THREE ESSAYS IN ENVIRONMENTAL AND RESOURCE ECONOMICS:

HYDRAULIC FRACTURING, WATER USE, AND MARINE RECREATIONAL

FISHING

A Dissertation

by

JESSE DONALD BACKSTROM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Richard T. Woodward
Committee Members,	Masami Fujiwara
	Steven L. Puller
	W. Douglass Shaw
	Reid B. Stevens
Head of Department,	Mark L. Waller

August 2019

Major Subject: Agricultural Economics

ABSTRACT

This dissertation includes three essays in environmental and resource economics. The first two pertain to important topics in relation to the upstream oil and gas industry in the U.S. The third pertains to marine recreational fishing demand in the Gulf of Mexico.

In the first essay, I study two interrelated issues related to the hydraulic fracturing industry's water use in Texas. Using a proprietary dataset of well-level completion reports, I show how firms' propensity to report detailed information on water use varies depending on whether the well is located within a groundwater conservation district— an area where groundwater availability is a concern. Next, I show a causal link between the industry's water use and declining local groundwater levels. These issues are important due to the rapid increase in hydraulic fracturing activity, along with its rising water use per well, which has created many local concerns about the impacts on water availability and the disposal of wastewater produced from these wells.

In the second essay, I use the results of a 2008 U.S. Geological Survey assessment of oil and gas resources as a natural experiment to study various aspects related to oil and gas leasing activity. Specifically, I study how upstream firms use free and publicly available information as an aid to acquire mineral rights in areas with better geology and negotiate more favorable terms on leases. The findings are important as they contribute to the literature on the value of information in the oil and gas industry, and suggest several implications of government-funded resource assessments related to welfare and bargaining in the oil and gas leasing market.

In the third essay, I study marine recreational fishing demand in the Gulf of Mexico. I make an empirical contribution by exploiting a new data set to study how various site characteristics and amenities influence the fishing site choices of private boat anglers. I also make methodological contributions by estimating expected catch for each site using a spatial-temporal technique not previously used in similar studies, and aggregating *less* relevant fishing sites in a new way, helpful to reduce computational burden by nearly three-fourths.

DEDICATION

To my parents, Ron and Kathy; brother, Ryan; and two sisters, Kelly and Lauren.

ACKNOWLEDGEMENTS

I would like to acknowledge and give thanks to my advisor and committee chair, Dr. Woodward, who gave me a chance to pursue a graduate school dream, the freedom to follow my passions and to travel extensively for invaluable research experiences, and who provided continued guidance and support throughout my five years at Texas A&M University. If not for his valuable advice and encouragement, and willingness to talk about any and all research projects and ideas, I would not be in the position I am today.

I am grateful to my other committee members: Drs. Fujiwara, Puller, Shaw, and Stevens for further research advice and support. There are also countless colleagues and other great faculty and staff in the Department of Agricultural Economics, the Department of Economics, and in other departments deserving considerable recognition. I am forever grateful for all of their support and opportunities, lasting relationships and friendships, encouragement, guidance, and collaborations during my time in graduate school.

I am indebted to several institutions who provided the data used in this dissertation, including Primary Vision, who provided the dataset in Section 2 and to Matt Johnson who eagerly provided answers to all of my questions and data inquiries. I am also grateful to DrillingInfo for the dataset used in Section 3, and to Chris Schenk of the U.S. Geological Survey for his helpful explanations of oil and gas resource assessments and their history. I am also extremely grateful to Lauren Dolinger-Few and Sabrina Lovell of NOAA Fisheries, who provided many extensive conversations and explanations of the MRIP data used in Section 4. None of this research would have been possible without their help.

I would like to thank former professors and colleagues at Utah State University and Strata Policy, including Drs. Randy Simmons and Ryan Yonk, who provided my first opportunity to pursue graduate school and prepared me for a Ph.D. program. I would also like to thank Drs. Ihsuan Li and Brad Cook, two professors at Minnesota State University, Mankato, for initially opening my eyes to the world of graduate school and encouraging me to pursue a challenge— one that has turned out to be the best decision I have made.

I would like to thank my mother and father, who have sacrificed and provided the foundation for my successes, and encouraged me every step of the way. My dairy-farming grandparents, Don and Henrietta, have also been instrumental in my path. Whenever I thought graduate school was hard, I always thought of them and remembered that I did not have to wake up at 2:30am to go outside in freezing weather to chisel ice off the barn doors just to get inside to milk the cows. Their influence has prevented entitlement and instilled patience, persistence, and perfectionism into every part of my life.

Lastly, I would like to thank my brother and other close family and friends in Minnesota and elsewhere for all of your trips to Texas. The visits were refreshing, always provided something to look forward to, and made many great memories. I am also a firm believer that the shared suffering and enjoyment throughout graduate school creates some of the purest friendships. I cherish many people that have become so important to me during my time in Texas, and am thankful for several other individuals and life-changing experiences that motivated me to pursue and finish this degree.

I hope to have made you all proud.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Rich Woodward, Douglass Shaw, and Reid Stevens of the Department of Agricultural Economics, Professor Steve Puller of the Department of Economics, and Professor Masami Fujiwara of the Department of Wildlife and Fisheries Sciences at Texas A&M University.

The depiction of regulations, jurisdictions, and other legal information pertaining to groundwater law in Texas in Section 2 and Appendix B was aided by an extensive outline from Jake Cottle, a former master's student at Utah State University. The calculation of total water volumes used in hydraulic fracturing across space in Section 2 was conducted in part by Professor Woodward. The calculation of expected catch rates in Section 4 was also conducted in part by Professor Woodward.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by assistantships from the Department of Agricultural Economics at Texas A&M University, research fellowships from Strata Policy, the Center for Growth and Opportunity at Utah State University, the Institute for Humane Studies, and a dissertation grant from the Charles Koch Foundation. Additional funding for conference travel was also provided by the Graduate Student Association of the Department of Agricultural Economics at Texas A&M University, Strata Policy, the Center for Growth and Opportunity at Utah State University, and the Institute for Humane Studies.

The research in Section 4 was supported in part by Texas Sea Grant under Grant Number NA14OAR4170102. Its contents are solely the responsibility of the author and do not necessarily represent the official views of Texas Sea Grant.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION

The behavior of economic agents and the value of information are central concerns in economics. Studying the behavior of firms, especially those whose operations can have direct impacts on the environment, is important in order to understand how they respond to incentives, and needed to design appropriate policies. Studying the behavior of humans is useful in many contexts, but particularly to understand the primary drivers of choice. Information, both private and public, is also valuable in many contexts, especially in bargaining and other settings where it is used to make strategic decisions. In this dissertation, I empirically study each of these concerns in three essays, each in a different setting, within the context of environmental and resource economics: hydraulic fracturing activity, oil and gas leasing activity, and marine recreational fishing. The findings of each respective essay are valuable to a wide range of stakeholders, including landowners in areas where groundwater availability is a concern, mineral rights holders, fisheries managers, and policy makers.

Activity in the hydraulic fracturing industry has increased rapidly in the U.S. over the last decade. During this time, firms have learned to complete wells that are more productive by increasing the amount of water, among other inputs, used in well stimulations. Since many unconventional oil and gas plays are located in relatively arid regions, this has created concerns over local water availability. However, management of water resources in these areas is complicated, partly due to state laws that largely allow unrestricted groundwater pumping by landowners, but also because the reporting of water use by the industry is not particularly transparent. In Section 2 of this dissertation, I study two interrelated issues on water use in this industry. First, using a unique dataset of completion reports for hydraulically stimulated wells in Texas, I analyze spillovers of local groundwater management regimes on the tendencies of upstream oil and gas producers when reporting water their use. Since freshwater is preferred due to little or no costly purification or treatments needed to make it usable, I argue firms also have preferences for a continued ability to use freshwater, amid concerns over its increased use. Given expectations about future regulations of the use of or access to freshwater, it is plausible that firms have an incentive to leave less of a paper trail when reporting water use. My findings show that producers do in fact provide less publicly available data about wells that are located in a groundwater conservation district, i.e. areas where groundwater availability is a concern. Second, I use a high-frequency dataset from the Texas Water Development Board and show a causal link between water use in hydraulic fracturing and declining local groundwater levels. My primary policy recommendations include expanding the reporting requirements of firms to include total water use per well by both source and type, and incentivizing the use of online water sourcing methods, which enable formal accounting for water transactions.

I study another aspect related to the upstream oil and gas industry in Section 3 of this dissertation. In 2008, the U.S. Geological Survey conducted the first large-scale geology-based assessment of the Williston Basin of North Dakota, Montana, and South Dakota. Although it had long been known that large quantities of unconventional oil and gas resources were present in the basin, the assessment estimated the total amounts in various discretely defined areas within. Given the high costs of private exploration typically needed to locate optimal subsurface geology, I treat the results of the assessment as an exogenous shock of free information and study how upstream oil and gas firms exploited the results as a strategic aid to narrow their search and acquire more economically preferable leased acreage, among other development outcomes. Post announcement of the survey results, in areas of the basin where “more” hydrocarbon resources were estimated I show economically significant changes in the number of lease acquisitions, the average lease size and, importantly, changes in the terms on new leases, relative to the same outcomes in areas of the basin where no assessments were conducted. These findings are important because they provide insight about how government-funded projects are used in industry, which is sought by the U.S. Geological Survey to determine the value of its work. Further, asymmetric information is important in this context because firms who internalized the assessment results quickly were plausibly able to exploit an arbitrage opportunity that may have enabled them to negotiate more favorable terms on lease agreements.

In the Section 4 of this dissertation, I study the behavior of private boat anglers when choosing marine fishing sites in three states along the coast of the Gulf of Mexico, a region whose economy benefits significantly from recreational anglers. In the marine recreational fishing literature, information on site characteristics is largely limited as it is often overlooked or too expensive to obtain. Most studies have incorporated only a small number of site characteristics in their empirical estimation, usually limited to the cost associated with traveling to a fishing site, and some form of expected catch as the only key factors explaining angler site choice. I make several contributions in this study of private-boat anglers who take one-day fishing trips and target Spotted Seatrout, one of the most frequently targeted species in the Gulf. I am the first to estimate a site-level recreation demand model with catch rates using Marine Recreational Information Program data, which is important as previous papers have aggregated sites to county level and were faced with site aggregation bias. I also make an empirical contribution by incorporating an extensive set of amenities and characteristics at fishing sites not previously analyzed, which is helpful to determine the influence of amenities such as the presence of bait and tackle shops, the number of boat ramps, ample parking space, and other characteristics that angler site choice. Lastly, I make two methodological contributions. I estimate expected catch for each site using a spatial-temporal technique not previously used in similar studies. I also estimate the same models using after aggregating *less* relevant sites in a new way, helpful to reduce computational burden by nearly three-fourths.

