.

of specifications result in insignificant estimates. Finally, greater *Income (log)* households, while often weakly significant, typically sign better leases.[37]

Controlling for these attributes, we are particularly interested in the role of race and ethnicity (Black and Hispanic), measured at the individual level, and the interaction of Hispanic with *% Spanish Speaking*. Results are described in Table 10, and we limit our analysis to lease signers who own their full estate (i.e., mineral and surface rights). Dependent variables (i.e., royalty rates, primary term lengths, and dummy variables indicating that a certain clause or category of clause is present in the lease) are listed in the first row of each panel, and regressors appear in the first column. Regression coefficients describe the likelihood of each lease clause as a function of signatory characteristics, house and property attributes, and year, spatial, and lessor fixed effects. The first rows of Tables 10a and 10b describe the effect of a Black signatory and the second and third rows describe the aggregate effects of being Hispanic. Across the board, both Black and Hispanic households are less likely to sign leases with higher royalties, shorter primary terms, and more landowner concessions (note that *Subsurface Easements* and *Free Water* are considered landowner "bads").[38]

Particularly important for the role of information asymmetries in the breakdown of Coasian bargaining, we find that the effects on Hispanic households are greater if the signatory is located in a neighborhood with a greater percentage of Spanish speakers.[39] While the household characteristics are primarily interpreted as determinants of willingness-to-pay to avoid exposure to externalities from shale gas development, race and ethnicity (conditioned upon other household characteristics) are not. If willingness-to-pay is indeed a function of race, ethnicity, and linguistic isolation after controlling for household attributes like income and education, this would suggest a breakdown of Coasian bargaining, driven largely by information asymmetry.

5.1.2 BISG Predicted Measures

HMDA data is administrative, and leasing firms may or may not be aware of their negotiation partners' race and/or ethnicity when they decide which contract to offer because, for example, the lease may be negotiated over the phone or firms make strategic leasing

[37]Note that households living in larger homes as measured by the interior square feet (i.e., living area) is positively and significantly correlated with better lease terms. In versions of the model that instead control for house size via the number of beds, the effect of income becomes much more positively and significantly correlated with better lease terms.

[38]Subsurface easements allow firms to lay pipelines and roads as necessary for transporting natural gas from the wellhead without necessarily extracting from the property itself, and thus compensating the landowner accordingly.

[39]Models excluding the Hispanic and % Spanish Speaking interactions result in larger, very significant estimates for all Hispanic households on average. In nearly all cases, the effect of being an Hispanic household is more negative than that of being a Black household.

Table 10: Coase Seemingly Unrelated Regressions with U.S. Census Tract Fixed Effects

	Royalty	Term Length	Environ.	Noise	Traffic	Surf. Rest. Protection
			(a) Part I			
Black	0.000	0.334**	-0.008	0.003	-0.006	-0.006
	(0.000)	(0.155)	(0.006)	(0.006)	(0.005)	(0.004)
Asian	-0.001	0.233	-0.005	0.014	-0.004	0.005
	(0.001)	(0.191)	(0.009)	(0.009)	(0.007)	(0.006)
Hispanic	0.001	0.259	-0.008	-0.004	-0.004	-0.010**
	(0.000)	(0.172)	(0.007)	(0.007)	(0.005)	(0.004)
Hispanic x % Spanish Speaking	-0.008***	1.644*	-0.087***	-0.088***	-0.061**	-0.006
	(0.002)	(0.936)	(0.031)	(0.034)	(0.025)	(0.016)
Income (log dollars)	0.000	-0.256***	0.000	0.002	0.003	-0.006**
	(0.000)	(0.096)	(0.004)	(0.004)	(0.003)	(0.003)
Land area (log sqft.)	0.001***	-0.346***	0.004	0.008	-0.002	0.009**
	(0.000)	(0.134)	(0.006)	(0.006)	(0.005)	(0.004)
Living area (log sqft.)	0.003***	-0.652***	0.042***	0.012	0.023***	0.017***
	(0.000)	(0.173)	(0.007)	(0.008)	(0.006)	(0.005)
Groundwater	-0.003***	1.468***	-0.050***	-0.034***	-0.013	0.000
	(0.001)	(0.303)	(0.013)	(0.013)	(0.009)	(0.008)
Observations	25,525	44,694	25,741	25,741	25,741	25,741
Pseudo-R^2	0.522	0.465	0.529	0.571	0.534	0.492
	Setback	Vertical Pugh	Post Prod. Cost	Lessee Defends	Subsurf. Easement	Free Water
			(b) Part II			
Black	-0.008	-0.011***	-0.010	-0.005	0.028***	-0.001
	(0.005)	(0.004)	(0.006)	(0.004)	(0.009)	(0.006)
Asian	-0.004	-0.006	-0.009	0.002	-0.001	-0.012
	(0.008)	(0.007)	(0.009)	(0.007)	(0.013)	(0.009)
Hispanic	-0.006	0.000	-0.007	-0.010**	-0.002	0.010
	(0.006)	(0.005)	(0.007)	(0.005)	(0.011)	(0.008)
Hispanic x % Spanish Speaking	-0.086***	-0.092***	-0.083***	-0.014	0.128**	-0.011
	(0.026)	(0.023)	(0.029)	(0.018)	(0.059)	(0.035)
Income (log dollars)	0.003	0.004	0.007*	-0.003	-0.001	-0.002
	(0.003)	(0.003)	(0.004)	(0.003)	(0.006)	(0.005)
Land area (log sqft.)	0.004	0.001	0.004	0.005	0.010	0.008
	(0.005)	(0.004)	(0.006)	(0.005)	(0.009)	(0.006)
Living area (log sqft.)	0.025***	0.020***	0.043***	0.023***	0.001	-0.013
	(0.006)	(0.006)	(0.007)	(0.006)	(0.011)	(0.008)
Groundwater	-0.006	-0.048***	-0.034***	-0.046***	0.048**	0.010
	(0.009)	(0.010)	(0.012)	(0.010)	(0.021)	(0.013)
Observations	25,741	25,741	25,741	25,741	25,741	25,741
Pseudo-R^2	0.583	0.519	0.517	0.557	0.344	0.406
City, U.S. Census Tract FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
Lessor FE	x	x	x	x	x	x
Other Controls	x	x	x	x	x	x

Notes: (i) Each column represents a separate regression of lease clause types on observable, household-level attributes; (ii) *Term Length*, *Subsurface Easement*, and *Free Water* are considered less protective clauses of landowners, so a positive coefficient suggests the household is worse off; (iii) *% Spanish Speaking* represents the fraction of households located in a U.S. Census Tract who speak Spanish as a primary language; (iv) other controls include the latitude and longitude of the property; (v) lessor fixed effects describe the firm identities signing the mineral rights leases with property owners; (vi) standard errors are robust.

decisions before approaching the landowners with contracts.[40] However, there are racial and

[40]There is significant suggestive evidence that firms strategically carve our geographic regions for proposed well development as opposed to ad-hoc leasing from house-to-house until enough acreage is amassed. Firms

Table 11: Coase Seemingly Unrelated Regressions with U.S. Census Tract FE & BISG Predictors

