Regulating the Boom: Strategic Behavior in Louisiana's Fracking Industry

Arie Beresteanu* Arifah Hasanbasri†

February 1, 2025

Abstract

We model the strategic interaction between a well owner and the Louisiana Department of Natural Resources (LDNR) as a simultaneous move, incomplete information game. Well owners choose whether to violate environmental regulations, and the LDNR chooses which wells it inspects. We characterize the Bayesian Nash Equilibrium as depending on payoff parameters. Using information on all wells in Louisiana, their operators as well as LDNR's inspections and their results from 2020 to 2023 we estimate the payoffs parameters of this game using a nested fixed point (NFXP) penalized maximum likelihood (PMLE). Our estimated parameters align intuitively with the observed inspection and violation rates and manage to reproduce average rates similar to those in the data. Furthermore, we find heterogeneous externalities when well owners violate but were not inspected. This study highlights the interactions between an enforcement agency and a well owner, providing important implications for regulatory oversight.

Keywords: Hydraulic Fracturing, Regulatory oversight, Bayesian Games estimation, Nested Fixed Point Algorithm, Penalized Maximum Likelihood.

JEL classification codes: C31, C57, Q35, L51

*Department of Economics, University of Pittsburgh, arie@pitt.edu.

†Department of Economics, University of Pittsburgh, ARH205@pitt.edu

1 Introduction

Hydraulic fracturing (fracking) became a widely used technique to extract natural gas from underground rock formations such as shell rock. Fracking involves injecting water, sand, and xchemicals at high pressure to fracture rock formations, allowing oil or gas to be extracted more easily. The combination of horizontal drilling and hydraulic fracturing in the early 2000s revolutionized oil and natural gas extraction (Bartik et al. (2019)). This innovation made reservoirs like shale formations accessible and potentially profitable for (see Gamper-Rabindran (2018)). However, the rapid growth of fracking has raised concerns regarding potential impacts on health outcomes, groundwater and drinking water (Currie et al. (2017); Olmstead et al. (2013); Mason et al. (2015)), prompting calls for increased regulatory oversight (Wynveen (2011)).

Fracking in Louisiana Louisiana has two main natural gas deposits: the Haynesville Shale, located in northwest Louisiana, and the Tuscaloosa Marine Shale, located in central Louisiana. The large-scale application of fracking in Louisiana started around 2007-2008, when companies like Chesapeake Energy and others began developing the Haynesville Shale, one of the richest natural gas formations in the U.S. This period marked the beginning of horizontal drilling and high-volume hydraulic fracturing in the state, leading to a significant boom in natural gas production.[1]

The primary regulatory body overseeing the fracking industry in Louisiana is the Louisiana Department of Natural Resources (LDNR), specifically through its Office of Conservation. The LDNR is responsible for regulating hydraulic fracturing wells throughout their life cycle. This agency approves and grants permits, monitors and can inspect any work done on cite, receives reports about chemical used in the fracturing process, monitors water use and management, waste management, and site deactivation. The LDNR can inspect the wells at any stage, from the moment work begins on the site to after the wells are deactivated and plugged. The LDNR maintains a dedicated website that provides up-to-date information on hydraulic fracturing activities, regulations, and related

[1]For an historical overview of the fracking technology and industry in the US, Europe and Asia see Gamper-Rabindran (2018).

environmental considerations. The LDNR documents all the inspection it conducts, their results and whether fines were levied on the well owners.

Empirical Framework. We model the strategic interaction between a well owner and the LDNR as a simultaneous move, incomplete information game (see Section 2). Well owners can choose to violate or not violate environmental regulations, and the LDNR can choose to inspect. We observe data on violations and inspections as reported by the LDNR and well owner's characteristics. We treat each well as an independent observation. We assume that well owners have private information on the benefits of violating the regulations (the amount of cost saved). The inspector (LDNR) has private information about the benefit of finding a violation in a certain location. The payoffs for the players are modeled as linear combinations of observables and depend on a finite-dimensional parameter. Since each player has private information and this is a simultaneous game, we use Bayesian Nash Equilibrium (BNE) as the equilibrium concept.

This paper uses Nested Fixed Point (NFXP) algorithm within a maximum likelihood estimator (MLE). For every guess of the structural parameters, the fixed-point algorithm finds the BNE for each game (observation) in the sample. Then the likelihood function for this proposed parameter is computed. The maximum likelihood procedure then finds the maximizing parameter. Generally, finding the equilibrium for each game in the sample for each parameter considered can be a daunting task (see Su and Judd (2012) and Bajari et al. (2013) for alternative approaches). we provide conditions similar to the conditions in Aradillas-Lopez (2010) to show that a unique equilibrium exists and, therefore, it is easy to find even for a large number of games. To improve the quality of the structural parameters and to assure convergence we use external moment conditions to the likelihood function as a penalty term. Specifically, we know the overall inspection rates and violation rates in the data. The penalty term is the distance between the model predicted frequencies and the observed frequencies (see Table 3).

Literature Review:

This paper contributes to the broader literature on environmental regulation. Harrington (1988) developed a model in which firms are divided into two groups, with one designated as the target group based on their historical compliance records. Friesen (2003) extended this framework by randomly assigning firms to groups, allowing firms to exit the target group based on observed compliance. These studies find that such enforcement models can minimize inspection costs while achieving the desired compliance rates. Shimshack and Ward (2008) empirically demonstrates that credible enforcement, including fines, can reduce waste discharge below legally permitted levels. Contributing to this literature, we present a case study of Louisiana, where regulators conduct routine inspections and impose fines for violations. This paper examines the strategic interactions between well owners, who weigh the benefits of cost reductions from violations against the risk of penalties, and regulators, who face inspection costs and the potential negative externalities of failing to regulate well owners effectively.