2. STRATEGIC REPORTING AND THE EFFECTS OF WATER USE IN HYDRAULIC FRACTURING ON LOCAL GROUNDWATER LEVELS IN TEXAS

2.1. Introduction

High-volume hydraulic fracturing, conducted after the drilling of a wellbore, is the primary stimulation technique used to produce oil and gas from low-permeability, unconventional reservoirs. Its use has increased rapidly in the U.S. over the last decade and a half and over this time, an increasingly large volume of water has been used and created concern over the impacts on local availability. In my dataset, median water use for a horizontal well was almost 11.2 million gallons per well in Texas in early 2017, which is up from about 3.8 million in 2012. In the Permian Basin in west Texas, where the largest share of unconventional oil is now being produced,[1] median water use per well in early 2017 was over 14.6 million gallons or the equivalent to supplying about 91,000 average 2-person U.S. households with water for a day, or 250 households with water for a year.[2]

A common complaint from communities located near unconventional oil and gas production is that they felt unprepared to handle the rapid pace of its development, especially the water use of the industry (Freyman 2014; Gold 2014), which U.S. EPA (2016) posits can have large impacts depending on the local balance between withdrawals, availability, and quality. Although informative to understanding the water use of the

---

[1] Source: https://www.eia.gov/petroleum/drilling/#tabs-summary-2.

[2] Calculation based on an average 2-person household consuming 160 gallons of water per day (USGS 2016).

industry, previous studies in the natural sciences literature have only descriptively analyzed the industry's historical water use and its future needs. In economics, qualitative studies of the impacts on local water quality and availability have been helpful to identify potential externalities (e.g. Burnett 2013; Muehlenbachs and Olmstead 2014; Olmstead and Richardson 2014; and Kuwayama et al. 2015). No study in either of these literatures has credibly quantified the impacts of the industry's water use.

In this paper, I am the first to empirically investigate two interrelated issues related to the industry's water use. First, using a unique dataset of completion reports for hydraulically stimulated wells in Texas,[3] I analyze spillovers of local groundwater management regimes on the tendencies of upstream oil and gas producers (operators hereafter) when reporting water their use. I test the hypothesis that completion reports submitted for wells located in a groundwater conservation district (GCD) are *less* detailed, i.e. contain the bare minimum information on water use that is required, relative to those for wells not in a GCD. Since operators prefer freshwater due to little or no costly purification or treatments needed to make it usable, I argue they also have preferences for a continued ability to use freshwater.[4,5] Given expectations about future regulations of the

[3] Dataset provided by Primary Vision in Houston, Texas: http://www.pvmic.com/.
[4] Freshwater is preferred since it is a less expensive (and usually more timely) input than recycled wastewater or other alternatives, even after paying for the disposal of wastewater.
[5] According to Stepan et al. (2010), transportation represents the largest component of water handling costs. In North Dakota, they estimate acquisition and transportation costs of raw water at $0.25-$1.05 per barrel (42 gallons) and $0.63-$5.00 per barrel, respectively, and transportation and deep-well injection costs associated with wastewater at $0.63-$9.00 per barrel and $0.50-$1.75 per barrel, respectively.

use of or access to freshwater, particularly in water-scarce areas (hence, the need for GCDs) operators have an incentive to leave less of a paper trail when reporting water use.[6]

Second, I use a high-frequency dataset from the Texas Water Development Board (TWDB) to test the hypothesis that the amount of water used in hydraulic fracturing is large enough to have an empirically discernable effect on groundwater levels near oil and gas development. Since surface waters are owned by the state of Texas and surface water rights are allocated, this issue is made more salient by the fact that lucrative markets have developed throughout the state where landowners sell their water rights and or pump groundwater and sell it to operators. Larger effects of water use are expected in the more arid regions, such as those where the Permian and Eagle Ford Basins are located, where the majority of hydraulic fracturing activity is occurring and groundwater is the primary source of water for the industry.[7]

To preview the results, I find that operators of wells located in a GCD area are more likely by 1.3-2.8 percentage points to omit key details of their water use. I also find a similar relationship with respect to actual water use, as the propensity to report more than a minimum amount of detail decreases with larger amounts of water used to stimulate a well, and for horizontally and directionally drilled wells relative to their vertically drilled

[6] In 2012, the costs associated with the removal of total suspended solids and total dissolved solids from hydraulic fracturing wastewater were estimated to be $3.00-$6.00 per barrel and $20.00 per barrel, respectively. Source: https://www.waterworld.com/articles/wwi/print/volume-27/issue-2/regional-spotlight-europe/shale-gas-fracking.html.

[7] Nicot et al. (2012) estimate that in the Permian and Eagle Ford Basins, groundwater accounts for 100% and 90%, respectively, of the water used in hydraulic fracturing. They estimate lower portions of groundwater use in the Anadarko (80%) and East Texas (70%) Basins. Nicot et al. (2014) estimate groundwater use at 50% in the Barnet Shale. The other portions come from surface water sources.

counterparts. In the model of groundwater levels, I show a causal link between the volume of water used in hydraulic fracturing and local groundwater levels. I estimate that a 73 million gallon increase (roughly five new 2017 Permian Basin wells) in the monthly amount of water used in hydraulic fracturing within a 10-mile radius of a groundwater monitoring station leads to a drop in the groundwater level by 1.93 feet. I also find heterogeneous impacts for cumulative water use in GCD vs. non-GCD areas, and in the Permian and Eagle Ford regions relative to other areas in the state. These findings are important as they provide the first credible evidence that water use in hydraulic fracturing affects local water availability, and they provide insight on the reporting tendencies of operators in Texas, each of which emphasizes the importance of transparency in reporting for water management.

2.2. Literature Review

A growing literature in economics has studied many of the positive and negative local economic impacts of the "shale boom." The general consensus is that the economic benefits are large, but further research is needed to understand the magnitude and extent of the negative effects (Mason et al. 2015). Bartik et al. (2017) study the local welfare consequences and estimate a willingness to pay to prevent reductions in local amenities, but also a willingness to pay for allowing unconventional oil and gas development, with significant heterogeneity across regions. These results align with the findings of other studies of residents' attitudes and risk perceptions toward hydraulic fracturing, which are predominantly based on prior experience and familiarity with the oil and gas industry, but also how well the impacts are understood by both stakeholders and non-stakeholders (e.g., Schafft et al. 2013; Boudet et al. 2014; and Boudet et al. 2016).

Real estate is one of the primary markets that has been studied in this context. It has seen positive shocks due to industry demand for leased acreage (or access to mineral rights), and negative shocks due to concerns over the potential for domestic water supply contamination and other issues associated with proximity to development (Muehlenbachs et al. 2015; Weber and Hitaj 2015; Weber et al. 2016; and He et al. 2017). The results of other studies investigating the positive impacts of unconventional oil and gas development mostly echo those in previous studies of resources booms. They have found significant increases in local employment rates, wages, and tax and royalty revenues (Feyrer et al. 2017), and large benefits of natural gas consumption for residential, commercial, industrial, and electric power sectors (Hausman and Kellogg 2015). Selected studies of the negative impacts have found that a close proximity to oil and gas development has adverse effects on fetal health (Currie et al. 2017), and that there can be increases and decreases in crime rates (James and Smith 2017; Street 2018). Significant increases in the number of vehicular accidents and significant road deterioration have also been found, due to new populations and heavy truck traffic brought in by hydraulic fracturing and related activities (Muehlenbachs and Krupnick 2013; Rahm et al. 2015; and Muehlenbachs et al. 2017).

Aside from these more 'general' economic impacts, there is an increasingly expansive literature on the localized *environmental* effects, which includes qualitative review papers that are helpful to characterize benefits, environmental risks, and other costs of unconventional oil and gas development (e.g. Fitzgerald 2013; Jackson et al. 2014; Burnett 2015; Krupnick and Gordon 2015; and Mason et al. 2015). Quantitative studies have primarily addressed the effects on: air quality and greenhouse gas emissions (e.g. Howarth et al. 2011; Knittel et al. 2015; and Holladay and LaRiviere 2017); induced seismic activity associated with wastewater disposal (Ellsworth 2013; Walsh and Zoback 2015); agricultural production (e.g. Hitaj et al. 2014; Farah 2017); and the effects of mechanisms aimed at internalize some of the externalities associated with development (Black et al. 2017; Lange and Redlinger 2019). In addition to studying the welfare implications, Hausman and Kellogg (2015) discuss limitations of the current regulatory environment of unconventional natural gas development, and emphasize how more data are needed on the extent and valuation of the environmental impacts.

Outside of studies on the effects on surface water quality (Olmstead et al. 2013), drinking water quality (Hill and Ma 2017), the displacement of water use in agriculture (Hitaj et al. 2017), and the effects of drought on hydraulic fracturing productivity (Stevens and Torell 2018), the only studies on water-related issues have come from a purely descriptive narrative. In particular, they have only generalized about water use trends and availability, estimated total consumptive water use (e.g. life cycle modeling of water withdrawals and use), or forecast future water use in the industry and discussed potential implications for local availability.[8] Empirical studies on the industry's water use have

[8] E.g., Mielke et al. 2010; Nicot 2012; Nicot and Scanlon 2012; Rahm and Riha 2012; Mitchell et al. 2013; Scanlon et al. 2013; Nicot et al. 2014; Rahm and Riha 2014; Scanlon et al. 2014a; Scanlon et al. 2014b; Vengosh et al. 2014; Small et al. 2015; Barth-Naftilan et al. 2015; Kondash and Vengosh 2015; Horner et al. 2016; Scanlon et al. 2016; Scanlon et al. 2017; Kondash et al. 2018; and Lin et al. 2018.

largely escaped the literature, and a causal link has not made between water use in hydraulic fracturing and local water availability.

The primary reason for the lack of empirical literature on this issue is due to data availability on two dimensions. First, high-resolution data on water availability is limited, and Texas is the only state with expansive, frequently collected data on groundwater levels in areas with and without hydraulic fracturing activity. Second, there is relatively poor quality in the data on water use reported by operators, and the level of detail varies significantly across states, mostly attributable to differences in reporting requirements, among other barriers.[9] The combination of horizontal drilling and hydraulic fracturing is one of the most important technological advancements to the oil and gas industry. However, it was not until February 2012 when operators in Texas became required to report total water use, chemical ingredients, and other additives in hydraulic fracturing fluids to *fracfocus.org*.[10,11] Even today, they are still not required to report detailed information on the type nor source of water used in well stimulations, which has led to a nontrivial amount of variability in the level of detail that is reported by operators *within* the state.

Given the typical high concentration of wells associated with unconventional oil and gas development, if many new wells in an area are due to be stimulated and operators

---

[9] The water *type*(*s*) used (i.e. fresh, salt, produced, or recycled water), the exact *source*(*s*) (i.e. surface water bodies, groundwater aquifers, freshwater or produced water storage pits, and municipal wastewater sources), and *location*(*s*) (i.e. grid coordinates) where operators obtain water have long been oilfield mysteries.

[10] The national chemical disclosure registry for the hydraulic fracturing industry, which is managed by the Groundwater Protection Council and Interstate Oil and Gas Commission.