	Royalty	Term Length	Environ.	Noise	Traffic	Surf. Rest. Protection
(a) Part I						
Pred. Black	-0.002***	-0.046	-0.018***	-0.013*	-0.012**	-0.003
	(0.000)	(0.173)	(0.007)	(0.007)	(0.005)	(0.004)
Pred. Asian	-0.000	0.129	0.004	0.008	0.014**	0.009**
	(0.000)	(0.144)	(0.007)	(0.006)	(0.006)	(0.004)
Pred. Hispanic	-0.001*	0.630***	-0.006	-0.007	-0.009**	-0.014***
	(0.000)	(0.140)	(0.006)	(0.006)	(0.004)	(0.004)
Pred. Hispanic x % Spanish-Speaking	-0.006***	0.943	-0.137***	-0.119***	-0.054***	0.004
	(0.002)	(0.664)	(0.022)	(0.024)	(0.017)	(0.012)
Land area (log sqft.)	0.001***	-0.321***	0.006*	0.010***	-0.004	0.002
	(0.000)	(0.084)	(0.004)	(0.004)	(0.003)	(0.002)
Living area (log sqft.)	0.002***	-0.747***	0.043***	0.012***	0.025***	0.008***
	(0.000)	(0.103)	(0.004)	(0.004)	(0.003)	(0.003)
Groundwater	-0.003***	1.213***	-0.048***	-0.032***	-0.014**	-0.012**
	(0.001)	(0.191)	(0.008)	(0.008)	(0.006)	(0.005)
Observations	63,636	107,576	64,202	64,202	64,202	64,202
Pseudo-R^2	0.517	0.468	0.526	0.566	0.520	0.484

	Setback	Vertical Pugh	Post Prod. Cost	Lessee Defends	Subsurf. Easement	Free Water
(b) Part II						
Pred. Black	-0.017***	-0.013***	-0.017**	-0.010**	0.010	-0.004
	(0.005)	(0.005)	(0.007)	(0.005)	(0.010)	(0.007)
Pred. Asian	0.005	0.002	-0.001	0.016***	-0.007	-0.022***
	(0.006)	(0.006)	(0.007)	(0.006)	(0.009)	(0.007)
Pred. Hispanic	-0.007	-0.005	-0.014***	-0.021***	0.011	0.012**
	(0.004)	(0.004)	(0.005)	(0.004)	(0.009)	(0.006)
Pred. Hispanic x % Spanish-Speaking	-0.102***	-0.094***	-0.100***	0.014	0.071*	-0.044*
	(0.018)	(0.016)	(0.020)	(0.014)	(0.039)	(0.024)
Land area (log sqft.)	0.005*	0.010***	0.011***	0.004	-0.004	-0.002
	(0.003)	(0.003)	(0.004)	(0.003)	(0.006)	(0.004)
Living area (log sqft.)	0.029***	0.020***	0.039***	0.018***	-0.002	-0.011**
	(0.003)	(0.003)	(0.004)	(0.003)	(0.007)	(0.005)
Groundwater	-0.005	-0.034***	-0.036***	-0.050***	0.037***	0.016**
	(0.006)	(0.006)	(0.008)	(0.006)	(0.013)	(0.008)
Observations	64,202	64,202	64,202	64,202	64,202	64,202
Pseudo-R^2	0.572	0.508	0.511	0.548	0.324	0.389
City, U.S. Census Tract FE	x	x	x	x	x	x
Year FE	x	x	x	x	x	x
Lessor FE	x	x	x	x	x	x
Other Controls	x	x	x	x	x	x

Notes: (i) Each column represents a separate regression of lease clause types on observable, household-level attributes; (ii) *Term Length*, *Subsurface Easement*, and *Free Water* are considered less protective clauses of landowners, so a positive coefficient suggests the household is worse off; (iii) *% Spanish Speaking* represents the fraction of households located in a U.S. Census Tract who speak Spanish as a primary language; (iv) *Predicted* race and ethnicity characteristics are generated through the BISG algorithm using landowner surnames; (v) other controls include the latitude and longitude of the property; (vi) lessor fixed effects describe the firm identities signing the mineral rights leases with property owners; (vii) standard errors are robust; (viii) the models control for household characteristics, including median income, via U.S. Census Tract fixed effects.

need to consider their own and other existing nearby well development, proximity to pipelines, and other production concerns. We account for these measures of spatial heterogeneity by estimating a large set of fixed effects that control for wellpad area and city boundaries, in addition to our temporal, lease year fixed effects.

ethnic indicators channelled through individuals' first and last names and the neighborhood in which they live. The Consumer Financial Protection Bureau, tasked with monitoring fair lending practices, utilizes a Bayesian Improved Surname Geocoding (BISG) algorithm that is based on U.S. Census data to monitor discriminatory lending practices in settings where there is no borrower demographic data. We borrow their method of classifying individuals as likely Hispanic, Black, White, Asian, etc... using their surnames and home address zip codes to assign probabilities as described in Section 3.3.

Estimates utilizing the BISG measures are reported in Table 11 using a format that mirrors the presentation of the HMDA estimates. The predicted measures are individualized based on signatory surnames, and models control for neighborhood characteristics, such as median household income, via U.S. Census tract fixed effects. As before, we also interact the probability that a household is Hispanic with *% Spanish Speaking* to test our linguistic isolation hypothesis in the expanded setting. Compared to HMDA, the BISG method to identify household characteristics allows us to expand our sample by more than twice the number of observations.

Across the board, both Hispanic and Black households receive lower royalty rates, longer primary terms, and fewer beneficial landowner concessions in their negotiated contracts. Further, the second row testing our linguistic isolation hypothesis confirms that Hispanics living in tracts with high levels of Spanish speakers receive even fewer concessions than those that do not. To interpret the results for an added *Environmental* clause, as the probability an individual's surname is Hispanic increases, the lease is less likely to include environmental clause protections (-0.006 (0.006)) and the effect is significantly more negative (-0.137 (0.022)) if the individual lives in an area where a high concentration of residents are Spanish speaking. The interpretation is strikingly similar to what we estimate for the *Environmental* clause using HMDA, household-level characteristics. More generally, the BISG estimates mimic those of the SUR and, in many cases, the race and ethnicity characteristic estimates have larger magnitudes and are still significant.

5.1.3 Bonus Payments

The final lease component is the bonus payment, which is a fixed payment to the landowner on the mineral leasing date. Bonuses are a component of T in our theoretical model of efficient bargaining. Firms are not required to report bonus payments, and only 0.6% voluntarily report them in the Drillinginfo data. We supplement the small sample size by extracting a term from the pdf which states that the extension bonus is the same as the original bonus, thereby permitting us to supplement the bonuses with the more widely reported extension bonuses.[41] Using the expanded bonus sample, we report estimates relating

[41] There is also a high, 0.9, correlation between reported and extension bonuses within the 0.6% sample.