Additionally, this paper contributes to the broader literature on the economic impact of fracking. Studies have shown that shale gas development has various negative effects. Muehlenbachs et al. (2015) found that shale gas development impacts the housing market, with negative effects on the values of groundwater-dependent homes located near wells. Other research has highlighted risks to surface water quality (Olmstead et al., 2013) and potential threats to drinking water under certain conditions (U.S. EPA, 2016).

Despite these negative impacts, fracking also generates positive welfare effects. Studies have found increased welfare for natural gas consumers and producers (Bartik et al. (2019); (Hausman and Kellogg, 2015)) and positive net financial impacts for local governments (Raimi and Newell, 2014). While most of the existing literature examines the impacts of shale gas development, relatively few studies focus on its regulatory aspects. This paper takes the insights from this literature to analyze the strategic behaviors surrounding fracking inspections.

2 Model

Suppose that each well pad,[2] i, is characterized by a discrete characteristic $r \in \{r_1, r_2, \ldots, r_n\}$ and a vector of other characteristics X. In the context of the game we model, r represents the type of the well owner. Specifically we consider a case where $r \in \{0, 1\}$ where $r = 1$ means that the well owner is one of the large operators and $r = 0$ otherwise. The type of the well owner affect the benefit (i.e. cost saving) that the well owner can experience if they choose to violate the regulations. We denote the value of violations as $v(r) = \bar{v}(r) + \xi$ where ξ is the well owner's private information. Consider the following 2×2 Bayesian game played between the monitoring agency (LDNR) and the well owners. The actions available to the LDNR are $A_{LDNR} = \{\text{Inspect, Do not inspect}\}$ and the actions available to the well owner are $A_{Well} = \{\text{Violate, Do not violate}\}$. For each $(a_1, a_2) \in A_{LDNR} \times A_{Well}$ let $U_{LDNR}(a_1, a_2)$ and $U_{Well}(a_1, a_2)$ denote the payoffs for player 1 (the LDNR) and player 2 (well owner), respectively. The following table makes the payoff functions explicit.

LDNR \ Well	violate	Do not violate
Inspect	$(-c, \bar{v}(r) - \xi - j)$	$(-c, 0)$
Do not inspect	$(-x'\beta + \varepsilon, \bar{v}(r) - \xi)$	$(0, 0)$

Table 1: Inspect-violate simultaneous game between the LDNR and well owners.

Equilibrium

We assume that the parameters $\theta = (\beta, j, c, \{\bar{v}(r)\}_{r \in R})$ are known to the players as well as the observables (r, x) but ξ is the private information of the well owner and ε is the private information of the LDNR. We assume that the players have the correct common prior $F_{\xi,\varepsilon}$ for the joint distribution of (ξ, ε). Let $\Xi \times \mathcal{E}$ be the support of (ξ, ε). A pure strategy for the LDNR in the above Bayesian game is a map $t_{LDNR} : \mathcal{E} \to \{Inspect, Do\ not\ inspect\}$. Let $\mathcal{T}_{LDNR}$ be the set of all these maps.

[2] A well pad refers to one location where potentially several wells are drilled in very close proximity. To simplify the discussion we use the term 'well' to mean a 'well pad'.

Similarly, a strategy for a well owner is $t_{Well} : \Xi \to \{Violate, Do\ not\ violate\}$. Let $\mathcal{T}_{Well}$ be the set of all these maps.

Definition 1 *In the above game, a pair of Bayesian Nash Equilibrium (BNE) strategies are two maps* $t^*_{LDNR} \in \mathcal{T}_{LDNR}$ *and* $t^*_{Well} \in \mathcal{T}_{Well}$ *such that*

$$i. \int U_{LDNR}\big(t^*_{LDNR}(\varepsilon), t^*_{Well}(\xi)\big) dF_{\xi,\varepsilon} \geq \int U_{LDNR}\big(t(\varepsilon), t^*_{Well}(\xi)\big) dF_{\xi,\varepsilon}, \ \forall t \in \mathcal{T}_{LDNR} \tag{1}$$

$$ii. \int U_{LDNR}\big(t^*_{LDNR}(\varepsilon), t^*_{Well}(\xi)\big) dF_{\xi,\varepsilon} \geq \int U_{LDNR}\big(t^*_{LDNR}(\varepsilon), t(\xi)\big) dF_{\xi,\varepsilon}, \ \forall t \in \mathcal{T}_{Well} \tag{2}$$

Both players use a threshold crossing strategy.In other words, there are $\bar{\varepsilon} \in \mathcal{E}$ and $\bar{\xi} \in \Xi$ such that for $\varepsilon \leq \bar{\varepsilon}$ the LDNR chooses $Inspect$ and for $\xi \leq \bar{\xi}$ the well owner chooses $Violate$. Both $\bar{\varepsilon}$ and $\bar{\xi}$ depend on the parameter vector θ. As a result of this threshold crossing strategy profile, we can write the probability that the LDNR chooses $Inspect$ and that the well owner chooses $Violate$ in the following way. Let $\phi_\theta = \Pr(Violate|x;\theta)$ and let $\rho_\theta = \Pr(Inspect|x;\theta)$ be the choice probabilities of the players. Then,

$$\phi_\theta = \Pr\left(\xi < \bar{v}(r) - \rho_\theta \cdot j\right) = F_\xi(\bar{v}(r) - \rho_\theta \cdot j) \tag{3}$$

$$\rho_\theta = \Pr\left(\varepsilon < x'\beta' x - \frac{\varepsilon}{\phi_\theta}\right) = F_\varepsilon\left(x'\beta' x - \frac{\varepsilon}{\phi_\theta}\right) \tag{4}$$

The inspected benefit for the inspection agency (LDNR) for inspecting a well pad with characteristics (r, x) is $(x'\beta - \varepsilon)\phi_\theta - c$ where $c > 0$ is the cost of inspection and ε is the well's unobserved characteristic.