[11] House Bill 3328, Texas Legislature. September 1, 2011.

obtain water from the same or a connected source, there is potential for aquifer drawdown or reductions in water availability to occur more quickly than optimal. This potential is greater in more arid regions, during times of drought, the summer months, and if the water source is a groundwater aquifer with little or no natural recharge. This paper contributes to the literature by showing a causal link between the industry's water use and groundwater availability, and providing further details on the underlying incentive structure for operators when reporting their water use.

2.3. Background

2.3.1. Water Use in Oil and Gas

A Water use in the oil and gas industry has a long history, yet the advent of hydraulic fracturing has made its use a new focus.[12] While the industry's aggregate water use is small when compared to other uses,[13] it can be significant at the local level, sometimes constituting over 50% of total water use— more than the combined use of domestic, agriculture, and other industries —in certain counties in Texas.[14,15,16] Figures 1 and 2 illustrate two trends in drilling activity and water use per well in the Permian Basin. Figure 1 shows a declining number of (less economical) vertical wells; and an increasing

---

[12] Anecdotal concerns over water availability in west Texas: https://www.texasobserver.org/big-spring-vs-big-oil/.
[13] Kondash and Vengosh (2015) estimate hydraulic fracturing accounts for 0.04% of total fresh water use per year in the U.S. Nicot and Scanlon (2012) estimate that its water use in shale gas extraction accounted for <1% of annual total water use in Texas, but in the Barnett shale, it was nearly 9% of the total water use by the city of Dallas.
[14] Source: https://www.scientificamerican.com/article/analysis-fracking-waters-dirty-secret/.
[15] Source: https://insideclimatenews.org/news/15082018/fracking-environmental-impacts-data-water-usage-oil-natural-gas-sand-pollution-study.
[16] Nicot and Scanlon (2012) project net water use by the industry to reach over 100% of total water use (i.e., if water must be transported in) for certain counties, most of which are located in drought-prone regions.

number of horizontal and directionally wells until roughly October 2014, and a slightly declining number thereafter.[17] The decline is likely due to the drop in the world oil price, but also from operators drilling fewer, but longer and more efficient wells, which achieve greater total oil and gas recovery that effectively lowers drilling and completion costs per unit of production.[18,19] Drilling fewer wells, and concentrating them to fewer well pads, is also desirable for local communities as there is less disturbance to the surface (i.e. changes to land use), and the impacts of other disamenities associated with well pad development are less severe.

Figure 2 shows how the median reported volume of water used has decreased over time for vertically drilled wells, but increased for horizontal and directionally drilled wells.[20] There are two reasons for the increasing amount of water used. The first is that operators began drilling longer horizontal wellbores, which systematically increases total water use. The second is that more water was used per horizontal foot of the wellbore, which has an immediate association with oil and gas recovery since more water and associated stimulation pressure create longer fractures that effectively expose more

---

[17] Many early wells in the Permian Basin were drilled vertically and in large clusters to maximize formation exposure.

[18] Source: https://info.drillinginfo.com/permian-basin-production/.

[19] Source: https://newsbase.com/topstories/exxonmobil-takes-lateral-drilling-new-lengths.

[20] Since operators do not know with certainty what the underlying geology of an oil and gas-bearing formation contains, Agerton (2019) posits that an initial vertical 'test' well is often drilled on a new lease to enable operators to learn about the geology below, inform completion designs for sequential wells, and develop more precise expectations about future production. These wells can also be used to extend the life of a lease if its drilling or production deadline is approaching (Herrnstadt et al. 2019). Since they are also typically smaller and use less water (indicating lower drilling and completion costs), these wells can serve a strategic purpose for operators and presumably contribute to the declining amount of water used in vertical wells.

pathway from the producing formation to the wellbore, making the well more productive (Abramov 2016; Bush 2017).

Although the impacts of large water withdrawals over a short period mostly depend on availability and competing water users at a given point, withdrawals still can exacerbate local water scarcity, especially during times of drought such as the one in Texas in 2011 (U.S. EPA 2016). Since a large quantity of water is needed several days to weeks before each well is stimulated (Nicot and Scanlon 2012), a concentration of new wells to be drilled can lead to an abrupt increase in water use in a relatively small area. Data on water use by the industry is therefore crucial in order to understand how its future development may affect local water availability, determine appropriate water management objectives, and aid in the design of socially efficient water policies, especially since the majority of the industry's water use is consumptive.[21]

2.3.2. Texas Oil and Gas Water Markets – Informal and Formal

The regions in Texas with the most hydraulic fracturing activity (west and south) are also areas that experience low rainfall and groundwater recharge, meaning that withdrawals in these areas typically have a larger impact on water availability. Figure 3 provides a map of rainfall in Texas.[22] The majority of hydraulic fracturing activity occurs in the arid parts of Texas,[23] and operators in these areas face two primary water problems.

---

[21] See Appendix B.A.3 for more information on the life cycle of water used in hydraulic fracturing.
[22] A map of groundwater recharge rates, which exhibits similar spatial characteristics to the rainfall map in Figure 3, is available in Estaville and Earl (2008) or at http://texasaquaticscience.org/aquifers-springs-aquatic-science-texas/.
[23] The major unconventional oil and gas plays and where hydraulic development is occurring in Texas can be seen in Figure 5. Figures 3 and 5 are helpful to characterize where water availability may affect water

The first is locating and acquiring, i.e. sourcing, water in a timely manner before hydraulic fracturing stimulation occurs; and the second is disposing, treating, or reusing wastewater that is produced throughout the life of a well (Carr 2017). Due to the institutions governing mineral and groundwater rights in Texas, which treat them as private property (more on this in subsection 2.3.3), informal (and some formal)[24] markets have developed where in addition to leasing their mineral rights, many landowners lease their water rights and or sell groundwater to the industry for use in hydraulic fracturing.

Evidence of the scale of the informal markets is apparent in Hitaj et al. (2017), who show that water use in hydraulic fracturing has displaced some agricultural irrigation water in several states, indicating that water is flowing to higher-valued uses. In Texas, Goldenberg (2013) reports anecdotal estimates from a landowner, who installed a groundwater pump and two storage tanks for use in a new business of selling water to the industry, and said that his well could pump enough to fill 20-30 water trucks for the industry each day. At $60 per truck, a back-of-the-envelope calculation shows this water provision could be worth nearly $40,000 per month in revenue.[25] Further, the landowner mentioned that if he was open to installing more pumps, he could easily increase capacity

use in hydraulic fracturing and vice versa. Appendix B.B describes each of the oil and gas plays in more detail.

[24] Formalism has developed more recently after companies, such as MidstreamH2O and Solar Midstream, recognized a need for more efficient water transportation systems. Adoption of midstream oil and gas transportation methods has occurred, including permanent and temporary water pipelines. Sourcewater.com, an online water marketplace connecting water suppliers and demanders, is another example of innovation occurring these markets.

[25] Similar water markets exist in the Williston Basin in Montana and North Dakota. During the peak of the boom there, many landowners invested roughly $150,000 to build a water depot, from which they pumped groundwater and sold it to operators (Kusnetz 2012). Some earned profits in excess of $25 million in a year supplying the industry water, with several small towns followed suit and earning $10 million in a year.

to fill 100 trucks per day. Due to large rents available to landowners with access to groundwater, a ‘race to pump’ began in areas with increasing hydraulic fracturing activity, creating concerns about depleting water resources too rapidly and imposing external costs[26] on other water users. This common pool resource dilemma appears to be more pronounced in Texas due to its rule of capture law (see subsection 2.3.3.1) and the aridness of the major unconventional oil and gas producing parts of the state. Additionally, in several counties, there are tax write-offs available for aquifer drawdown, meaning that landowners with water rights are able to profit from selling water, which in effect depletes the aquifer below their property, yet they are able to write this off on taxes.[27]

2.3.3. Groundwater Institutions in Texas

2.3.3.1. Common Law: Rule of Capture

Texas groundwater management has been shaped by court cases aimed at protecting private property and legislative efforts to conserve and protect the state’s natural resources. In a landmark decision, Texas adopted the rule of capture for groundwater in the case of *Houston & Texas Central Railroad Co. v. W.A. East* (1904). The railroad company drilled a water well on its property to support its operations, which dried up its neighbor's domestic well. The neighboring landowner sued the railroad company for damages and the case made its way to the Texas Supreme Court in 1904. The

[26] Under common-pool water resource regimes, two externalities are prevalent. Namely, the stock externality, which occurs to due water used today being unavailable tomorrow, and the pumping cost externality, where costs of water extraction increase as the resource is depleted (Provencher and Burt 1993).
[27] Source: https://www.propublica.org/article/irs-tax-loophole-reward-excessive-water-use-drought-stricken-west.

court chose the rule of capture,[28] which grants landowners the right to pump water from beneath their property regardless of the effects on neighboring wells (without malicious intent or intentional waste), over the American Rule, or the rule of reasonable use. Although this decision has subsequently brought many court cases, the rule brings few restrictions with respect to water use, and has proved favorable to the hydraulic fracturing industry.

2.3.3.2. Groundwater Conservation Districts

Many of the Texas Legislature's efforts to conserve water resources have been on the heels of drought. First created in Texas in 1949, GCDs are legal entities charged with providing for the conservation, protection, recharging, and prevention of waste of groundwater resources within their jurisdiction.[29] To manage groundwater, they are empowered with three primary legislatively mandated duties, including the permitting of water wells, developing a comprehensive management plan, and adopting rules to implement the plan.[30] A GCD can be created in one of four ways: (1) action of the legislature, (2) landowner petition, (3) by the Texas Commission on Environmental Quality (TCEQ) on its own motion in a designated Priority Groundwater Management Area (PGMA),[31] and (4) an alternative to creating a new GCD is to add territory to an existing district.[30] Figure 4 provides a map of existing GCD areas in Texas.

---

[28] More information on this decision is available in Appendix B.A.1.
[29] Source: https://texaswater.tamu.edu/groundwater/groundwater-conservation-districts.html.
[30] Source: https://www.tceq.texas.gov/assets/public/permitting/watersupply/groundwater/maps/gcd_text.pdf.
[31] See Appendix B.A.2 for more on PGMAs.

2.3.3.3. Senate Bill 1

The rule of capture allows landowners to freely pump groundwater, but the Texas Legislature has passed laws aimed at encouraging the establishment of more GCDs. Following a three-year drought, the state created its first omnibus water bill in 1997, Senate Bill 1 (SB1), which consolidated all laws governing GCDs into Chapter 36 of the Texas Water Code, and affirmed them as the state's preferred method for groundwater management (Hubert and Bullock 1999). The ruling increased GCDs' statutory power to limit water withdrawals by authorizing them to require a permit for new wells, a statement of purpose in permit applications, users to report metered water use, and to deny out-of-basin transfers. SB1 also authorized the exemption of certain wells from needing a permit, namely those drilled for domestic and livestock use, but also rig supply wells.[32] Since water use in hydraulic fracturing is not directly regulated, the ruling has become heavily debated, and questions have arisen over whether it exempts water used for high-volume well stimulations, or if it only intended to exempt water used for drilling and smaller rig needs.[33]

---

[32] In Chapter 36 Section 117(b)(2) of the Texas Water Code: "A district (GCD) shall provide an exemption from the district requirement to obtain a permit for drilling a water well used *solely* to supply water for a rig that is actively engaged in drilling or exploration operations for an oil or gas well permitted by the Railroad Commission of Texas provided that the person holding the permit is responsible for drilling and operating the water well and *the water well is located on the same lease or field associated with the drilling rig*." This statutory exemption of oil and gas rig water supply wells was originally passed in Texas in 1971.