Table 12: Race/Ethnicity and Bonus Regressions

HMDA Sample	Bonus (log)	Bonus/ Acre (log)	*BISG Sample*	Bonus (log)	Bonus/ Acre (log)
Black	0.053	0.053	Pred. Black	0.064*	0.064*
	(0.037)	(0.037)		(0.038)	(0.038)
Asian	0.031	0.031	Pred. Asian	-0.018	-0.018
	(0.051)	(0.051)		(0.042)	(0.042)
Hispanic	0.018	0.018	Pred. Hispanic	-0.068	-0.068
	(0.048)	(0.048)		(0.043)	(0.043)
Hispanic x	-0.405	-0.405	Pred. Hispanic x	-0.082	-0.082
% Spanish Speaking	(0.281)	(0.281)	% Spanish Speaking	(0.231)	(0.231)
Income (log dollars)	-0.034	-0.034	Median HH Income	U.S. Census	
	(0.030)	(0.030)		Tract FE	
Land area (log sqft.)	0.050	1.050***	Land area (log sqft.)	0.023	1.023***
	(0.044)	(0.044)		(0.025)	(0.025)
Living area (log sqft.)	0.016	0.016	Living area (log sqft.)	0.060*	0.060*
	(0.051)	(0.051)		(0.031)	(0.031)
Groundwater	0.149	0.149	Groundwater	0.014	0.014
	(0.178)	(0.178)		(0.078)	(0.078)
Observations	3,347	3,347		8,158	8,158
Pseudo-R^2	0.888	0.904		0.887	0.903
City, U.S. Census Tract FE	x	x		x	x
Year FE	x	x		x	x
Lessor FE	x	x		x	x
Other Controls	x	x		x	x

Notes: (i) The bonus regressions mimic the analysis in Tables 10 and 11; (ii) *Bonus* includes both reported bonuses and extension bonuses, which are highly correlated with reported bonuses (0.9); (iii) *Bonus/Acre* scales bonuses by lease parcel acreage; (iv) the first two columns of estimates restrict to the HMDA sample with household characteristics, while the second set of columns utilize BISG predicted race/ethnicity sample; (v) other controls include the latitude and longitude of the property; (vi) lessor fixed effects describe the firm identities signing the mineral rights leases with property owners; (vii) standard errors are robust; (viii) the BISG models control for household characteristics, including median income, via U.S. Census Tract fixed effects.

race and ethnicity to the bonus (log) and bonus per acre (log) amounts in Table 12. The first two columns use HMDA data to identify race and ethnicity, while the last two columns use BISG predicted values. Both HMDA and BISG models find that the effects of *Hispanic*% Spanish Speaking* are negative and larger than the effects of *Hispanic* alone, while the effects of *Black* are positive. Additional analysis reveals that the positive effects attributed to Black households are dependant on whether we estimate the model with lessor fixed effects, and when lessor fixed effects are omitted, the effect of being a Black household is negative.[42] In general, we find that our bonus estimates point in a consistent direction to our findings across all other lease dimensions, even though the estimates are less precise. Further, we

[42]Models that interact the lessor fixed effects with whether a household is Black confirms that the bonuses are incrementally lower for the majority of lessors negotiating leases with Black households relative to leases negotiated with white households.

demonstrate that it is unlikely minority households are being monetarily compensated for signing fewer legal and environmental protections: in fact, their leases appear to be worse along all dimensions.

5.1.4 Minority well exposure

Whether or not minority households negotiate differing lease terms matters more if minority households live nearer to drilling activity, exposing them to greater environmental harm and negative externalities. While we summarize well-exposure measures by race and ethnicity in Table 2a, we also run simple well exposure regression models that control for other observable and unobservable neighborhood and household characteristics. In particular, Table 13 reports estimates for linear-probability regressions of well exposure on race, ethnicity, income, house attributes, and year and spatial fixed effects. Across the board, the regression estimates suggest that minority households are more likely to live within two kilometers of a well (compared to white households).[43]

Table 13: Race/Ethnicity and Well Exposure Regressions

	HMDA		BISG	
	(1)	(2)	(3)	(4)
Hispanic	0.023***	0.004	0.012***	0.005**
	(0.003)	(0.002)	(0.003)	(0.002)
Black	0.027***	0.005*	0.027***	0.006*
	(0.004)	(0.003)	(0.004)	(0.003)
Asian	0.018***	0.002	0.020***	0.006*
	(0.006)	(0.004)	(0.004)	(0.003)
Income (log)	-0.022***	-0.006***	-0.086***	
	(0.003)	(0.002)	(0.003)	
Groundwater	-0.060***	-0.010*	-0.048***	-0.004
	(0.009)	(0.006)	(0.006)	(0.004)
Land area (log sqft.)	-0.025***	-0.010***	-0.029***	-0.011***
	(0.004)	(0.003)	(0.002)	(0.002)
Observations	47,439	47,439	115,368	115,368
R-squared	0.363	0.678	0.350	0.634
Year FE	x	x	x	x
City, U.S. Census Tract FE		x		x

Notes: (i) Each column is a regression of a dummy describing households with wells located within 2000 meters on covariates describing the race, ethnicity, and income level of the head of the household; (ii) the first two columns use the HMDA measures of household characteristics, while the last two use BISG predicted race and ethnicity and U.S. Census Tract-level median household income; (iii) all minority categories have a positive relationship with exposure that lessens in statistical significance when controlling for U.S. Census Tracts; (iv) standard errors are robust.

[43]We lose statistical significance in the models estimating U.S. Census Tract fixed effects (columns 2 and 4), which we attribute to lower variation in exposure measures as tracts are quite small, especially in population-dense parts of Tarrant County.

5.1.5 Moment Forest & Heterogeneous Effects

To further demonstrate how the negative implications of being a minority household varies with other neighborhood characteristics, i.e., the heterogeneous effects on lease quality that stem from being an Hispanic or Black household, our final lease quality analysis uses the Chi-square Automatic Interaction Detector (CHAID) in the context of the Moment Forest estimator proposed by Asher et al. (2016).[44,45] Expanding on Asher et al. (2016), we implement the Moment Forest estimator in a setting with multiple outcomes, or in our case, contract attributes. With this method, we are able to estimate how the minority effects differ along with other observable, neighborhood characteristics without imposing modeling or functional form assumptions on the relationship (e.g., a linear relations as implied by our coefficient estimates on *Hispanic* x *% Spanish Speaking*). Moreover, it allows us to enrich the dimensions of heterogeneity in the model to include other, potentially determinative factors like educational attainment and median household income, as well as linguistic isolation. The new method finds consistent evidence that more linguistically isolated Hispanic households receive lower lease quality, and additionally, minorities that live in neighborhoods with lower income and educational attainment tend to sign leases with fewer protective clauses.

Building on the model presented in Section 5.1.1, we now allow for more flexible coefficient estimates that depend on other observable components of the data denoted by Z, which includes U.S. Census Tract-level variables *% Spanish Speaking*, *Median Household Income*, and *% College Graduated*.[46]. These tract-level characteristics in Z are called the *splitting variables*, and Equation (8) describes the lease outcomes, $y_{i\in l}$, as function of the splitting variables and coefficients that vary with the values of them, i.e. $X_{i,l}\beta_l$.

$$y_{i,l,b} = X_{i,l}\beta_l\omega_{l,b}(Z) + \mu_{t,l}\omega_{l,b}(Z) + \delta_{j,l}\omega_{l,b}(Z) + \epsilon_{i,l}, \forall l = 1, ...L, b = 1, ...B \qquad (8)$$

The Moment Forest estimator is designed to search for splits of the observations in the data across all splitting variables that result in similar effects between being a minority household and the outcome variable, or measure of lease quality. It searches for these splits

[44] The CHAID decision tree analysis is a recursive partition algorithm that was originally part of Gordon Kass's PhD thesis in Kass (1980). Work in Biggs and Suen (1991) expands Kass's framework by introducing improved partition selection methods. Both of these papers were used to build the code in STATA that implements the CHAID "tree finder," which is authored by JosephAsher et al. (2016) Luchman at BeMavior Statistics Lead.

[45] Asher et al. (2016) presents econometric consistency and convergence results for the two-step Moment Forest estimator. It also demonstrates the method by estimating heterogeneous treatment effects in a regression discontinuity design.