Equations (3) and (4) need to be jointly satisfied. Given that both unobservables ξ and ε have contiguous distributions, a solution for the above system of equations exists and is unique.

We assume that ξ_i and ε_i are independent of each other and accross observations ($i.i.d.$ sample) and distributed Logistic with a cummulative distribution function $\Lambda(t) = (\frac{e^t}{1+e^t})$. The equilibrium conditions in (3 and (4) can be written as,

$$\phi_\theta = F_1(\rho_\theta; r, x) - \phi_\theta \tag{5}$$

and

$$\rho_\theta = F_2(\phi_\theta; r, x), \tag{6}$$

where $F_1(\rho_\theta; r, x) = \Lambda(\bar{v}_r - \rho_\theta \cdot j)$ and $F_2(\phi_\theta; r, x) = \Lambda(x'\beta - \frac{c}{\phi_\theta})$. Equations (5) and (6) represent the best response functions of the LDNR and Well owners, respectively. Fix (r, x) and denote by $(\phi^*_\theta, \rho^*_\theta)$ the BNE of the game for the parameter $\theta = (\beta, j, c, \{\bar{v}_r\}_{r\in R})$. Since F_1 is monotonically decreasing in ρ_θ and F_2 is monotonically increasing in ϕ_θ, there is a unique fixed-point solution to the equation system in (5) and (6).

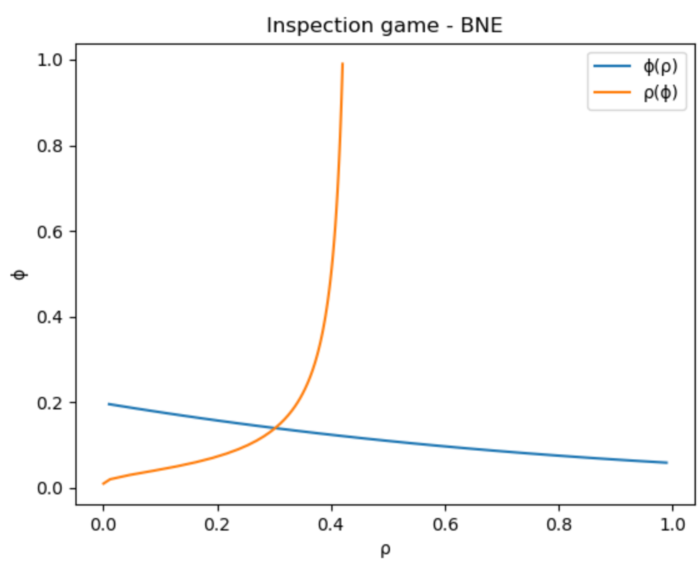

Figure 1: LDNR and Well owners best response functions

Maximum Likelihood

Denote by $y^i_{LDNR} \in \{0,1\}$ the choice by the LDNR to inspect or not inspect well i. Denote by $y^i_{Well} \in \{0,1\}$ the choice by the well owner i to violate or not violate. These choices as well as the covariates (r, x) are observed for each well i.

Let $y^i_{LDNR} = 1$ if well i was inspected $y^i_{LDNR} = 0$ otherwise. Let $y^i_{Well} = 1$ if a violation was detected in well i and $y^i_{Well} = 0$. We cannot observe violations or non-viloations in wells that were not inspected. Therefore, given the parameter

	Violate	**Don't Violate**
Inspect	$\rho^*\phi^*$	$\rho^*(1-\phi^*)$
Don't Inspect	1-ρ^*	

Table 2: Inspection Decision Table

Therefore, the likelihood function is,

$$\begin{aligned} L(\theta) = \prod_i \Big(& [1-\rho^*(r_i, x_i; \theta)]^{(1-y^i_{LDNR})} \\ & \cdot [\rho^*(r_i, x_i; \theta)\phi^*(r_i, x_i; \theta)]^{y^i_{LDNR} \cdot y^i_{Well}} \\ & \cdot [\rho^*(r_i, x_i; \theta)(1-\phi^*(r_i, x_i; \theta))]^{y^i_{LDNR} \cdot (1-y^i_{Well})} \Big) \end{aligned} \tag{7}$$

The log-likelihood is defined by taking a log of the product in (7). In addition to maximizing the likelihood function we use the observed frequencies of inspections and violations. We, therefore, had a penalty term to the log-likelihood function that captures deviations from the observed frequencies. Let $\hat{\phi}$ and $\hat{\rho}$ be the sample frequencies of inspections and violations, respectively. Denote by $\tilde{\phi}_\theta$ and $\tilde{\rho}_\theta$ the model induced frequencies of inspections and violations, respectively, given a parameter θ.

We then find θ that maximizes,

$$
\begin{aligned}
LL(\theta) = \sum_i \Big(& (1 - y^i_{LDNR}) \cdot \log\left[1 - \rho^*(r_i, x_i; \theta)\right] \\
& + y^i_{LDNR} \cdot y^i_{Well} \cdot \log\left[\rho^*(r_i, x_i; \theta)\phi^*(r_i, x_i; \theta)\right] \\
& + y^i_{LDNR} \cdot (1 - y^i_{Well}) \cdot \log\left[\rho^*(r_i, x_i; \theta)(1 - \phi^*(r_i, x_i; \theta))\right] \Big) \\
& + \lambda\big[(\tilde{\phi}_\theta - \bar{\phi})^2 + (\tilde{\rho}_\theta - \bar{\rho})^2\big],
\end{aligned}
\tag{8}
$$

where $\lambda > 0$ is a tuning parameter. We denote $\hat{\theta} = \arg\max LL(\theta)$ to be the estimator of the game payoffs parameters. More details are given in Appendix A.