[33] Mentioned in phone conversation with attorney Jim Bradbury (https://www.bradburycounsel.com/): interpretation of this exemption is different across GCDs, where its lack of clarity allows opposing views to "maneuver around it." However, since there is disagreement between GCDs on whether water used in well stimulation is associated with drilling, exploration, or production, no GCD has wanted to be the 'test case' and exclude this type of water use from the statutory exemption. It was mentioned that, in instances of dispute, settlements are typically made and no cases ever reach the courtroom.

2.3.3.4. Senate Bill 2 and A (Water) Data Problem

Although treating mineral and groundwater rights as private property is both industry and (mostly) landowner friendly, many parts of Texas were underprepared for the new and collectively increasing water use of operators. Senate Bill 2 (SB2) was passed by the legislature in 2001 and it updated and strengthened the initiatives in SB1. The bill reduced some permitting powers of GCD by prohibiting them from denying a permit solely on the basis that the user planned to export groundwater out of the district; instead, it authorized them to place an export fee on such transfers (Hardberger 2016). SB2 also expanded GCDs' permitting and enforcement powers by authorizing them to regulate water well spacing to minimize interference between wells and set production limits based on tract size or pumping capacity.[34] Reporting requirements however, are still limited and water use is often reported infrequently in most GCDs.[35]

2.3.3.5. House Bill 3328

In response to the industry's increasing water use and public concerns over the chemicals used in hydraulic fracturing fluids contaminating groundwater, House Bill 3328 (2011) changed the future of the industry's reporting in the state. It directed the Texas Railroad Commission (TRC) to adopt rules requiring the disclosure of the fluids and additives used in hydraulic well stimulations. The new disclosure rules required operators

---

[34] The pumping capacity of a well is usually based on a pumping rate such as gallons per minute or acre-feet per acre, and is a primary determinant of the need for a permit, and owners of larger wells can be subject to user and export fees by GCDs (Lesikar et al. 2002). To my knowledge, there are no physical pumping limits directly imposed on these water users, but the permitting process has been one way of limiting pumping in GCD areas.

[35] Most well users are only required to report an annual volume, according to phone conversations with Jim Bradbury.

of wells, for which drilling permits were issued by the TRC on or after February 1, 2012, to submit a completion report for each well to FracFocus that included, at a minimum, the information in Table 1 (Cavender 2011). Important to this study are requirements (8), (9), and (11), which detail the information that pertains to water and chemical ingredients in hydraulic fracturing fluids. Although a total water volume must be reported per requirement (8), operators are not required to provide any detail on the water type used nor its source. I use this requirement, or lack thereof, to create two metrics for the level of reporting exhibited by operators (see subsection 2.5.1).[36] Requirement (8) also made water use data available for all hydraulically fractured wells in Texas, which I use to study the impact of the industry's water use on local availability.

2.3.3.6. Conceptual Framework and Intuition for Reporting Water Use

Taking into account the ambiguity of the 'total volume of water' reporting requirement, a proposition I make is that if operators have preferences to continue using freshwater as a primary water source, then strategic tendencies should exist when reporting water use to FracFocus. Strategic reporting in response to mandatory or non-mandatory disclosure requirements is not a new phenomenon. Lyon and Maxwell (2004 and 2008) offer a theoretical framework to explain industry behavior in response to disclosure regulations, mostly out of an effort to appear "green." They posit how firms may respond to such interventions in various ways, including some that may be costly in

---

[36] The reporting of chemicals used in hydraulic fracturing has been studied previously by Fetter (2017) and Fetter et al. (2017).

the short run but are strategically beneficial in the long run, particularly if they forestall other regulations that are expected to be costlier.[37]

In the context of reporting in hydraulic fracturing, Fetter (2017) studies how state-level disclosure regulations affect operators' use of toxic additives in hydraulic fracturing fluids, and provides evidence of behavioral changes in oil and gas development activity and reporting to FracFocus. Other studies have also found episodes of strategic behavior by operators in the oil and gas industry, including experienced firms exhibiting less deterrence from regulations (Maniloff 2019) and general avoidance behavior towards obligations for the environmental remediation of oil and gas wells (Muehlenbachs 2012). However, the impacts of such regulations can be disproportional. In a study of the effects of changes in environmental regulations on oil and gas development in North Dakota and Montana, Lange and Redlinger (2019) show compositional shifts occur within the industry and that smaller operators are more burdened by such regulations and frequently exit the industry.

Given plausible operator concerns over water availability in Texas, possible knowledge on the effects of their water use, expectations over future regulations that may

[37] Other examples of firm responses to disclosure laws: Chatterji and Toffel (2010) show that an organization's responses to institutional pressure are heavily influenced by both the marginal cost and perceived benefits of responding. They find that, among firms receiving poor environmental ratings, those in environmentally sensitive industries are especially likely to improve performance, given their heightened scrutiny and potential to be inspected. Doshi et al. (2013) study an environmentally important disclosure program, the Toxic Release Inventory and later expansions requiring establishments to report waste, transfers, and releases of certain toxic chemicals under the U.S. Emergency Planning and Community Right-to-Know Act of 1986, and find heterogeneous responses by firms mandated to disclose certain toxic chemicals. Such programs are also common in relation to carbon emissions and climate change. For example, Matisoff (2013) studies the effectiveness of state-based mandatory carbon reporting programs and the voluntary Carbon Disclosure Project, and finds that *how* information is reported to stakeholders has important implications for program effectiveness.

increase the costs to access and or acquire water, and a concern for disclosing too much information on hydraulic fracturing fluids to the public or competitors (e.g., Fetter et al. 2017), it is clear that incentives exist for operators when reporting water use. My focus is to distinguish how the reporting of information on water use varies for oil and gas wells located in areas where water availability is of greater concern compared to where it is less of a focus (i.e., in GCD and non-GCD areas of Texas). I hypothesize that strategies come in the form of operators reporting less detailed information on their water use.[38]

2.4. Data

I make use of datasets from several sources to conduct the analyses in this paper. The first is a set of completion reports for hydraulically fractured wells. The second is a high-resolution dataset on groundwater levels. I also obtained several weather-related datasets, along with data on large water users, and shape files for counties and GCDs in Texas. I describe each dataset in the following subsections.

2.4.1. Primary Vision

I obtained a dataset from Primary Vision (PV), a company based in Houston, Texas, which contains a unique set of well-level records and completion report

[38] For example, operators might report an ambiguous total "water" volume, making the actual water type used unascertainable in order to strategically limit transparency, or even limit the potential for interaction with local institutions that regulate water use or access to water, plausibly in an attempt to delay or prevent future regulations on accessing freshwater. Further, since FracFocus data are publicly available, the platform also acts as a source to learn about water, sand, and chemical use in hydraulic fracturing fluids, and operators may wish to keep their fluid mix as proprietary in order to limit other operators from gaining knowledge about hydraulic fracturing fluids. Although water type and source are not required on completion reports submitted to FracFocus, it makes rational economic sense for an operator that used freshwater to be less transparent in its reporting if it prefers a continued ability to use freshwater without regulations. This is in contrast to an operator that predominantly uses alternative water types, such as saltwater or recycled wastewater, who intuitively would have preferences to make public its smaller freshwater footprint.

information on the use and composition of water in hydraulic fracturing fluids. It included other data as well, such as the operator, start and end dates of the completion, drilling orientation (i.e. horizontal or directional vs. vertical), the reported total volume of water (in gallons) used in well stimulation. It also included an indicator for whether the well record associates with a new completion or a re-fracture of an existing well (commonly known as a 'refrac'), and a unique hydraulic fracturing fluid mass (HFFM) variable calculated by PV. The complete dataset included records for nearly 124,000 hydraulically fractured wells in several states over 2011-May 2017, and was constructed by combining the FracFocus database with data from other public sources such as the TRC. Figure 5 shows the spatial extent of the 59,578 wells completed in Texas over this time.[39] In the analysis of groundwater levels, I included all of these wells, but for the analysis of reporting, I dropped well records prior to February 1, 2012[40,41] since reporting was not mandatory and including these wells would bias my results due to inherent differences in characteristics between operators that voluntarily reported and those that did not.[42]

2.4.1.1. Other Controls and Data Manipulations

Using the well records in this dataset, I also created several additional variables including the cumulative number of wells completed in each county-month by *all* operators, by the *largest five* operators,[43] and by *each* operator. I then created variables for

---

[39] For detailed information on each major hydraulic fracturing region in Texas, see Appendix B.B.
[40] This is in contract to Fetter (2017), who included wells prior to reporting mandates in each state.
[41] Dropping these observations reduced my sample size to 53,182 wells.
[42] Only a limited number of reports were available Texas in 2011.
[43] I define size by the total number of wells each operator has in Texas, and the largest five operators have ~20% of the wells in my sample.

the number of wells completed in *each* county-month by the *five largest* operators and by *each* operator, as well as for the total number of wells completed in each county by *each* operator. Collectively, these variables help me to control for learning about reporting over time, and potential peer effects since smaller operators may observe and attempt to learn from larger operators, such as by locating in the same areas and adopting similar development practices. Since it is common for operators to drill a preliminary test well in a new location to learn about geology (Agerton 2019), I also create indicators for whether the well was an operator's first (in my sample) and whether it was an operator's first in a given county. These controls are important because if a well was drilled for testing or exploratory purposes, or simply to extend a lease (or hold it by commencing drilling activity or production), then an operator might be less careful when reporting water use since these wells are typically smaller and use less water.

2.4.2. Groundwater Monitoring Stations

Texas has one of the most comprehensive statewide groundwater databases in the U.S.[44] I obtained an unbalanced panel dataset from the TWDB on groundwater levels at 273 groundwater-monitoring stations located throughout the state (Figure 6). The panel is unbalanced due to new stations being built and others being shut down, both occurring at different points in time. These monitoring stations record the distance from the surface to the groundwater level at a daily frequency. As groundwater withdrawals occur within the vicinity of the monitoring station, the distance from the surface to the groundwater level

---

[44] Source: http://www.twdb.texas.gov/groundwater/data/index.asp.

will increase, and vice versa when the aquifer recharges. I used the daily observations to create a monthly average distance to the groundwater level below each monitoring station, which I use as my outcome variable in the analysis of groundwater levels. The dataset also contains information on the county, aquifer, whether the aquifer is confined or unconfined,[45] and whether the monitoring station is located within a GCD.