[46] SUR models that include an interaction between Hispanic and/or Black households and levels of neighborhood *Median Household Income*, *% College Graduated*, and *% Spanish Speaking* reveal there is a significant statistical relationship between a household's racial and ethnic composition and these neighborhood attributes.

based on pairing different quantiles of % *Spanish Speaking*, *Median Household Income*, and % *College Graduated* and retaining those splits that best fit the data. For example, a potential split of the data is characterized in Figure 1 for a lease clause that restricts noise. In this example, the model selects splits of the data based on the deciles of the three Z variables. In terms of the relationship between being an Hispanic household and receiving a noise restriction, the Moment Forest estimator finds that the data is better fit:

- When observations from the 1^{st} through 4^{th} decils of % *Spanish Speaking* are estimated separate from observations in the 5^{th} through 10^{th} deciles.

- When higher % *Spanish Speaking* deciles are estimated with lower *Median Household Income* deciles (i.e., 1^{st} through 5^{th} following the path on the left side of Figure 1).

- When the effects of % *College Graduates* are estimated separately for the 1^{st} through 3^{rd} deciles and 4^{th} through 5^{th} (i.e., the split that occurs in the final row of Figure 1).

Moving down the right-hand side of Figure 1, the Moment Forest estimator finds a more common relationship between being an Hispanic household and receiving a noise restriction among households with lower deciles of % *Spanish Speaking* and higher deciles of both *Median Household Income* and % *College Graduates*.

Figure 1: Sample Forest for Noise Clause
$\omega_{i\in noise}(Z|data_1^b)$

Classification dataset: $data_1$

- % Spanish Speaking (5,6,7,8,9,10)
  - Median Household Income (1,2,3,4,5)
    - % College Graduates (1,2,3)
    - % College Graduates (4,5)
- % Spanish Speaking (1,2,3,4)
  - Median Household Income (6,7,8,9,10)
    - % College Graduates (6,7,8,9,10)

In practice, the observations in the data are split in a way that results in more similar relationships between being a minority household and receiving leases of a particular quality based on the CHAID-algorithm. It uses a logistic model to compare model fit across different combinations of % *Spanish Speaking*, *Median Household Income*, and % *College Graduates* for a given lease attribute, which amounts to about $2^{20} \text{x} 2^{20} \text{x} 2^{20}$ potential models in our setting where each U.S. Census tract-level characteristics is broken into twenty quantiles. Instead of testing the model fit for every quantile combination, the model uses a bootstrapping technique to calculate the distribution of estimated effects of being an Hispanic household along the domain of tract-level characteristics.[47] Once the model specifies how to stratify the data for each lease attribute, the data are stacked and used to estimate single heterogeneous effects for each group and each lease attribute.[48,49] Each model estimates l (clause)-equation specific city and year fixed effects, along with other household (income), property (groundwater and land and living areas), and U.S. Census Tract-level characteristics (population > 60 and < 19).

As in Section 5.1.1, the Moment Forest models are estimated using a set of (1) individual clauses in addition to a discretized royalty and primary term length, as well as for (2) four categories of clause types in addition to discretized royalty and primary term length.[50] For some specifications, we are even able to include a discretized bonus measure using our small bonus sample. We present two sets of results wherein the first set estimates the model along one dimension of heterogeneity, namely, the % *Spanish Speaking*, and the second set allows estimates to vary across all three U.S. Census tract-level dimensions.

The first set of results are presented in Figure 2. Each sub-figure plots the estimated effects of being Hispanic on a given lease quality attribute, wherein the heterogeneous coefficient estimates are represented with dots.[51] The dashed blue, horizontal lines mark the average effects of being an Hispanic household when not controlling for heterogeneous effects. Finally, beneath each plot are correlations describing the relationship between the sets of

[47]To begin, we sub-divide the total data sample into two where the first is used to select the decision tree structure using the CHAID algorithm, $data_1$, and the second is used to estimate the heterogeneous effects, $data_2$. Using $data_1$, we bootstrap sample observations for each lease outcome, l, and run the CHAID algorithm *splitting* along three, neighborhood characteristics dimensions in Z. The output sub-categorizes $data_1$ based on the twenty quantiles for each of the three neighborhood characteristics, or $\omega_{i \in l}(Z|data_1^b)$, where ω denotes the selected household observations with Z characteristic quantiles that best predict a given lease clause, l, is signed in a bootstrapped dataset, b (notation used to describe the example "tree" in Figure 1). The CHAID algorithm searches for the *optimal CHAID decision tree* using 25% of the bootstrapped data for each bootstrap draw and for each lease attribute. The outcome variables are discrete, and the CHAID algorithm uses a (multinomial) logistic model to estimate the chi-squares used to compare potential models or decision tree structures.

[48]After searching for an optimal $\omega_{i \in l}(Z|data_1^b)$ for each lease attribute, we then jointly estimate the set of lease outcome coefficients for each relevant sub-classification of the data using $data_2$.

[49]We find an optimal CHAID decision tree for each bootstrapped sample, and we can use the bootstrapped estimates to construct standard errors, as described in Asher et al. (2016).

[50]Clause bundles are constructed variables that describe whether a lease contains sets of legal, surface, externality, and "bads" clauses that are based on groupings of the individual clauses used to estimate the SUR and BISG models.

[51]The estimates are consistent when lease attributes are measured as individual clauses, as well.

heterogeneous, Moment Forest estimates of being an Hispanic household and the level of % *Spanish Speaking* in the U.S. Census Tract. In Panel 2a, the correlation between royalty estimates and % *Spanish Speaking* is 0.31, which suggests that the effect of being Hispanic on royalty rates is more negative in tracts with more Spanish speakers. Similarly, there is a negative correlation in Hispanic estimates for bonuses and externality, surface, and legal clauses, whereas there is a positive correlation for "bads" clauses that benefit the firm. These results are consistent with the findings in our SUR and BISG frameworks.

Figure 3 reports the histograms of coefficient estimates when the effects of being an Hispanic household vary along three neighborhood dimensions: % *Spanish Speaking*, *Median Household Income*, and % *College Graduates*. Histograms from the Moment Forest estimator are plotted for both Black (black outline) or Hispanic (green fill) households, and the dashed black lines describe the average effect of being an Hispanic household (without heterogeneity).[52] Consistent with the SUR and BISG modeling frameworks, the distributions of Hispanic and Black estimates are almost all negative for landowner concessions reported for royalty rates, bonuses, externality clauses, surface clauses, and legal clauses). Conversely, the distribution of effects are more positive for "bad" clauses reported in panel (c). Moreover, the Hispanic and Black estimates are mostly overlapping with two exceptions: in panel (a), Hispanic households receive slightly higher royalties than Black household, though both still receive lower royalties than white households (the baseline category), and in panel (c) Hispanic households appear more likely to receive "bads" clauses in their leases.

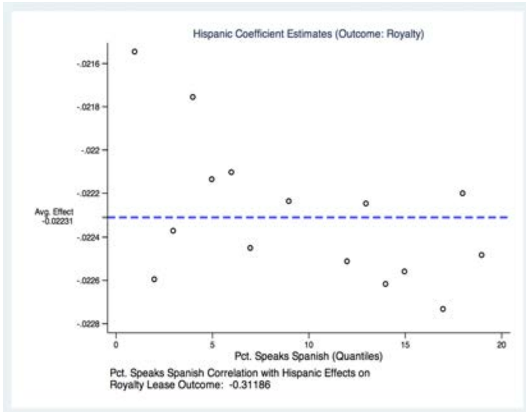
Similar to Figure 3, we report correlations in Table 14 between Hispanic estimates and the levels of three dimensions of heterogeneity. We find that estimates are more negative in U.S. Census Tracts with more Spanish speakers and less negative in tracts with higher median household incomes and more college graduates with two exceptions where the clause is considered more favorable to the firm, i.e., term length and "bads" bundle. Table 14 reports correlations for the clause model wherein subsurface easements, no litigation, free water, and longer term lengths are categorized as "bads" for landowners and thus are a disamenity when include in the leases. We find that there are greater negative effects of being Hispanic in U.S. Census Tracts comprised of more Spanish speakers, and smaller effects in areas with higher median incomes and more college graduates.