3 Data

We collected data from the Strategic Online Natural Resources Information System (SONRIS), maintained by Louisiana's Department of Energy and Natural Resources. SONRIS provides public access to comprehensive information on the state's natural resources, including oil and gas records, well logs, production data, GIS maps, and coastal use permits. Our analysis focuses on fracking well pads and their inspection reports. The well pads data includes details such as location, status history, operators, well testing results, and production records. The inspection data covers all routine lease facility inspections, including those for wells that were compliant and non-compliant, along with associated penalty information.

From SONRIS, we extracted data on 6,057 fracking well pads with wells historically opened between 1967 and 2023. Using longitude and latitude coordinates from SONRIS and Census TIGER/Line shapefiles, we mapped the locations of natural gas well pads across Louisiana (Figure 2). The majority of well pads in our data are concentrated in Louisiana's northwest region, which coincide with the region of the Haynesville Shale. The map we compile from the data we downloaded has a similar spatial distribution to the well fracking distribution in FracFocus (2024), maintained by Ground Water Protection Council.

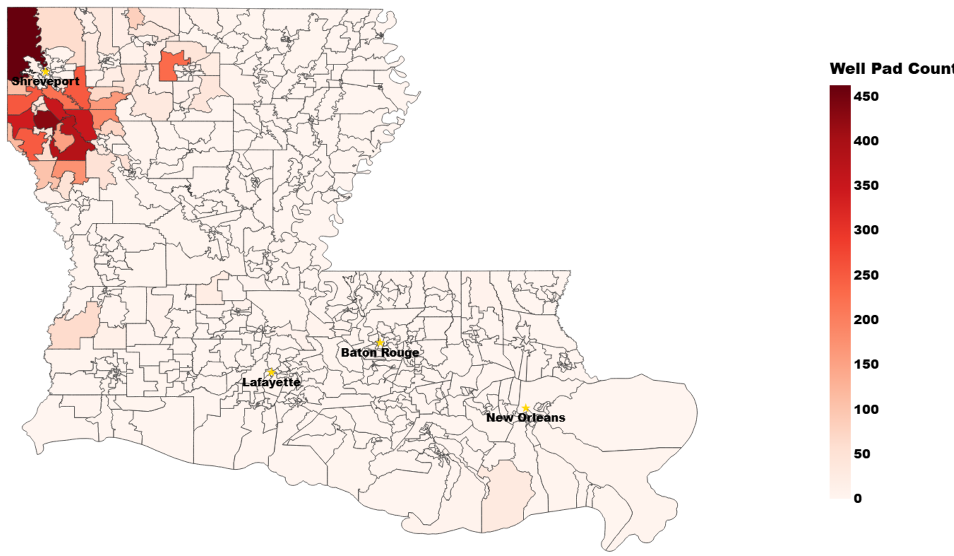

Figure 2: Hydraulic Fracturing Wells in Louisiana

Additionally, well pad data provides information on the fracking operators servicing the wells. Using these data, we computed the market share of fracking operators by ranking them based on the number of well pads they service. The cumulative well pad count is then calculated by rank and market share is determined by dividing the cumulative well pad count by the total number of well pads. Figure 3 illustrates the cumulative market share of well owners in 2020. The top ten biggest fracking operators account for approximately 43% of the well pads, indicating a relatively low level of market concentration.

This cumulative market share that we find in our data aligns with Wang and Krupnick (2015), who report that the top thirty firms owned about 78% of natural gas wells. Moreover, for 2020, the HHI was as low as 282, showing that this market is not concentrated. Market concentration and market shares remained relatively stable from 2000 to 2023. In terms of the composition of the biggest then operators, before 2020 there was a great deal of variation from year to year. From 2020 to 2023, however, the top ten remained relatively stable, with some variation in their rankings.

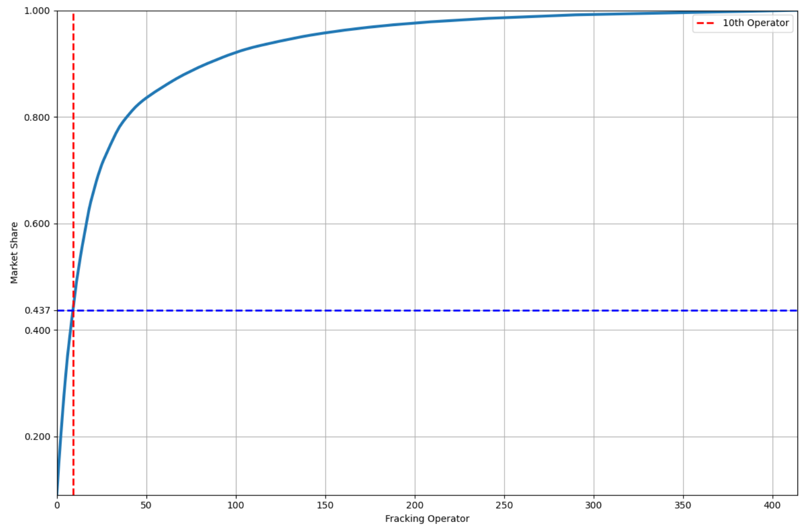

Figure 3: Well Pads Market Share in 2020

For inspection reports, we analyzed routine lease facility inspections conducted between 2020 and 2023, totaling 24,718 inspections over four years. Table 3 provides a detailed breakdown of inspection counts and violation rates, conditional on inspection, by year. Notably, the inspection rate in 2020 was relatively higher than in any other year (29.2%), despite having the lowest violation rate (3.5%). This noticeable difference may be attributed to the impacts of the COVID-19 pandemic.