2.4.3. GCD Indicator

I estimate differences in the level of reporting of water use using an indicator for the location of a well within a GCD or non-GCD area. This variable was created in ArcMap using the grid coordinates reported for each oil and gas well and overlaying them with shape files of Texas counties and GCDs, which come from the Texas Department of Transportation[46] and the TWDB,[47] respectively. Four of the existing GCDs were established during my sample period. However, only three of these four areas had a hydraulically fractured well in the area of the GCD either before or after it was established,[48] and only 386 wells total were completed in these areas *after* the establishment of the respective GCDs, providing only a small amount of 'post' variation. Although it would be desirable to have more variation in GCD establishment date, my

[45] Surface water may seep into an unconfined aquifer and recharge it, whereas for a confined aquifer there is an impermeable layer of dirt or rock located above it that prevents such seepage from occurring.
[46] Source: http://gis-txdot.opendata.arcgis.com/datasets/8b902883539a416780440ef009b3f80f_0.
[47] Source: http://www.twdb.texas.gov/mapping/gisdata.asp.
[48] Of the four GCDs established during my sample period, Reeves County GCD had the most wells with 1,655, followed by Terrell County GCD with 14, Calhoun County GCD with 3, and Comal County GCD had 0. These totals however, shrink after reducing the sample to only include operators with wells in both GCD and non-GCD areas.

sample is limited due to most GCDs having been established before the reporting of water use was required.

Table 2 provides detailed information on the number of wells completed in Texas since 2012 and the reported water volumes used, both by drill orientation and across GCD status. Given that more wells were completed in GCD areas than non-GCD areas in each year of the sample, it is clear that the majority of hydraulic fracturing activity is occurring in more arid parts of the state, where the geology is more favorable, but groundwater availability is of greater concern.[49] Median reported water volumes used in horizontal and directionally drilled wells located in GCD areas were also higher than those in non-GCD areas for 2012, 2013, and 2015, but were otherwise higher in non-GCD areas, potentially indicative of more water availability, easier access to water, or both.

2.4.4. Weather Data

I obtained data from the U.S. Drought Monitor, and created an average monthly index of five levels of drought in each county.[50] I merged these variables with both the oil and gas well and groundwater monitoring station datasets. In the former, I use them to control for potential changes in reporting during times of low water availability when operators may have more of an incentive not to disclose freshwater use. An additional explanation for this is that a large lag exists between the time a well is completed and

[49] One exception is the Permian Basin, where more wells were drilled in non-GCD areas each year, although the entire basin is located in a region with little annual rainfall and low groundwater recharge (see Appendix B.B.1).
[50] Source: https://droughtmonitor.unl.edu/Data/DataDownload/ComprehensiveStatistics.aspx.

when the completion report was submitted.[51] Operators may observe drought (or water availability in general) during this time and adjust reporting accordingly. I also obtained additional weather data from NOAA's National Centers for Environmental Information.[52] I used these data to create average monthly rain, temperature, and wind levels for each county. These data were also collected from monitoring stations located throughout the state of Texas, although temperature and wind were recorded by fewer monitoring stations, which limits my sample since data for all counties do not exist.

2.4.5. Other Data

Since hydraulic fracturing is not the only use that may affect groundwater levels, I also obtained data on other water users. First, I obtained data on annual irrigated acreage for corn, cotton, sorghum, and wheat in each county in Texas, and the annual total acreage planted for each of those four crops plus rice, all of which come from the U.S. Department of Agriculture's Quick Stats database.[53] Controlling for these terms is important as irrigation is the largest water user in Texas, and these five crops are known to be higher on the water-use intensity distribution of irrigated crops. To control for municipal water use over time, I also obtained annual population data for each county in Texas from the U.S. Census Bureau.[54]

---

[51] The average time to submit a completion report to FracFocus is 79 days following the completion date on the record. Source: https://fracfocus.org/node.
[52] Source: https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GHCND.
[53] Source: https://quickstats.nass.usda.gov/#E0F9B6F0-3138-3F90-B8F8-8F27943CB593.
[54] Source: https://www.census.gov/data/tables/2017/demo/popest/counties-total.html.

2.5. Empirical Strategy

This paper examines the reporting tendencies of operators when detailing information on water use in completion reports submitted to FracFocus, and the causal effects of the industry's water use on local availability. The following subsections outline my empirical approach. First, I use a linear probability model of reporting aimed at testing the conceptual framework where firms respond to disclosure regulations in strategic ways. Second, I use a fixed effects strategy tied to hydrogeology to model changes in groundwater levels. The models complement each other well since the majority of hydraulic fracturing activity in Texas occurs in more arid regions— areas that are more susceptible to impacts from large water withdrawals. Hence the reason for GCDs and incentives to forestall future regulation of access to freshwater.

2.5.1. Reporting Water Use

2.5.1.1. Metrics for Reporting Water Use

I created two metrics for the level of reporting exhibited by an operator in a completion report. First, I made use of a variable created by PV that estimates the total HFFM used in well stimulation. This variable is important because it directly relates to the amount of freshwater used in each well, attributable to the total volume of water by type(s), but also the density associated with each water type, frac sand (or proppant), and chemical additives used. Each operator has private beliefs on what this mix should consist of, but for each well record that contained voluntarily-reported information on total water volume by type(s) PV used this along with other information on the FracFocus completion report and other proprietary information, and applied a density estimate for each water type used in order to estimate a total HFFM. For well records with insufficient information available for PV to estimate a HFFM (i.e., when only an ambiguous total 'water' volume was reported, or the other reported information was inadequate), the HFFM was coded as unknown as PV was unwilling to make assumptions about the water type(s) used or other fluid characteristics.[55] An unknown HFFM is important because it indicates than an operator only reported the baseline amount of information on fracturing fluids, instead of voluntarily disclosing more information. I use this variable to create a binary outcome for each well as follows:

$$Y_{1,i} = \begin{cases} 1 & if\ hydraulic\ fracturing\ fluid\ mass\ estimated \\ 0 & if\ hydraulic\ fracturing\ fluid\ mass\ unknown \end{cases}.$$

Roughly, 95% of the well records in my final dataset had enough information for PV to estimate a HFFM. However, the actual proportion of well records detailing any information on water use by type is lower than this. Using an alternate metric for the level of reporting, equal to one if *any* information on water type was reported and zero otherwise, about 79% of well records contained at least some information on water type beyond a generic total water volume:

$$Y_{2,i} = \begin{cases} 1 & if\ any\ information\ on\ water\ type\ was\ reported \\ 0 & if\ no\ information\ on\ water\ type\ was\ reported \end{cases}.$$

Figure 7 provides information on the reporting of water use over time in GCD and non-GCD areas, and for the two metrics of reporting. In both cases, reporting has improved

[55] Since water and sand make up the largest components of the fluid mass, erroneous assumptions about water types used can lead to large over or underestimates of the true fluid mass. PV verified that when a HFFM was coded as unknown, it meant that they were unable to gather additional information on water types used or other information about the well from other industry sources that was needed in its calculation.

over time, potentially attributable to learning or a more mechanized and increasingly manufacturing-style approach to drilling and completions. In the first image, PV estimated a HFFM for a larger proportion of well records over time, and in the second image, an increasing proportion of well records contained information on water use by type. However, there is still a clear difference in the level of reporting across GCD status, which persists over the duration of the sample and provides preliminary evidence of strategic reporting.[56]

2.5.1.2. Model Choices

Using each binary indicator for the level of reporting for well record *i*, of operator *j* during month *t*, I first estimate linear probability models to obtain more easily interpretable results. I then re-estimate each model using a logit specification as a robustness check for the functional form assumptions associated with the linear probability model. My estimating equation is as follows:

$$Y_{i,j,t} = \delta \cdot GCD_{i,j,t} + \boldsymbol{x\beta} + \gamma_j + \lambda_t + \varepsilon_{i,j,t}, \tag{2.1}$$

where $GCD_{i,j,t}$ is a treatment indicator equal to one if well *i* of operator *j* is located in a GCD during the month of completion and zero otherwise, and $\boldsymbol{x}$ is a vector of controls that includes well characteristics and other local influences, which are important to include as some wells are fundamentally different in ways that affect reporting.[57] I also include

[56] Although my data suggest only a minimal amount of recycled wastewater is being reused, reporting the use of freshwater alternatives can be beneficial to operators in order to give them, and the hydraulic fracturing industry in general, a “greener” public image. However, if it is actually being used, I am not seeing it being reporting by operators in Texas.

[57] Note: well characteristics are at the well level, but the other local influences are at the county level during the month of completion of the respective well.

operator and month of sample fixed effects, γ_j and λ_t, respectively, to control for average differences in reporting between operators and time-specific confounders common to operators of wells in all areas. The coefficient of interest is δ, which estimates the difference in the probability of a well record containing detailed information on water use based on whether it is located in a GCD or a non-GCD area.

2.5.2. The Causal Effects of Hydraulic Fracturing on Local Groundwater Levels

All unconventional oil and gas development in Texas uses groundwater, but the Permian and Eagle Ford Basins use it in the largest proportions at 100% and 90%, respectively (Nicot et al. 2012).[58] The law of conservation of mass dictates that groundwater levels respond to withdrawals, irrespective of their magnitude. My goal is to show that water use in hydraulic fracturing is large enough to have an empirically discernable effect on a local scale.

2.5.2.1. Naïve Approach

As U.S EPA (2016c) notes, the impacts of groundwater use depend on the withdrawal amounts and water availability at a given point. To estimate the impact of hydraulic fracturing on groundwater levels, I implement a fixed effects strategy similar to Smith (2018) who studies the impacts of groundwater withdrawals across space. I use the PV dataset to create a set of variables for dynamically changing total water use in oil and gas wells within the vicinity of groundwater monitoring stations (see Figure 8 for a visual characterization). My preliminary estimating equation is:

---

[58] See Appendix B.B for more details on the industry's use of groundwater across the state.

$$GWL_{i,t} = \delta \cdot TWV10_{i,t} + \sum_{j=1}^{8} \alpha_k \cdot TWVring_{k,i,t} + \boldsymbol{x\beta} + \lambda_t + \gamma_i + \varepsilon_{i,t}. \quad (2.2)$$

The outcome, $GWL_{i,t}$, is the groundwater level below monitoring station i at time t. The total water volume term, $TWV10_{i,t}$, represents the total water volume used in oil and gas wells within 10 miles of monitoring station *i* in month t. The k^{th} interval, $TWVring_{k,i,t}$ is the total water volume used in oil and gas wells beyond the initial 10-mile radius, but within 10-15, 15-20 … and 45-50 mile annuli (or rings) around groundwater monitoring station i in month t. The set of controls, $\boldsymbol{x}$, includes variables for weather and other water users in the county of monitoring station *i* over time.[59] Lastly, λ_t is a set of month-by-year fixed effects included to absorb time-specific confounders common across all stations, and γ_i is a set of monitoring station fixed effects, included to account for differences in average groundwater levels across monitoring stations. Identification requires that the counterfactual trajectory of groundwater levels in regions with shale absent of hydraulic fracturing would have followed a trajectory similar to the groundwater levels in regions that do not have shale.