5.2 Violations

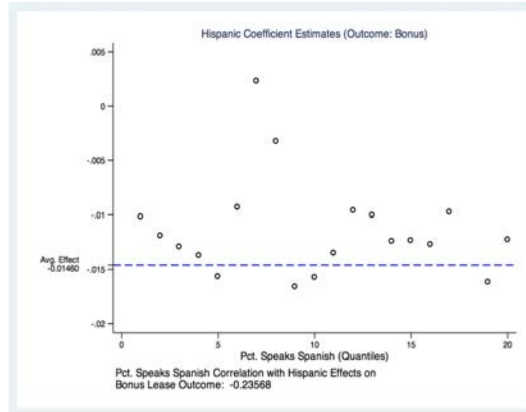
In the above analysis, we find that lease quality is not distributed evenly across the population in our data. We also find that race and ethnicity explain some of this variation, even after controlling for income, suggesting environmental injustice. Moreover, the quality

[52]There is a unique mean estimate for each CHAID-optimal split of the data for each lease outcome and summarized across all bootstrapped samples.

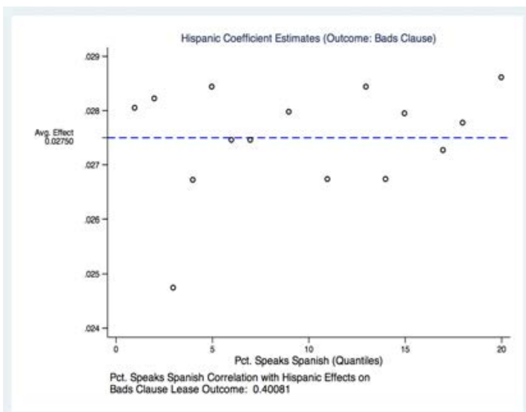
Figure 2: Heterogeneous Effects Varying with the Percent Spanish Speaking



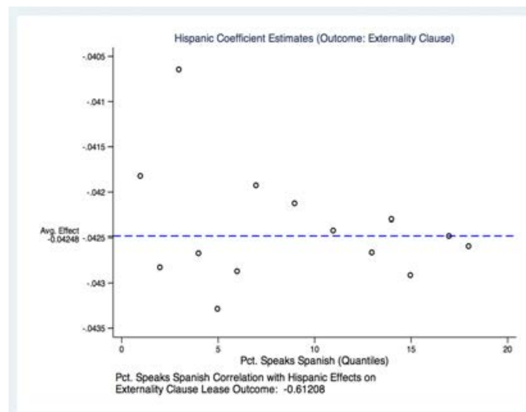
(a) Royalty Estimates



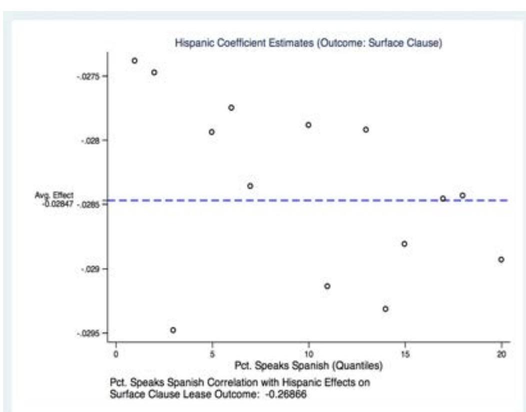
(b) Bonus Estimates



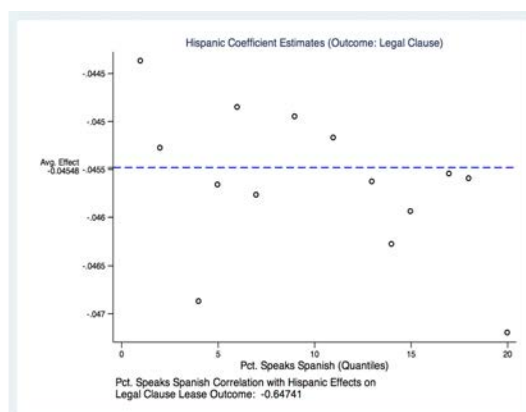
(c) Bads Clause Estimates



(d) Externality Clause Estimates



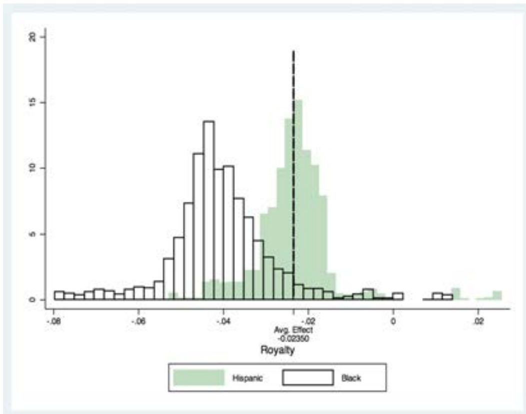
(e) Surface Clause Estimates



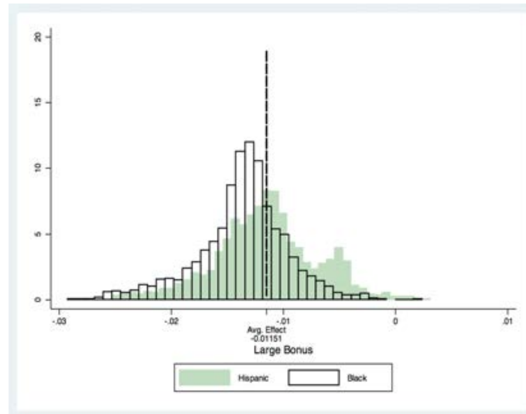
(f) Legal Clause Estimates

Notes: (i) Each dot describes an estimate of the effect of being Hispanic on a lease outcome; (ii) dots are estimated for each of the 20 quantiles of percent Spanish speaking in the data; (iii) the correlation between the estimates and the percent Spanish speaking are reported below each sub-figure, and a negative value suggests that Hispanic households are less likely to receive a particular lease outcome when they live in a region with more Spanish speakers; (v) the blue, dashed lines report the average effect of being Hispanic on lease outcomes; (vi) all models are estimated with clause interacted measures of income, all race and ethnic categories, land size, house size, likely accessing ground water, U.S. Census Tract-level measure of individuals > 65, and year and city fixed effects.