Year	2020	2021	2022	2023
Inspected	29.21%	24.48%	25.16%	23.28%
Violated	3.56%	9.10%	8.60%	8.23%

Table 3: Inspection and violations annual frequencies

Figure 4 presents the age distribution of active wells from 2020 to 2023, where an active well is defined as one that has not been plugged back. All producing wells are active well, but active wells are not necessarily producing. If wells are not producing, they could be set for production in the future and still labeled as active. Existing wells—those active before the given year—are shown in blue, while new wells that became active that year are highlighted in red. If present, black bars indicate wells operated by companies not included in our fracking operators list. However, the number of wells from non-fracking operators is negligible.

The number of new wells introduced each year during the period of our data is relatively small compared to the total number of wells. New well additions remained above 200 wells until 2021 before declining in subsequent years. Despite these variations, the overall shape of the age distribution remains consistent across years, reflecting a stable pattern of well active and non-active wells.

Across all four years, the age distribution exhibits a right-skewed pattern, with the majority of wells concentrated in the younger age brackets but a noticeable proportion persisting for over a decade. A consistent peak is observed around the 10–12 year age range, suggesting that many wells are producing or remain to produce in the future for an extended period. The distribution extends up to approximately 17 years.

To complement the SONRIS dataset, we merged 2020 population data (United Nations-adjusted) from WorldPop, which provides high-resolution gridded population estimates at a fine spatial scale (e.g., 100m $\times$ 100m). For each well pad, we calculated population density within three zones: a 1-kilometer radius, a 1–5 kilometer radius, and a 5–10 kilometer radius from the well pad's location. Figure 5 visualizes the population distribution, showing that relatively higher densities align with the four largest cities: New Orleans, Baton Rouge, Lafayette, and Shreveport.

In general, the industry of hydraulic fracturing in Louisiana has been stable throughout the years in our sample. Moreover, this industry is not concentrated (HHI=282). In addition, there were no regulatory changes during the period of our changes.

4 Results

Mapping our data to the model, β includes standardized population densities within 1 km, 1–5 km, and 5–10 km radii. For $\bar{v}(r)$, we assume that it follows the form below. We interpret γ as an additional benefit that top fracking operators gain from violating regulations.

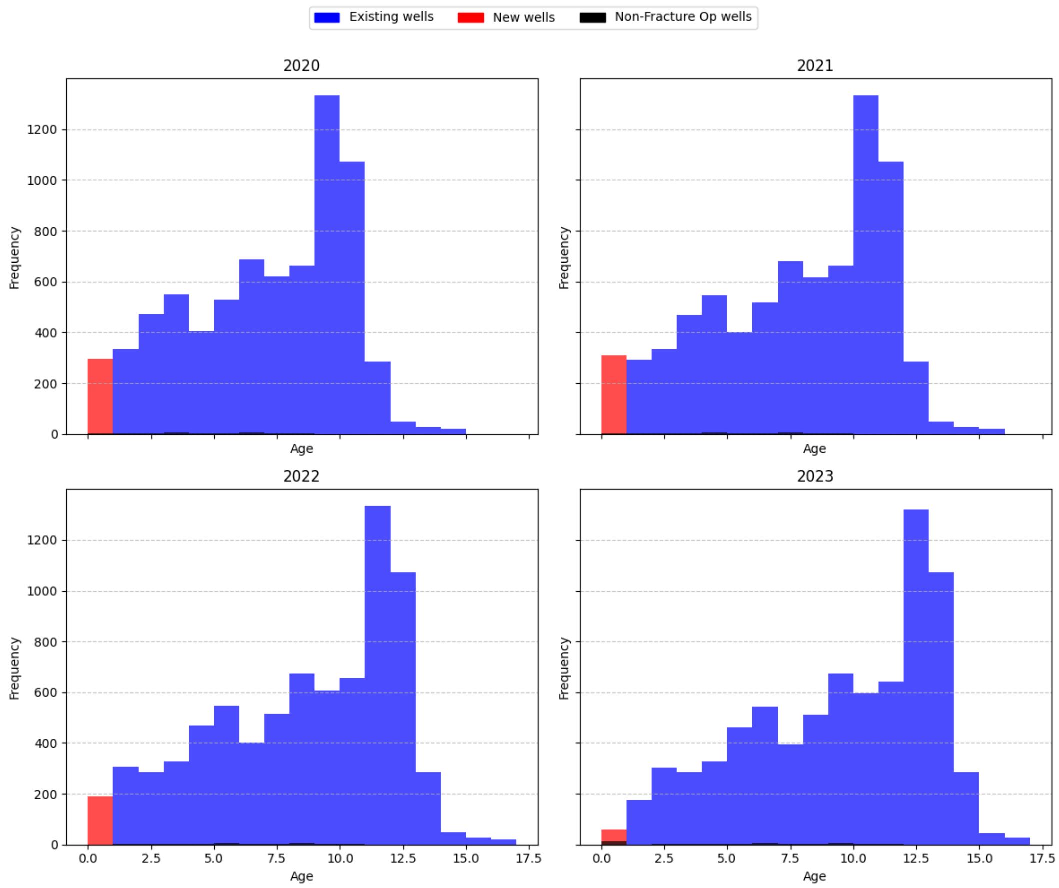

Figure 4: Age Distribution of Wells

$$\bar{v}(r) = \alpha + \gamma \cdot \mathbb{1}(\text{Top 10 Operators}_{y-1}), \quad y \in \{2020, 2021, 2022, 2023\} \tag{9}$$

All estimations in Table. 4 include only well pads where the earliest injection occurred after 2005.[3] The parameters were estimated separately for each year from 2020 to 2023. To derive confidence intervals, we employ an empirical bootstrap method with stratification at the parish level. The bootstrap estimates are then sorted to determine the 2.5^{th} and 97.5^{th} percentiles, which define the confidence interval bounds. For each year, 100 bootstrap samples are used to compute the confidence intervals.

[3] We also estimated the model using well pads that applied for a hydraulic fracturing stimulation permit and obtained similar results.