Figure 8 provides a visual of the spatial aspect. The blue dot in the middle circle is a groundwater monitoring station (purple dots are oil and gas wells). Although estimating the spatial dispersion of the effects of water withdrawals is not of first priority for this analysis, I use this approach in a baseline model to show how the magnitude of the effect of water withdrawals on groundwater levels should diminish across space. In

[59] Note: all weather variables are at a monthly frequency, but total irrigated acreage, total acreage planted, and county population are all at an annual frequency.

other words, the cone of depression associated with each groundwater withdrawal should have less of an effect on the groundwater level read by the monitoring station the farther away it occurs.

There are a few potential concerns here. First, it is unknown whether the water used in each well truly comes from a surface, ground, or another source. However, given previous studies of the primary water sources used to supply the industry in Texas, it is reasonable to assume that a large portion comes from groundwater. Especially in the arid west and south Texas, groundwater is the largest source of water for all users in these regions. Second, there is measurement area associated with the location of where the withdrawal of water used to supply each well actually occurred. For each radius and ring, I sum water volumes used in oil and gas wells in these areas assuming that the water came from some area around that well, but within the vicinity of the respective radius or ring of the monitoring station. This assumption is reasonable, but there is clearly measurement error introduced with this approach, since water supplies for a well can come from nearby sources or others located farther away.[60] Lastly, and of most concern, the use of contemporaneous total water volumes as the treatment variable may result in a reverse causality problem, and in the next subsection I outline my strategy to circumvent this issue.

---

[60] One company trucked 3.5 million gallons of water from 50 miles away to a drilling site, paying about $68,000, and a fraction of the $3.5 million cost to complete the well. Source: https://fuelfix.com/blog/2011/10/06/parched-texans-impose-water-use-limits-for-fracking-gas-wells/.

2.5.2.2. Alternative Approach

If declining groundwater levels (i.e., water availability) are observable to operators, they may respond by using less water in well stimulations,[61] which biases (albeit attenuates towards zero) the parameter estimates on the total water use terms.[62] The ideal solution would be to adopt an instrumental variable (IV) strategy, where I would use an IV that is correlated to total water volumes in hydraulic fracturing and only affects groundwater levels through that pathway.[63] Given a lack of data for an IV at the time of the analysis, in the spirit of Granger (1969) I instead opt for the use of lagged total water use terms to identify the effect on groundwater levels.[64] I specify this as follows:

$$GWL_{i,t} = \sum_{j=0}^{m} \delta_{-j} \cdot TWV10_{i,t-j} + \boldsymbol{x\beta} + \lambda_t + \gamma_i + \varepsilon_{i,t}. \tag{2.3}$$

---

[61] Stevens and Torell (2018) find that during an exceptional drought in Texas during 2011 and 2012, operators completed fewer wells and completed wells using less water, which had an immediate impact on production.

[62] Picture a chart with distance to groundwater level on the vertical axis and time on the horizontal axis, which has two trend lines representative of groundwater levels in two otherwise equal aquifers with no recharge. The first is trending slightly upward but at a constant rate, and the second is trending similarly until time *t*, at which point it becomes more upward sloping. When groundwater extraction occurs at time *t*, it causes the distance from the surface to groundwater level to increase, and higher withdrawal costs should be observed at time *t*+1. If operators are responsive to these increasing costs, subsequently smaller withdrawals will occur in periods *t*+*k*. Hence, this distance to groundwater level should increase at a *decreasing* rate. Relative to the groundwater level in the aquifer absent withdrawals for hydraulic fracturing, this should cause the magnitude of the effect of withdrawals to decline over time, attenuating the average effect across the whole period.

[63] I have yet to find a valid IV that is only related to groundwater levels through water use in hydraulic fracturing. One approach to explore in future research is to obtain isopach maps for each shale play in Texas. These maps provide information on geological quality, such as the thickness of the formation, which can be interpreted as an indicator of the productive potential at given points across the formation. I will use this spatial variation in geological quality below monitoring stations interacted with time to identify an effect on groundwater levels (the interaction with time enables use in a monitoring station-level fixed effects model). This approach is valid so long as geological quality is strongly correlated with growth in hydraulic fracturing across space, but is uncorrelated with time-varying shocks to groundwater levels. Regarding this assumption, shale geology is time invariant, as long as the thickness of the formation is not treated as a depletable resource stock. The relationship between geology and water use in hydraulic fracturing has certainly changed over time as technology improved and learning in the industry occurred.

[64] The essence of Granger causality can be explained by using lagged terms of independent variable(s), and using t-tests and or f-tests to test for statistical significance on the lagged terms.

Since the effect of water use in the contemporaneous period is attenuated towards zero, I expect to see an insignificant parameter estimate on this term, but to see statistical significance on lagged terms. Intuitively, both the magnitude of the estimated effect and the level of statistical significance should eventually decline with each additional lagged term, since previous water withdrawals for hydraulic fracturing have *less* of an effect on contemporaneous groundwater levels the longer ago they occurred. This approach overcomes the endogeneity problem so long as *future* water availability is unobservable (or unimportant) to operators in the *contemporaneous* month, but the water use by operators in *previous* months still affects *contemporaneous* groundwater levels.

2.5.3. Cumulative Effects

After testing for a causal effect from hydraulic fracturing water use occurring over a short period, then it is clear that cumulative water use within the vicinity of a monitoring station is likely to have causal a relationship with groundwater levels as well. Using the total water volumes in each month, I create a new term, $CTWV10_{i,t}$, which represents the cumulative water use that has occurred within 10 miles of monitoring station *i* at time *t*, and I re-specify equation (2.2) as follows, omitting the annulus terms:

$$GWL_{i,t} = \sigma \cdot CTWV10_{i,t} + \boldsymbol{x\beta} + \lambda_t + \gamma_i + \varepsilon_{i,t}. \tag{2.4}$$

The specification enables me to estimate the cumulative effects of water use in hydraulic fracturing, relative to monitoring stations absent nearby hydraulic fracturing activity.[65]

[65] Ex-ante, it was unclear about which effect should be bigger, the effect from lagged (contemporaneous) or cumulative water use. After discussions with a hydrologist, it was made clear that the effect of lagged contemporaneous water use should be expected to be larger, since under the cumulative case, aquifers are able to recharge more over a longer period, so a one-unit increase in monthly total water use should have a

In alternate specifications of equation (2.4), I interact $CTWV10_{i,t}$ with indicators for whether monitoring station *i* is located within a GCD ($CTWV10_{i,t} \cdot GCD_{i,t}$), overlays the Permian Basin ($CTWV10_{i,t} \cdot Permian$), or overlays the Eagle Ford shale ($CTWV10_{i,t} \cdot Eagle\ Ford$). Although I estimate separately the specification with the GCD interaction term from the specification with interactions with each shale region, the interpretations are similar. The parameter on the GCD interaction term is informative to show how water use in hydraulic fracturing occurring in GCD areas impacts groundwater levels, relative to groundwater levels in non-GCD areas. The parameters on the Permian and Eagle Ford interaction terms reflect the impacts of water use in each respective region, relative to groundwater levels outside of these areas. I hypothesize that the parameters on each of the three interaction terms will be positive, indicating that impacts are larger in these areas where water availability is more of a concern (GCD areas), and where the majority of hydraulic fracturing activity is occurring (Permian and Eagle Ford).

2.6. Results

2.6.1. Strategic Reporting of Water Use

In Table 3, the results from four linear probability models are presented to analyze the reporting tendencies of operators using the first metric for reporting. The final dataset includes 47,521 well records from 255 operators over February 2012-May 2017, which reflects the omission of well records for operators that do *not* have a well in both GCD and non-GCD areas, enabling me to include a full set of operator fixed effects and isolate

---

larger immediate effect than a one-unit increase in the cumulative total water use that may be spread across a longer horizon.

the within-operator effect of locating a well within a GCD.[66] Since Konschnik and Dayalu (2015) find that rates of withheld chemical information from FracFocus completion reports increased from 2013-2015, I also include month-of-sample fixed effects to absorb confounders related to time that are common to wells in GCD and non-GCD areas.

In each model, I find that if a well was stimulated in a GCD there was a *decline* in the likelihood of an operator reporting detailed information on water use— statistically significant at the 10% level in all specifications. In column 2, total water volume was added and shows that a marginal increase in the total water volume (TWV) used in well stimulations is associated with a small, but statistically significant *decline* in the likelihood of an operator reporting information on water use beyond what is required. In column 3, an indicator was added to control for whether the well was a re-fracture, as well as additional controls for the number of wells drilled by the largest three operators in the same county-month that well *i* was completed, whether the well was an operator's first well, and the operator's first well in that county.

In column 4, my preferred specification, I add controls for rain and drought, and estimate that if a well was stimulated in a GCD area there was a *decline* of 1.29 percentage points in the likelihood of an operator reporting detailed information on water use. I find a more extreme result with respect to drill orientation. Relative to vertically drilled wells, for horizontal and directionally drilled wells there is a decline in the likelihood of detailed

[66] Additional summary statistics are available in Table 22 in Appendix B. A total of 255 unique operators had wells in both GCD and non-GCD areas, 253 unique operators only had wells in GCD areas, and 148 unique operators had wells in non-GCD areas only.

reporting by 4.93 percentage points. For a 100,000-gallon increase in TWV there is an associated with a decline in the likelihood of reporting of .013 percentage points.

In Table 4, the results are presented for the same models, but estimated using the second metric for reporting, and I find that the likelihood of a record for a well located in a GCD containing any information on a water type was even lower. In column 4, I estimate a difference in the likelihood of reporting information on water types to be 2.8 percentage points lower for wells located in a GCD relative to a non-GCD area. I find a similar, although statistically insignificant coefficient on the total water volume term, and an even larger negative coefficient on the drilling orientation term (statistically significant at the 1% level). I estimate that a completion report for a horizontally or directionally drilled well is 8.81 percentage points less likely to contain information on a water type than a well that was drilled vertically. The other terms were not statistically significant.

In Appendix B.D.1, I report the results for the same specifications as those in Tables 3 and 4, but estimated via logit.[67] Although the interpretation of each coefficient's magnitude is unintuitive, the sign on each estimate remained unchanged from those in Tables 3 and 4, and additional statistical significance was gained for the treatment variable. I again find that if a well was stimulated within a GCD area, there was a *decrease* in the likelihood of an operator providing additional information on water use, which is statistically significant at the 5% level, as shown in column 4 in each respective table. Similar significance levels were found on the estimates for both total water volume and

[67] See Tables 23 and 24 in Appendix B.

refrac, indicating that the model passed one robustness check.[68] These results are important as they indicate some awareness on behalf of operators' towards an important environmental concern. Such strategic responses by operators are very much rational, but they make water management and policy-making more difficult as poorer quality data are available to base policy decisions on that may be helpful to change incentive structures towards water conservation.