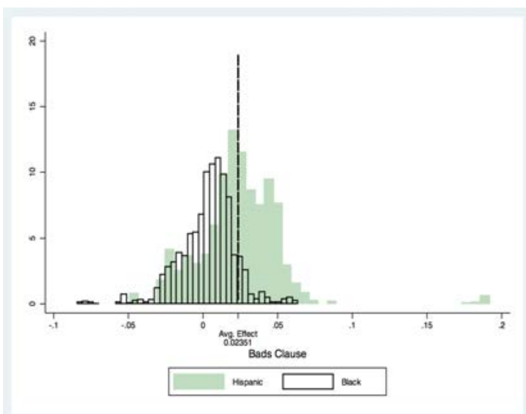
Figure 3: Heterogeneous Effects Varying with the Percent Spanish Speaking, Median Income, and Percent College Graduates



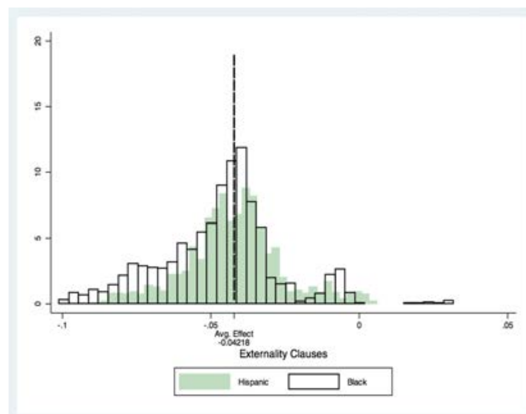
(a) Royalty Estimates



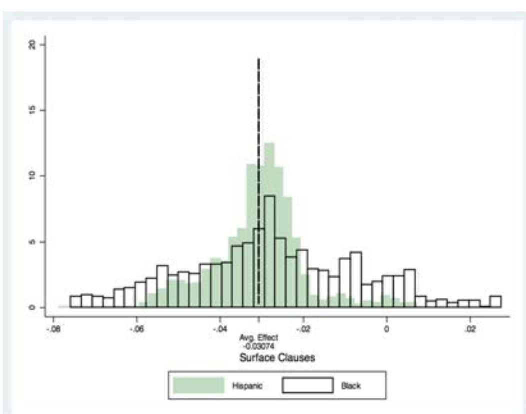
(b) Bonus Estimates



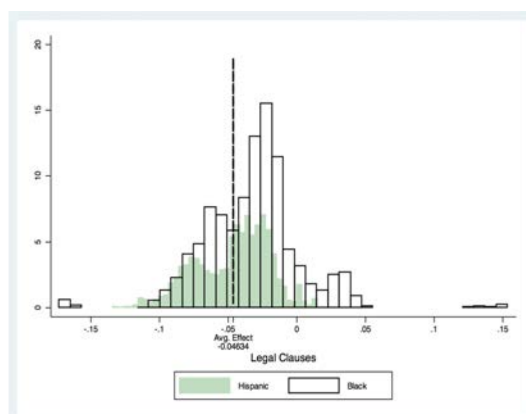
(c) Bads Clause Estimates



(d) Externality Clause Estimates



(e) Surface Clause Estimates



(f) Legal Clause Estimates

Notes: (i) Each histogram plots the frequency of mean, bootstrapped estimates of the effect of being a Hispanic or Black household on a lease outcome; (ii) the data are split into groups (via the Chaid-forest algorithm) across 20 quantiles describing the percent Spanish speaking, median income, and percent of college graduates measured at the U.S. Census Tract-level; (iii) the dashed lines denote the average effect of being a Hispanic or Black household estimated using the full data and no heterogeneous effects; (iv) all models are estimated with clause interacted measures of income, all race and ethnic categories, land size, house size, likely accessing ground water, U.S. Census Tract-level measure of individuals > 65, and year and city fixed effects.

Table 14: Heterogeneous Coefficient Correlations with Household Attributes

(a) Measuring Lease Quality in Bundles

	% Spanish Speaking	Median HH Income	% College Graduates
Large Bonus	-0.296	-0.023	0.066
Royalty	-0.295	-0.068	0.025
Term Length	0.305	-0.008	-0.214
Externality Bundle	-0.348	0.036	0.156
Surface Bundle	-0.295	0.017	0.029
Legal Bundle	-0.397	0.168	0.313
Bads Bundle	0.226	-0.143	-0.350

(b) Measuring Lease Quality in Clauses

	% Spanish Speaking	Median HH Income	% College Graduates
Royalty	-0.146	0.037	0.035
Term Length	0.369	-0.080	-0.464
Environmental Clause	-0.206	-0.142	0.058
Noise Clause	-0.227	0.074	0.144
Vertical Pugh	-0.153	0.152	0.012
Post-Production Costs	-0.337	0.305	0.291
Lessee Defends	-0.122	0.156	0.160
No Surface Access	-0.234	0.254	0.271
Surface Restriction	-0.177	-0.062	-0.053
Setback	-0.193	0.176	0.244
Subsurface Easement	0.115	-0.046	-0.256
No Litigation	0.144	-0.149	-0.202
Free Water	0.277	-0.166	-0.390

Notes: (i) Table 14a describes the correlations between the coefficients describing the effects of being a Hispanic household on lease outcomes, measured in bundles, and U.S. Census Tract attributes; (ii) Table 14b describes the correlations between the coefficients describing the effects of being a Hispanic household on lease outcomes, measured in clauses, and U.S. Census Tract attributes; (iii) we expect positive correlations between landowner protections and greater median household income and percent college graduates, while we expect negative correlations with percent Spanish speaking; (iv) among "bads" like term length, subsurface easements, free water, and no litigation, we expect the opposite, aka positive correlations with percent Spanish speaking and negative correlations otherwise.

of the protections in municipal ordinances is positively correlated with the quality of lease protections. However, inequitable distribution of protections only matters practically if we find that legal restrictions stipulated in leasing documents or ordinances effectively deter bad drilling behavior that leads to negative externalities borne by households located near to drilling activity. We build suggestive evidence that lease contents matter for future drilling behavior by estimating regressions relating the prevalence of landowner protections at the U.S. Census Tract-level to future violations.

We begin by demonstrating that, on average, minorities tend to live in cities with fewer

Table 15: Ordinance & Household Characteristic Correlations

(a) Household Characteristics Correlations

City-level mean	Ground-water	Living area	Land area	Drilling 2km	Income	White	Hispanic	Asian
Living area (log sqft.)	0.251							
Land area (log sqft.)	-0.101	0.491						
Drill within 2km	-0.265	-0.637	-0.105					
Income (log)	0.183	0.974	0.504	-0.666				
White	0.019	0.502	0.490	-0.366	0.630			
Hispanic	-0.192	-0.767	-0.358	0.472	-0.838	-0.801		
Asian	0.314	0.276	-0.336	-0.299	0.232	-0.212	-0.165	
Black	0.134	-0.084	-0.270	0.282	-0.235	-0.776	0.334	0.194

(b) Ordinance & Household Characteristic Correlations

City-level mean	Air	Soil	Water	Flood	Noise	Frack Hours	Work Hours	Pit Rest.	Saltwater Rest.
Black	0.067	-0.206	-0.021	0.361	0.259	0.287	0.180	-0.022	0.348
Asian	0.511	0.450	0.511	0.498	0.440	0.259	0.375	0.455	0.604
Hispanic	-0.288	-0.338	-0.329	-0.214	-0.412	-0.165	-0.338	-0.228	-0.413
White	0.045	0.230	0.096	-0.150	0.032	-0.120	0.037	0.049	-0.035
Income (log)	0.541	0.287	0.507	0.187	0.456	0.227	0.339	0.321	0.464
Drill within 2km	-0.576	-0.518	-0.704	-0.030	-0.256	-0.310	-0.302	-0.508	-0.253
Land area (log sqft.)	0.160	-0.066	0.010	-0.333	-0.210	-0.243	-0.360	-0.137	-0.153
Living area (log sqft.)	0.585	0.209	0.510	0.239	0.480	0.239	0.329	0.293	0.506
Groundwater	0.202	0.236	0.169	0.276	0.227	0.199	0.351	0.349	0.416
Observations	30								

Notes: (i) Table describes the city-level, mean correlations across household characteristics; (ii) Table describes the raw correlations between mean household characteristics in a city and the stringency of the passed ordinance; (iii) Hispanic and Black households own smaller homes and have lower average income levels; (iv) larger properties (*Land Size*) have less stringent regulations, which is consistent with more rural property not falling under municipal jurisdiction; (v) Hispanic households are less likely to be protected with more stringent ordinances.

local ordinances on the books.[53] Table 15 summarizes raw correlations between city-level household characteristics and the prevalence of particular ordinance restrictions. Notably, the Hispanic correlations in Table 15b are all negative, while the income correlations are all positive. Black households also tend to live in cities with more local ordinances, while cities with more drilling activity overall tend to have fewer local ordinances that regulate the drilling activity.