Table 4: Estimated Parameters by Year

	2020			2021		
Parameters	(1)	(2)	(3)	(1)	(2)	(3)
α	-2.853	-2.852	-2.847	-2.088	-2.089	-2.089
	(-3.07, -2.65)	(-3.08, -2.58)	(-3.39, -2.58)	(-2.3, -1.93)	(-2.34, -1.92)	(-2.26, -1.93)
γ	0.699	0.699	0.699	0.699	0.699	0.699
	(0.7, 0.7)	(0.69, 0.71)	(0.68, 0.72)	(0.7, 0.7)	(0.7, 0.7)	(0.7, 0.7)
j	1.567	1.567	1.571	1.278	1.277	1.273
	(1.51, 1.63)	(1.49, 1.7)	(-0.17, 1.88)	(1.23, 1.33)	(1.23, 1.34)	(1.23, 1.32)
c	0.028	0.028	0.028	0.092	0.092	0.093
	(0.02, 0.03)	(0.02, 0.04)	(0.02, 0.04)	(0.07, 0.11)	(0.07, 0.11)	(0.08, 0.11)
Pop 1 radius km	-0.016	0.065	0.092	-0.096	-0.116	-0.148
	(-0.09, 0.05)	(-0.03, 0.14)	(-0.02, 0.18)	(-0.19, -0.03)	(-0.21, -0.03)	(-0.26, -0.05)
Pop 1-5 radius km		-0.137	-0.553		0.033	0.354
		(-0.24, -0.02)	(-0.8, -0.35)		(-0.08, 0.13)	(0.15, 0.55)
Pop 5-10 radius km			0.41			-0.354
			(0.23, 0.57)			(-0.56, -0.18)
Log likelihood	-3570.0	-3566.0	-3554.0	-3509.0	-3509.0	-3501.0
Penalty	10^8	10^8	10^8	10^8	10^8	10^8
Mean ρ	0.309	0.309	0.309	0.248	0.248	0.248
Mean ϕ	0.034	0.034	0.035	0.083	0.083	0.083
N	5374	5374	5374	5570	5570	5570

(a) 2020 & 2021

	2022			2023		
Parameters	(1)	(2)	(3)	(1)	(2)	(3)
α	-2.122	-2.1	-2.1	-2.116	-2.116	-2.114
	(-2.31, -1.95)	(-2.31, -1.95)	(-2.3, -1.94)	(-2.31, -1.96)	(-2.3, -1.95)	(-2.28, -1.95)
γ	0.699	0.699	0.699	0.699	0.699	0.699
	(0.7, 0.7)	(0.7, 0.7)	(0.67, 0.72)	(0.7, 0.7)	(0.7, 0.7)	(0.7, 0.7)
j	1.306	1.385	1.396	1.3	1.302	1.304
	(1.35, 1.44)	(1.35, 1.44)	(1.36, 1.48)	(1.27, 1.35)	(1.27, 1.35)	(1.25, 1.35)
c	0.084	0.084	0.084	0.098	0.098	0.098
	(0.07, 0.1)	(0.07, 0.1)	(0.07, 0.1)	(0.08, 0.11)	(0.08, 0.12)	(0.08, 0.12)
Pop 1 radius km	0.007	0.019	0.008	-0.037	-0.023	-0.035
	(-0.05, 0.08)	(-0.07, 0.12)	(-0.12, 0.1)	(-0.09, 0.02)	(-0.12, 0.07)	(-0.15, 0.04)
Pop 1-5 radius km		-0.02	0.257		-0.025	0.218
		(-0.1, 0.06)	(0.12, 0.47)		(-0.12, 0.08)	(0.07, 0.41)
Pop 5-10 radius km			-0.314			-0.275
			(-0.5, -0.16)			(-0.44, -0.13)
Log likelihood	-3577.0	-3577.0	-3569.0	-3392.0	-3392.0	-3387.0
Penalty	10^8	10^8	10^8	10^8	10^8	10^8
Mean ρ	0.257	0.257	0.257	0.232	0.232	0.232
Mean ϕ	0.079	0.079	0.079	0.082	0.082	0.082
N	5586	5585	5585	5585	5585	5585

(b) 2022 & 2023

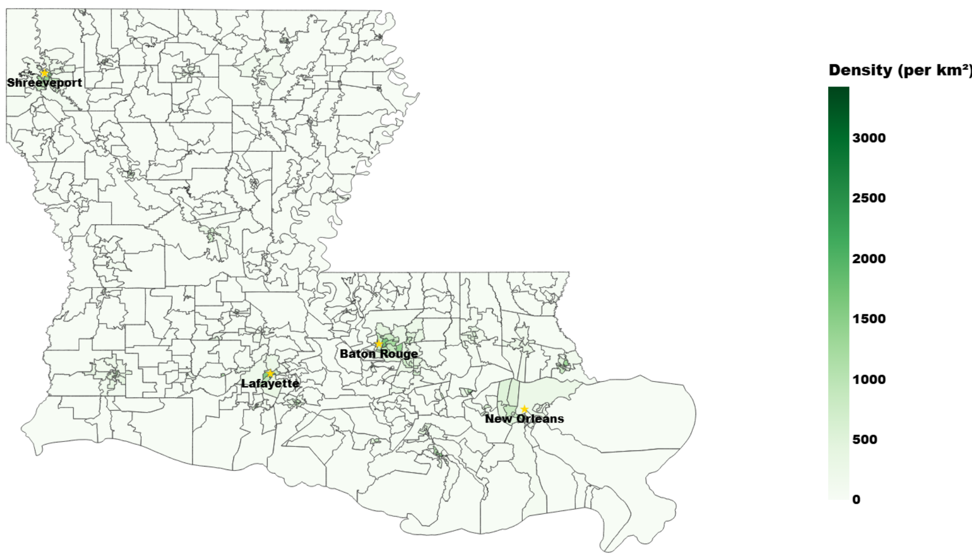

Figure 5: Population Distribution in Louisiana

First, consider the estimated parameters α and γ, which constitute the term $\bar{v}(r)$ in the well owner's payoff. As noted in the data section, the top ten fracking operators remained unchanged from 2020 to 2023. Consequently, γ remained consistent across these years, while slight variations in α were driven by new well pads entering the sample. From 2021 to 2023, the number of new well pads was relatively stable, resulting in a nearly identical α across those years. However, a significant increase in new well pads from 2020 to 2021 led to a noticeably lower estimate for α in 2020.