2.6.2. The Causal Effects of Hydraulic Fracturing on Local Groundwater Levels

2.6.2.1. Preliminary Spatial Evidence

The ideal natural experiment to test whether water use in hydraulic fracturing affects groundwater levels would involve using spatially concentrated data on groundwater levels in areas with and without a proximity to shale geology (both before and after hydraulic fracturing), as well as data on total water use dating back to the first high-volume well stimulation in each region. In other words, I would prefer to have data on each of these variables from the early 2000s until the present. However, since I have a limited amount of data on water use in hydraulic fracturing pre-2012 when reporting was not required; my analysis of the effects on groundwater levels is restricted to the time period of 2011-May 2017,[69] which also reduced the number of monitoring stations in my dataset to 267. Before presenting the main results from the groundwater models, I show

---

[68] Although I do not provide the regression results, I also re-estimated these same specifications for both the linear probability and logit models using the full sample of operators, i.e. those with wells in GCD areas only, those with wells non-GCD areas only, and those with wells in both GCD and non-GCD areas. I find very small changes in the magnitudes of the parameters and no changes in the significance levels.

[69] Note: I include total water volumes for 2011, even though reporting was not required at this time. This is reasonable since the operators that reported in this year still used water that was pulled from the ground, yet my total water use calculations for 2011 are likely underestimated since not all wells reported.

evidence that the magnitude of the effect of water use in hydraulic fracturing dissipates across space. Figure 9 shows the coefficients from equation (2.2), which estimate the effects on groundwater levels due to water use within a 10-mile radius and outer 5-mile annuli.[70]

Although informative to confirm previous hypotheses, I drop the annulus terms in subsequent models for two main reasons. First, my goal is not to estimate a particular distance at which water use in hydraulic fracturing affects groundwater levels. Instead, it is to show that within the vicinity of a monitoring station, it is large enough to have an empirically discernable effect. Second, there are several issues with the spatial aspects (annulus terms) in equation (2.2).[71] Given that groundwater levels and withdrawals are correlated across space, one potential remedy is to adjust the standard errors, although I have not made such adjustments since the focus of these results is not those across space.[72]

In Table 5, I present the results from other various specifications of equation (2.2). Columns 1-3 provide the estimates for a baseline specification and two others using different sets of drought, rain, temperature, and wind controls. Although I would prefer to

---

[70] Regression results available in Table 25 in Appendix B.

[71] First, the surface area in the initial radius and in each distant ring area not the same. This means that I am measuring the effects of water use coming from different sized areas, so in this specification some terms systematically have more wells, and therefore more total water use, than others. Although the goal of this study is not to determine an exact distance or extent that water use in hydraulic fracturing affects groundwater levels, to circumvent this issue, the area in each term should be standardized in order estimate water use coming from equal-sized areas.

Second, there is measurement error in the total water use terms and noise in my estimates if water withdrawals did not occur within the 10-mile radius (or subsequent annuli) that I am assuming they did. Similarly, there is measurement error if the withdrawals are not coming from the same aquifer from which the monitoring station is reading.

[72] Conley standard errors (Conley and Molinari 2007) may help to correct for spatial and temporal correlations associated with the error my approach introduces.

include temperature and wind in each model, not many stations across counties actually recorded this information; so as a result, I lose a large number of observations during estimation of the specification in column 3. I omit these two variables in column 4, but add a set of controls for population, irrigated corn, cotton, sorghum, and wheat acreage, and the total number of acres of corn, cotton, sorghum, wheat, and rice in the county of each groundwater monitoring station. In column 5, I estimate the same specification as in column 4, but include interaction terms between total water use and indicators for the Permian and Eagle Ford instead of total water use for all areas.

The signs on the coefficients for water use within 10 miles of a monitoring station (TWV10) are all in the predicted direction. The positive coefficient means that for a 100-barrel (or 4,200 gallon) increase in the monthly total water use in hydraulic fracturing within 10 miles of a monitoring station, the groundwater level declines due to extraction, therefore *increasing* the distance from the surface to the water level below the monitoring station. However, they are all statistically insignificant with the exception of column 3, which includes temperature and wind controls and was estimated with few observations. In column 5, I interact TWV10 with indicator variables for whether the monitoring station is located within either of the two most prominent unconventional oil and gas-producing areas. The significance levels on these estimates provide evidence that, at least in the two areas where significant hydraulic fracturing activity is occurring in Texas, groundwater levels are responsive to water use in the contemporaneous period. However, if operators observe water availability (and associated changes in the price of water) and respond by using less water in hydraulic fracturing stimulations during times of drought, I am facing a simultaneity issue in each of these specifications. Meaning, my contemporaneous parameter estimates on the total water volume term in columns 1, 2, and 4 are biased, although attenuated towards zero.

2.6.2.2. The Short Term Causal Effects

In Table 6, I present the estimates for various specifications of equation (2.3), using lagged terms for monthly total water use, drought, and rain. In column 1, I model groundwater levels only as a function of total water use and monitoring station and year-month fixed effects. In column 2, I control for drought and rain in the contemporaneous period as well. Column 3 in Table 6 is the same specification as column 3 in Table 5, where I control for temperature and wind since it is reasonable to assume that each of these variables has a relation to groundwater levels, but I lose nearly one third of the observations in estimation due to a lack of data for these variables. In column 4, I control for additional water use from other sources, as well as lagged drought terms. In column 5, I add a set of rain lags as well. Following Granger causality, I would expect parameter estimates on total water use in the contemporaneous period to be statistically insignificant, but an effect to be seen in previous periods that gets smaller with more distant lags since it is intuitive that past withdrawals should have less of an effect the farther back in time they occurred.

Looking at columns 1, 2, 4, and 5, the results are somewhat mixed, and do not explicitly confirm the intuition on the magnitudes nor significance levels on the lagged terms. Yet, past water use in hydraulic fracturing still is shown to be large enough to have a persistent effect on future groundwater levels, particularly in lags 3-5 in column 5, my preferred specification. I chose to use six lags as statistical significance was lost on the sixth lag. Each column in Table 6 shows that water withdrawals for use in hydraulic fracturing occurring in the 3-5 previous months all have a statistically discernable effect on groundwater levels in the contemporaneous period. The results in column 2 show that withdraws occurring in the previous two months do as well. To interpret, I show that for a 73 million gallon increase in the monthly total water volume used in hydraulic fracturing (roughly five new 2017 Permian Basin wells) within 10 miles of a monitoring station, the distance from the surface to the groundwater level below the station increases by about 1.93 feet.[73]

As a robustness check for the results above, Angrist and Pischke (2009) discuss how the inclusion of leading terms in the above equation should not change the results if there is a causal relationship. That is, if water use in hydraulic fracturing causes declines in groundwater levels but not vice versa, then significant coefficients on lagged terms should remain and the leading terms for total water use should *not* have a statistically significant effect. I specify this model as follows, where each leading term is effectively a placebo used to test equation (2.3) for robustness:

$$GWL_{i,t} = \sum_{j=0}^{m} \delta_{-j} \cdot TWV10_{i,t-j} + \sum_{j=1}^{n} \delta_{+j} \cdot TWV10_{i,t+j} + \boldsymbol{x\beta} + \lambda_t + \gamma_i + \varepsilon_{i,t}. \quad (2.5)$$

In words, conditional on monitoring station and year-by-month fixed effects, this specification is designed to show that past treatments (water withdrawals) should predict

[73] To estimate this effect, I changed the units on total water use in the 4th lagged term by 73 million gallons (14,600,000*5 gallons = 73 million gallons = five 2017 Permian Basin wells) and re-estimated the specification in column 5 of Table 6.

contemporaneous groundwater levels, but future withdrawals do not. In my estimation of equation (2.5), I included six leading total water volume terms in addition to the lags and the same sets of controls in Table 6 except omitting the specification with temperature and wind. As shown in Appendix B.D.2.B, all leading terms across all specifications are statistically *in*significant, as well as for contemporaneous TWV10, yet statistical significance on the lagged terms remains.[74] I conducted additional robustness checks as explained in subsection 2.7, and the results provide strong evidence that water use in hydraulic fracturing occurring in a short time interval affects local groundwater availability in Texas.

2.6.2.3. The Long Term Causal Effects

After investigating the effects of water use occurring over a short period, it is also of interest to see how these results compare to the cumulative effects, which essentially is another form of a robustness check if the estimated parameters for short term and cumulative water use match expectations. In Table 7, I present results for various specifications of equation (2.4). In each specification, I use a complete set of control variables (as in column 5 of Table 6), as well as monitoring station and year-by-month fixed effects. In column 1, I estimate the effects of cumulative water use on contemporaneous groundwater levels. In columns 2 and 3, I estimate the effects of cumulative water use occurring in GCD areas, and in the regions of the Permian and Eagle Ford shale, respectively. As shown above, attenuation bias rendered insignificant the

---

[74] Regression results available in Table 26 in Appendix B.

parameter estimates for the effect of water use in a given month on contemporaneous groundwater levels. However, for cumulative water use, the estimated parameters are all statistically significant at the 1% level in the contemporaneous period.

Comparing the effects of *monthly* total water use in in Table 6 to those for *cumulative* total water use in column 1, I find that the magnitudes match expectations. That is, the effect of a 100-barrel increase in the *cumulative* water (column 1, Table 7) use has a much smaller effect on groundwater levels than does a 100-barrel increase in lagged *monthly* total water use (the effect is smaller than each contemporaneous and lagged TWV coefficient in columns 1-5 in Table 6). I find that for a 73 million gallon increase in the *cumulative* total water use in hydraulic fracturing (i.e. roughly five new 2017 Permian Basin wells) within 10 miles of a monitoring station, the distance from the surface to the groundwater level below the station increases by about .41 feet. In column 2, I estimate a model specified to show the heterogeneous impacts of water use in hydraulic fracturing occurring in GCD and non-GCD areas and, as expected, I find that water use in GCD areas has a larger impact than in non-GCD areas. I estimate that for a 73 million gallon increase in the cumulative total water use in hydraulic fracturing within 10 miles of a monitoring station that is located in a GCD area, the distance from the surface to the groundwater level increases by .404 feet, relative to non-GCD areas with and without hydraulic fracturing activity.[75]

In column 3, I find similar results for water use in the major hydraulic fracturing regions in Texas, and show heterogeneous impacts in the Permian and Eagle Ford regions. I estimate that for a 73 million gallon increase in the cumulative total water use in hydraulic fracturing within 10 miles of a monitoring station, the distance to the groundwater level increases by an average of .19 and .59 feet in the Permian and Eagle Ford regions, respectively, relative to other areas with and without hydraulic fracturing. The larger magnitude on the effect for water use occurring in the Eagle Ford indicates that the aquifers in this area are more susceptible to depletion since groundwater levels are more responsive, given an equal change in water use. Collectively, the results in this section are helpful to show that there is a valid reason for operators to be concerned with future water availability. They also make clear that in areas where water scarcity is a concern, incentives exist for operators to report water use with less detail if they wish not to be identified as a culprit for impacts on water availability.