We then estimate a regression model that captures the effect of the average level of lease protections on the incidence of future violations, and the results are reported in Table 16.

[53]Note that our analysis accounts for the timing of the introduction of the ordinance.

Table 16: The Effect of Lease Clauses on Violations

	Violations After Signing				Violations per Producing Well			
	Total	Complaint	Air	Waste	Total	Complaint	Air	Waste
Legal bundle	-0.173**	-0.174**	-0.170**	-0.012	-1.201***	-1.077***	-1.153***	-0.034***
	(0.073)	(0.070)	(0.072)	(0.008)	(0.118)	(0.107)	(0.117)	(0.001)
Surface clause bundle	-0.413***	-0.363***	-0.384***	-0.024***	-0.922***	-0.831***	-0.895***	-0.020***
	(0.033)	(0.032)	(0.032)	(0.004)	(0.053)	(0.048)	(0.052)	(0.001)
Externality Bundle	-0.495***	-0.483***	-0.500***	-0.015*	0.012	0.016	0.038	-0.016***
	(0.077)	(0.074)	(0.076)	(0.008)	(0.125)	(0.113)	(0.123)	(0.001)
Environmental Clause	-0.102***	-0.099***	-0.098***	0.007*	-0.072	-0.060	-0.057	-0.010***
	(0.035)	(0.034)	(0.035)	(0.004)	(0.057)	(0.052)	(0.057)	(0.001)
Noise Restriction	-0.421***	-0.386***	-0.410***	-0.024***	-0.134*	-0.121*	-0.119*	-0.010***
	(0.042)	(0.041)	(0.042)	(0.005)	(0.069)	(0.062)	(0.068)	(0.001)
Fence	0.468***	0.179**	0.368***	0.112***	0.399***	0.385***	0.397***	0.001
	(0.079)	(0.076)	(0.078)	(0.009)	(0.127)	(0.116)	(0.126)	(0.001)
Full set of clauses bundle	-0.393***	-0.358***	-0.374***	-0.022***	-0.731***	-0.656***	-0.698***	-0.023***
	(0.053)	(0.052)	(0.053)	(0.006)	(0.087)	(0.079)	(0.086)	(0.001)

Notes: (i) Each estimate and standard error represents a separate regression of violations on the U.S. Census Tract-level frequency of lease clauses or at the city level; (ii) the frequency of lease clauses in a U.S. Census Tract and city are measured each year and month; (iii) violations are measured for those occurring after leases are signed; (iv) standard errors are robust.

The models are designed to test whether cities with more local ordinances are less likely to have well violations that can lead to environmental harm (or are negative externalities). Each estimate and standard deviation in Table 16 represents a separate regression. In particular, for each U.S. Census Tract, we calculate the monthly, cumulative mean incidence of a particular landowner protection. Similarly, we match our aggregate lease quality measures to cumulative violations (by month) occurring after the leases are signed. We find that higher frequencies of landowner protections are negatively related to future violations as indicated by the negative and statistically significant estimates (the *Fence Clause* is an exception).[54] Our results are consistent across different measures of landowner protections (rows) and violation types (columns), which include landowner complaints and air and waste violations measured by U.S. Census Tract (first four columns) and violations per well by U.S. Census Tract (last four columns).[55]

We estimate similar relationships between U.S. Census Tract-level violations and levels of

[54]The *Fence Clause* requires a physical barrier be built around the well site. In contrast to the other landowner protections like legal restrictions, environmental, and noise clauses, fences appear ineffective at mitigating future negative drilling disamenities (or could be a "protection" suggested by operators who are more likely to commit future violations).

[55]The first four columns use violation count as the dependent variable while the last four columns use violations per producing well as the dependent variable.

Table 17: The Effect of Ordinances on Violations

	Violations After Signing				Violations per Producing Well			
	Total	Complaint	Air	Waste	Total	Complaint	Air	Waste
Air Restriction	-0.743***	-0.698***	-0.705***	-0.036	-0.359*	-0.337*	-0.357*	-0.001
	(0.137)	(0.134)	(0.136)	(0.026)	(0.194)	(0.180)	(0.192)	(0.002)
Water Restriction	-0.922***	-0.845***	-0.855***	-0.075***	-0.512***	-0.485***	-0.509***	-0.002
	(0.137)	(0.134)	(0.136)	(0.026)	(0.194)	(0.180)	(0.193)	(0.002)
Pit Restriction	-0.268***	-0.180*	-0.189**	-0.075***	-0.472***	-0.438***	-0.468***	-0.003**
	(0.095)	(0.093)	(0.095)	(0.018)	(0.134)	(0.125)	(0.133)	(0.001)
Soil Restriction	-0.974***	-1.059***	-1.068***	0.098***	-0.561**	-0.539**	-0.569**	0.009***
	(0.176)	(0.172)	(0.175)	(0.033)	(0.250)	(0.232)	(0.248)	(0.002)
Saltwater Restriction	-0.349***	-0.359***	-0.357***	0.003	0.373***	0.326***	0.370***	0.004***
	(0.089)	(0.087)	(0.088)	(0.017)	(0.125)	(0.116)	(0.124)	(0.001)
Noise	0.123	0.158**			0.622***	0.551***		
	(0.078)	(0.077)			(0.110)	(0.102)		
Fracturing Hours	-0.178***	-0.172***			0.042	0.029		
	(0.057)	(0.056)			(0.081)	(0.075)		
Work Hours	-0.940***	-0.959***			0.183	0.118		
	(0.311)	(0.305)			(0.440)	(0.408)		
Ordinance Dum.	0.005	-0.004	-0.001	0.004*	-0.012	-0.027***	-0.027***	0.017***
	(0.018)	(0.018)	(0.018)	(0.002)	(0.009)	(0.009)	(0.009)	(0.003)

Notes: (i) Each estimate and standard error represents a separate regression of violations on the U.S. Census Tract-level rate of ordinances passed at the municipal level by tract; (ii) violations are measured for those occurring after leases are signed; (iii) standard errors are robust.

ordinance stringency to test whether there is a deterrent relationship between ordinances and future drilling behavior. Table 17 reports negative and statistically significant relationships between *Air*, *Water*, and *Pit* restrictions. Further, restrictions on work hours are negatively correlated with complaints, which is intuitive.[56]

6 Conclusion

We analyze the determinants of privately negotiated lease characteristics that arise from a bargaining process that assigns rights to extract shale gas resources. With the dramatic growth in U.S. shale gas development over the last decade (particularly in populated areas), lease negotiations have become an important part of the energy landscape, and constitute a primary potential source of benefits for homeowners through royalty and bonus payments. This is significant, as there are also many potential drawbacks to living near a shale gas well enumerated throughout the paper. These characteristics of the oil and gas private bargaining process share many features of a classic Coasian bargaining framework, wherein a Coasian

[56]Specifications using the count of future violations as the dependent variable and controlling for U.S. Census Tract-level fixed effects are similar, whereby including tract-level fixed effects controls for the mean, tract-level lease quality signed.

bargaining process would suggest an efficiency result that would reduce the need for costly government oversight of the leasing process. We use our data to test whether oil and gas lease negotiations do indeed exhibit characteristics of Coasian *efficiency* in one of the most active shale gas counties in the U.S. - Tarrant Co., Texas.