Second, consider the parameters j, which represents the penalty for violations. Among the four years, j is highest in 2020, although it exhibits greater variance in column (3). Theoretically, this is reasonable given that the percentage of inspected well pads was lowest in 2020. Assuming this is not due to COVID-19-related leniency, it may suggest that the penalty was sufficiently high to incentivize well owners to remain compliant. In subsequent years, j is lower and remains virtually unchanged, which aligns with the similar violation rates observed during these years.

The parameter c, representing the cost of inspection, follows a similar intuition as j. The inspection rate was highest in 2020, resulting in the lowest estimated c among the four years. When the cost of inspection is relatively low, LDNR can conduct more inspections in that year. Consequently, as the inspection rate declined in subsequent years, it is reasonable to expect that the cost of inspection increased.

Regarding the covariates, population density within a 1 km radius exhibits high variance, with most models including zero in the confidence interval. The more relevant covariates are the population densities within the 1–5 km and 5–10 km radii. In most cases, the 1–5 km variable has a positive coefficient, while the 5–10 km variable has a negative coefficient. Interestingly, in 2020, the signs of these two covariates are reversed. The reason for this reversal is unclear, but it could be year-specific issue.

We interpret the coefficient of the covariates to be the externalities that are considered by the regulator. From 2021 to 2023, we interpret the positive coefficient for the 1–5 km radius as indicating that higher population density in this range exacerbates negative externalities when well owners fail to comply. This is intuitive, as residents closer to well pads are more likely to be directly affected by fracking operations. However, it is interesting that the coefficient for the 5–10 km radius is negative. This suggests that higher population density in this range is associated with a positive externality. One possible explanation is that when fracking operators violate regulations, the benefit that the fracking operators gain from cutting costs can contribute to local government or community (Raimi and Newell (2014)), which may benefit communities farther from the well pads. Unlike those within the 1–5 km radius, these communities do not experience the same negative externalities from fracking operations.

Table 4 also presents the average ρ and ϕ, representing the mean probabilities of inspection and violation predicted by the model using the estimated parameters, respectively. We find that these estimates are reasonably close to the actual inspection and violation rates reported in Table 3. To ensure that the mean values of ρ and ϕ remain aligned with the observed rates, we incorporate a penalty term, as described in the penalized MLE in (8). For more details on the algorithm's mechanics, please refer to Appendix A.

Overall, the model provides reasonable estimates. The stability of γ and the observed variations in α highlight the role of industry structure and new well developments in shaping well owners' payoffs. The patterns in j and c suggest that regulatory enforcement and inspection costs fluctuate in response to changes in inspection rates and compliance behavior. Furthermore, the population density reveals heterogeneous externalities, with closer communities bearing the brunt of negative impacts while those farther away may experience indirect benefits through the well owners cutting costs. The reversal of covariate signs in 2020 remains an open question, potentially linked to the COVID-19 pandemic. Taken together, these results contribute to a broader understanding of how regulatory mechanisms influence compliance decisions and externalities in the fracking industry, offering valuable implications for policymakers aiming to balance enforcement efficiency and community welfare.

5 Discussion

We explore the interactions between LDNR and well pad owners in the context of fracking in Louisiana. Our findings demonstrate the relationship between enforcement mechanisms and externalities, having potential implications for policymakers. However, several aspects warrant further consideration. For example, the sensitivity of our results to tuning parameters, using additional covariates of interest in the payoff functions, and data challenges.

In terms of the estimation method, as pointed out by Bajari et al. (2013), the nested fixed-point algorithm is highly sensitive to tuning parameters. The algorithm used to estimate the model's parameters relies on a tuning parameter — specifically, a penalty term—that helps ensure convergence to a fixed point for ρ and ϕ. Without a sufficiently high penalty term, the algorithm may fail to find the fixed points, leading to divergence in the estimated parameters. To address this, we aim to explore alternative methods that are more robust to tuning parameters.

The current model includes only nearby population density as a covariate, but we plan to incorporate additional factors such as proximity to water sources, including rivers and lakes. These variables may influence the probability of inspection or violation due to concerns from local communities and LDNR. Drawing from the fracking literature, we will include more relevant covariates that can capture the externalities that the LDNR face.

Our findings find peculiarity in the 2020 estimates, suggesting that the COVID-19 pandemic may have impacted the inspection and enforcement process for that year. If the relatively lower violation rate was due to regulatory leniency rather than an increase in penalty severity, this could introduce bias into our estimates. The higher inspection rate could be a re-allocation of resources from tasks that require higher interaction with people to tasks that require less interactions, which lease inspections can be done in safe distance. To address this, we plan to collect and analyze 2024 data to determine whether the trends observed from 2021 to 2023 persist.

Our analysis has thus far considered a static model, but potential dynamic effects can be furthered explored. For instance, well pads inspected in one year may be less likely to be inspected the following year. Additionally, findings from the environmental regulation literature suggest that compliance history could influence the probability of inspection. Well owners with a history of violations would face higher frequencies of inspections. Expanding the model to incorporate a historical inspection and compliance records would provide a more comprehensive understanding of regulatory behavior over time.