Measures of wealth, such as household income, land area, and housing square footage, are important determinants of willingness-to-accept payment in exchange for experiencing the nuisances associated with shale gas development. Our results suggest that these factors do indeed affect bargaining outcomes. This result is indicative of Coasian efficiency - those who we would expect to have a larger willingness-to-accept negotiate for stronger lease terms. However, the story does not end there. The argument in favor of Coasian efficiency becomes harder to sustain when we find similar results for race and ethnicity. Flexibly conditioning upon income and other sociodemographic characteristics, like obtained education and age, it is not clear why we would expect different racial and ethnic groups to have different willingnesses-to-accept to avoid environmental harm, so the fact that we find significant differences in lease quality across racial and ethnic groups suggests a realization of environmental injustice.

The question then arises as to what enables the observed injustice. One possibility is the leverage that operators can exert in lease negotiations due to forced pooling or market concentration (Vissing, 2017). A second possibility is that it is the result of information asymmetry. There are lots of potential sources of information asymmetry that one could think of in the lease bargaining process, but the simplest (which is also most relevant for the most disadvantaged minority group) is linguistic – i.e., a household will have less ability to bargain effectively if it is not English speaking. Our results suggest that this is indeed a source of the outcome experienced by the Hispanic group in the bargaining process. Asymmetric information presents a clear mechanism for the breakdown of efficient Coasian bargaining.

Moving forward, there are several alternative paths for communities tasked with negotiating leases because of new natural gas discoveries. Access to information about lease negotiation that details terms and potential implications is increasingly available. Usually, these are documents drawn by attorneys that represent property owners, and reading this information decreases the asymmetry in information between property owners and firms, who are likely seasoned negotiators.

More broadly, states could adopt uniform leasing requirements whereby leases are required to have a minimum set of terms deemed fair to households relinquishing the rights to their subsurface minerals for natural gas extraction. Establishing a minimum requirement ensures that all households are protected up to a minimum threshold established by the regulating body that could be based on the experience of communities with currently active natural gas industries.

Uniform leasing is not yet observed in practice. However, stricter regulations overseeing the natural gas industry and their treatment of surface and subsurface estates during and after production can substitute for it. Until recently, local municipalities in the state of Texas were allowed to exercise “home rule,” which meant they could pass stricter rules governing firms operating within their jurisdictions. These rules were more or less strict across cities, with some requiring additional environmental testing, remediation and protection of the landscape, local permitting standards, and restricting noise and the hours of operation. We find, however, that cities with protective ordinances also tend to be cities where residents were able to negotiate their own protective lease clauses. HB40 may therefore have not had much of an impact because local ordinances had largely just duplicated private protections. Further, most oil and natural gas drilling occurs outside municipal jurisdiction. One solution would, therefore, be to establish baseline protections in laws passed at the state level, so that individual landowners are no longer responsible for negotiating their own basic level of protection from potentially harmful violations.

References

Asher, S., D. Nekipelov, P. Novosad, and S. P. Ryan (2016). Classification trees for heterogeneous moment-based models. Technical report, National Bureau of Economic Research.

Bamberger, M. and R. E. Oswald (2012). Impacts of gas drilling on human and animal health. *New solutions: a journal of environmental and occupational health policy 22*(1), 51–77.

Banzhaf, H. S., L. Ma, and C. Timmins (2018). Environmental justice: Establishing causal relationships. *Annual Review of Resource Economics 11.*

Banzhaf, S., L. Ma, and C. Timmins (2019). Environmental justice: The economics of race, place, and pollution. *Journal of Economic Perspectives 33*(1), 185–208.

Biggs, D., d. V. B. and E. Suen (1991). A method of choosing multiway partitions for classification and decision trees. Technical report.

Coase, R. (1960). The problem of social cost. *Journal of law and economics 3*(1), 1–44.

Colborn, T., C. Kwiatkowski, K. Schultz, and M. Bachran (2011). Natural gas operations from a public health perspective. *Human and ecological risk assessment: An International Journal 17*(5), 1039–1056.

Cole, L. W. and S. R. Foster (2001). *From the ground up: Environmental racism and the rise of the environmental justice movement.* NYU Press.

Fambrough, J. (1987). *Hints on Negotiating an Oil and Gas Lease: Technical Report.* Texas A & M University, Real Estate Center.

Fields, J. (2006). Resolving conflicts between surface and mineral development.

Hamilton, J. T. (1995). Testing for environmental racism: prejudice, profits, political power. *Journal of Policy Analysis and Management 14*(1), 107–132.

Hill, E. and L. Ma (2017a). Shale gas development and drinking water quality. *American Economic Review Papers and Proceedings 107*(5), 522–525.

Hill, E. L. (2013). Shale gas development and infant health: evidence from pennsylvania. *Charles H. Dyson School of Applied Economics and Management, Cornell University, Working Paper. Available at*.

Hill, E. L. and L. Ma (2017b). Does shale gas development impact infant health through drinking water? *Working Paper*.

Jenkins, R. R., K. B. Maguire, and C. L. Morgan (2004). Host community compensation and municipal solid waste landfills. *Land Economics 80*(4), 513–528.

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. Technical report.

Maxwell, B. A. (2008). Texas tug of war: A survey of urban drilling and the issues an operator will face. *Tex. J. Oil Gas & Energy L. 4*, 337.

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

Medema, S. G. (1994). *Ronald H. Coase.* Macmillan Press.

Medema, S. G. (2014). The curious treatment of the coase theorem in the environmental economics literature, 1960–1979. *Review of Environmental Economics and Policy*, ret020.

Merrill, R. and F. Merrill (2013). Ownership of mineral rights under texas law. pp. 1–26.

Rahm, D. (2011). Regulating hydraulic fracturing in shale gas plays: The case of texas. *Energy Policy 39*(5), 2974–2981.

Richardson, N., M. Gottlieb, A. Krupnick, and H. Wiseman (2013). *The state of state shale gas regulation.*

Resources for the Future Washington, DC.

Schweizer, U. (1988). Externalities and the coase theorem: Hypothesis or result? *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 245–266.

Vanham, A. and W. Riley (2011). Shift in power: Why increased urban drilling necessitates a change in regulatory authority. *Saint Mary's LawJournal 43*, 229–288.

Vengosh, A., R. B. Jackson, N. Warner, T. H. Darrah, and A. Kondash (2014). A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the united states. *Environmental science & technology 48*(15), 8334–8348.

Vissing, A. (2017). Forced cooperation versus competitive holdouts: A study of efficient land use policies in the oil and natural gas industry. *Working Paper*.

Zelder, M. (1998). The cost of accosting coase: a reconciliatory survey of proofs and disproofs of the coase theorem. In *Coasean Economics Law and Economics and the New Institutional Economics*, pp. 65–94. Springer.