Furthermore, the model could be extended to account for potential network effects, either through spatial proximity or shared ownership. If one well pad is inspected, nearby well pads or those operated by the same owner may be less likely to violate compared to those outside the network. Incorporating these interdependencies would help capture broader regulatory spillover effects and improve the accuracy of enforcement models.

In summary, this study highlights the interaction between enforcement agency and well owners, providing important implications for regulatory oversight. While the current analysis offers valuable insights, several refinements and extensions would improve its robustness and scope. Future research can provide a more comprehensive understanding of the factors shaping enforcement outcomes.

References

Aradillas-Lopez, A. (2010): "Semiparametric estimation of a simultaneous game with incomplete information," *Journal of Econometrics*, 157, 409–431.

Bajari, P., H. Hong, and D. Nekipelov (2013): "Game Theory and Econometrics: A Survey of Some Recent Research," in *Advances in Economics and Econometrics: Tenth World Congress*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press, Econometric Society Monographs, 3–52.

Bartik, A. W., J. Currie, M. Greenstone, and C. R. Knittel (2019): "The Local Economic and Welfare Consequences of Hydraulic Fracturing," *American Economic Journal: Applied Economics*, 11, 105–155.

Currie, J., M. Greenstone, and K. Meckel (2017): "Hydraulic fracturing and infant health: New evidence from Pennsylvania," *Science Advances*, 3.

FracFocus (2024): "The national hydraulic fracturing chemical disclosure registry." Accessed: 2024-01-29.

Friesen, L. (2003): "Targeting enforcement to improve compliance with environmental regulations," *Journal of Environmental Economics and Management*, 46, 72–85.

Gamper-Rabindran, S., ed. (2018): *The Shale Dilemma, A Global Perspective on Fracking and Shale Development*, University of Pittsburgh Press, Pittsburgh PA.

Harrington, W. (1988): "Enforcement leverage when penalties are restricted," *Journal of Public Economics*, 37, 29–53.

Hausman, C. and R. Kellogg (2015): "Welfare and Distributional Implications of Shale Gas," *Brookings Papers on Economic Activity*, 71–125.

Mason, C. F., L. A. Muehlenbachs, and S. M. Olmstead (2015): "The economics of shale gas development," *Annu. Rev. Resour. Econ.*, 7, 269–289.

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

Olmstead, S. M., L. A. Muehlenbachs, J.-S. Shih, Z. Chu, and A. J. Krupnick (2013): "Shale gas development impacts on surface water quality in Pennsylvania," *Proceedings of the National Academy of Sciences*, 110, 4962–4967.

Raimi, D. and R. G. Newell (2014): "Shale Public Finance: Local Government Revenues and Costs Associated with Oil and Gas Development," Retrieved from Duke University.

Shimshack, J. P. and M. B. Ward (2008): "Enforcement and over-compliance," *Journal of Environmental Economics and Management*, 55, 90–105.

Su, C.-L. and K. L. Judd (2012): "CONSTRAINED OPTIMIZATION APPROACHES TO ESTIMATION OF STRUCTURAL MODELS," *Econometrics*, 80, 2213–2230.

U.S. EPA (2016): "Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report)," Tech. Rep. EPA/600/R-16/236F, U.S. Environmental Protection Agency, Washington, DC.

Wang, Z. and A. Krupnick (2015): "A retrospective review of shale gas development in the United States: What led to the boom?" *Economics of Energy & Environmental Policy*, 4, 5–18.

Wynveen, B. J. (2011): "A thematic analysis of local respondents' perceptions of Barnett Shale energy development," *Journal of Rural Social Sciences*, 26, 2.

A Algorithm

The algorithm used to estimate the parameters follows two main steps.

First, given an initial guess θ_0, we find the fixed points for ρ_{θ_0} and ϕ_{θ_0}. To guarantee that we find the fixed points for $\rho_{\theta_0,k}$ and $\phi_{\theta_0,k}$ for each well pad k, we include a penalty score in the algorithm to prevent $\bar{\rho}_\theta$ and $\bar{\phi}_\theta$, the averages of $\rho_{\theta_0,k}$ and $\phi_{\theta_0,k}$, from deviating too far from $\hat{\rho}$ and $\hat{\phi}$, which represent the probabilities of inspection and violation in the data. The penalty score is given by

$$\lambda\left[(\hat{\rho}-\bar{\rho}_\theta)^2+(\hat{\phi}-\bar{\phi}_\theta)^2\right],$$

where λ is the penalty term that assigns weight to the penalty score. The larger λ is, the more we prioritize keeping $\bar{\rho}_\theta$ and $\bar{\phi}_\theta$ close to $\hat{\rho}$ and $\hat{\phi}$.

Second, we maximize the likelihood in (7) while incorporating the penalty score to obtain a new set of parameters, θ_1, using ρ_{θ_0} and ϕ_{θ_0}. The updated parameters are given by

$$\theta_0 + a(\theta_1-\theta_0),$$

where a is the learning rate, initially set to 1. These two steps are repeated iteratively until the parameters θ_l at iteration l converge within a tolerance of 10^{-5}. The likelihood maximization is performed using the SciPy Python package.

If, during the iteration process, θ_l leads to any of the following: (1) a drastically worsened log-likelihood, (2) a drastically worsened penalty score, or (3) a failure to compute $\rho_{\theta_l,k}$ and $\phi_{\theta_l,k}$, the algorithm reverts to the last best set of parameters, denoted as θ_b. Each occurrence of any of these three cases causes the algorithm to update a by multiplying it by 0.8, i.e.,

$$a \leftarrow 0.8\times\alpha.$$

The rule for determining θ_b is that θ_{b+1} changes by more than a magnitude of 1. After reverting, the iteration continues from θ_b with small perturbations. If the reversion to θ_b occurs three consecutive times, the algorithm terminates after five such occurrences.