2.7. Discussion

The main results in this paper uncover details on several important issues related to water use in hydraulic fracturing. In subsection 2.6.1, I show that operators are less likely to report detailed information on water use when a well is located within a GCD area. The findings are robust to a logit specification, omitting observations for operators

---

[75] Similar to the effects for total monthly water use, to estimate this effect, I changed the units on cumulative total water use by dividing by 73 million gallons (14,600,000*5 gallons = 73 million gallons = five 2017 Permian Basin wells) and re-estimated the specification in column 1 of Table 7.

that do not have wells in both GCD and non-GCD areas, and including operator and year-month fixed effects. Since the hydraulic fracturing fluid mass calculation can be considered a more restrictive metric for reporting, I find that the effect under this specification is likely to be a lower bound on the difference in reporting. In the alternative specification, where I use a reporting metric specific to completion reports containing *any* information on water types, the difference in reporting is larger, and operators are even less likely to detail this information if the well is in a GCD.

The results from the groundwater models in subsection 2.6.2 show that short-term water use in hydraulic fracturing has a causal effect on local groundwater levels, but the magnitudes and significance levels of the estimated parameters in Table 6 do not perfectly exhibit what was expected for water use occurring in previous months. They also show that there are long term or cumulative effects of this use, which vary by region and whether or not the water use is occurring within a GCD area. However, the magnitude of the estimates on water use could be too high or too low for several reasons.

My estimates could be too high if there were other groundwater withdrawals occurring that are correlated with withdrawals for hydraulic fracturing. To circumvent this issue, I gathered data that reflects the impacts on groundwater levels from other users, including the population, and total irrigated acres and total acres of five water-intensive crops for each county in each year. Although these data are only available in an annual frequency, they are still important as I obtain positive coefficients on irrigated corn and cotton acreage, indicating that as the number of irrigated acres increases, groundwater levels decline. I also find similar coefficients for the effects of drought, and a correct sign on the rain variable. As an additional robustness check, my results pass leave-one-out and leave-many-out tests, where I arrive at roughly the same conclusions when omitting groundwater-monitoring stations located with the Permian and or Eagle Ford Basins (see Appendix B.D.2.C).[76] Since it is plausible that the severe drought occurring in Texas in 2011 and 2012 may affect my results, or at least might affect water use by operators the most in these years, I also conduct a leave-2011-and-2012-out test, which my findings are robust to as well (see Appendix B.D.2.D).[77]

My estimates could be too low, i.e. attenuated towards zero, as there are various avenues for measurement error associated with my total water volume variables. First, since I am using reported estimates of total water volumes, which I did not manipulate in any way, the total water volumes I attribute to groundwater use would be too large if a large portion of the water used in oil and gas wells came from an alternative source. Similarly, my total water volumes could be too low if operators underreported them, whether intentionally or due to a miscalculation in some other way.[78] Lastly, Lin et al. (2018) state that total water volume used for shale oil development is underestimated when using only the hydraulic fracturing water use derived from the oil and gas well databases, particularly since databases such as FracFocus only reflect water used in well stimulation. They do not include water use associated with drilling, rig workers, or other well pad uses or temporary populations that would likely not have occurred in the absence of oil and gas

---

[76] Regression results available in Table 27, Table 28, Table 29, Table 30, Table 31, and Table 32 in Appendix B.
[77] Regression results available in Table 33 in Appendix B.
[78] There is currently no monitoring of the accuracy of reporting, so underreporting is plausible.

development.[79] Hence, my estimates of the effect of water use would be attenuated if the net of these unobserved avenues for measurement error were non-zero.

My primary policy recommendation is for Texas and other states or countries with hydraulic fracturing to expand the reporting requirements of operators to conform to those in Louisiana, which require the disclosure of the water source(s) (or types) and associated volume(s) used in hydraulic fracturing stimulations (Hanson 2011).[80] This level of detail would provide useful information to policy makers about where operators obtain water and, when combined with improved real time data, can be used to better understand how withdrawals affect water stocks. Measurement error on completion reports may still occur, intentionally or unintentionally.[81,82] Yet, when the policy goal is to manage groundwater resources, the empirical shortcomings in this paper emphasize the importance of and the need for precise data on the water use of the industry (and other users), as well as on water availability.

Future work in this area might investigate options to induce accurate reporting such as through random monitoring protocols during well completions or subsequent monitoring if a completion report is not submitted to FracFocus within a required

---

[79] They estimate that the water use of rig workers in the Bakken shale is equivalent to ~15% of annual industrial water use for shale oil development there.

[80] Within twenty days after completion or recompletion operations information must be included on the 'Well History and Work Resume Report'.
Source: http://www.dnr.louisiana.gov/assets/OC/eng_div/Forms/WH-1.pdf.

[81] Groundwater data availability would be improved by constructing more groundwater monitoring stations or other infrastructure, such as satellites, capable of recording high-frequency data (e.g. Richey et al. 2015; Chen et al. 2016).

[82] Discussions with the industry have made it clear that 'fat fingering' sometimes occurs during reporting, where operators may intentionally smudge information on completions reports. Other times, there may be unintentional human error occurring throughout the measurement and or reporting process.

timeframe.[83] Also needed is a better understanding of operator responsiveness to changes in the price of water (or availability), which is difficult when there is poor transparency in the water use and prices the industry pays for water. Incentivizing the use of online water marketing platforms such as *sourcewater.com* might be one way to collect meaningful data on water use and prices, since it would provide a formal accounting mechanism for water transactions. Alternatively, it how providing tax incentives to operators for using freshwater alternatives influences water use.

2.8. Conclusion

Water scarcity is one of the biggest constraints imposed on economic development, and it has become a growing concern in arid regions of the U.S. and other parts of the world. The boom in unconventional gas and oil development has created major concerns about its impacts on water resources, especially since it is occurring in many areas with falling aquifer levels. Understanding how the industry interacts with and influences local water levels is complicated. This is due to variation, complexity, ambiguity and poor cohesion in state and local policies governing ground and surface water rights and use,

[83] To develop a reasonable estimate about when a well stimulation and completion will occur, relatively accurate information on the location of new wells is known when drilling permit applications are submitted. Well pad development can also be observed from satellite imagery, similar to what sourcewater.com uses to locate and create its database of 'frac ponds'.

poor and inconsistent quality in the data on the industry's water use,[84,85,86] and minimal data on water availability.[87] These complexities have previously made it difficult to evaluate the industry's water use, let alone establish accurate impacts associated with its use or design appropriate regulatory responses.[88]

In this paper, I overcome several of these hurdles to study two timely issues related to water use and hydraulic fracturing. I show that operators report water use with less detail when a well is located in a groundwater conservation district, and that the water used by the industry affects local availability. It is neither my goal nor conclusion to say that the industry uses too much water, because it is likely the water it uses is going to the use with the highest monetary value.[89] However, my findings are valuable because they provide credible evidence of the impacts from one industry whose water use is often of local concern, and they have direct policy implications for areas new to unconventional oil and gas development.[90]

---

[84] Reporting to FracFocus is now required in most states.
Source: http://fracfocus.org/sites/default/files/fracfocus_reporting_states_2-7-18-01_1.png.

[85] There is significant heterogeneity in state-level reporting requirements, yet FracFocus provides only one version of its completion report form for operators of wells in all states to fill out online and submit.
Source: https://stateimpact.npr.org/texas/2013/04/26/harvard-report-gives-failing-grade-to-fracfocus-texas-regulators-respond/.

[86] A lack of monitoring of completions at the well pad, along with no checks for accuracy of the information submitted to FracFocus, have created a number of pathways for measurement error to occur in completion reports.

[87] Conversations with the TWDB indicate that the current number of monitoring stations is limited by staffing to maintain them, but also funding (usually from GCDs) to establish new stations.

[88] Without this information, along with water prices, it is difficult to evaluate the marginal user cost of water, or the forgone future value of a marginal unit of water due to its use today.

[89] Allen et al. (2014) estimate the value of water used in hydraulic fracturing is several-fold greater than other uses such as agriculture.

[90] For example, the Vaca Muerta formation in Argentina holds some of the world's largest deposits of shale gas, and similar to the Permian Basin, it is located in a relatively water scarce region.

References

Abramov, A. 2016. 'Crude, Condensate Output Refuse to Decline; Average Well Configuration, Productivity Continue to Grow.' Oil & Gas Financial Journal Back Issues. Available at: https://www.ogj.com/articles/ogfj/print/volume-13/issue-5/features/permian-basin.html. [Accessed 7/19/17].

Agerton, M. 2019. "Learning Where to Drill: Drilling Decisions and Geological Quality in the Haynesville Shale." Working Paper. Available at: http://markagerton.com/files/Agerton_LearningWhereToDrill.pdf. [Accessed 4/14/19].

Allen, W.T., R.D. Lacewell, and M. Zinn. 2014. "Water Value and Environmental Implications of Hydraulic Fracturing: Eagle-Ford Shale." Texas Water Resources Institute Paper TR-466. Available at: https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/152429/TR-466%20Water%20Value%20and%20Environmental%20Implications%20of%20Hydraulic%20Fracturing.pdf?sequence=1&isAllowed=y. [Accessed 1029/17].

Angrist, J., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st. ed. New Jersey: Princeton University Press.

Barth-Naftilan, E., N. Aloysius, and J.E. Saiers. 2015. "Spatial and Temporal Trends in Freshwater Appropriation for Natural Gas Development in Pennsylvania's Marcellus Shale Play." *Geophysical Research Letters*, 42: 6348-6356.

Bartik, A.W., J. Currie, M. Greenstone, and C.R. Knittel. 2017. "The Local Economic and Welfare Consequences of Hydraulic Fracturing." NBER Working Paper 23060. Available at: http://www.nber.org/papers/w23060. [Accessed 9/28/18].

Black, K.J., S.J. McCoy, and J.G. Weber. 2017. "When Externalities Are Taxed: The Effects and Incidence of Pennsylvania's Impact Fee on Shale Gas Wells." *Journal of the Association of Environmental and Resource Economists*, 5(1): 107-153.

Boudet, H., C. Cake, D. Bugden, E. Maibach, C. Roser-Renouf, and A. Leiserowitz. 2014. ""Fracking" Controversy and Communication: Using National Survey Data to Understand Public Perceptions of Hydraulic Fracturing." *Energy Policy*, 65: 57-67.

Boudet, H., D. Bugden, C. Zanocco, and E. Maibach. 2016. "The Effect of Industry Activities on Public Support for 'Fracking'." *Environmental Politics*, 25(4): 593-612.

Burnett, J.W. 2013. “Hydraulic Fracturing and U.S. Water Policy.” Global Water Forum Discussion Paper 1309. Available at: http://www.globalwaterforum.org/wp-content/uploads/2013/03/Hydraulic-fracturing-and-US-water-policy-GWF-1309.pdf. [Accessed 4/3/18].

Burnett, J.W. 2015. “Unconventional Oil and Gas Development: Economic, Environmental, and Policy Analysis.” *Agricultural and Resource Economics Review,* 44(2): 3-27.

Bush, T. 2017. ‘West Texas inSANDity Continues in the Permian’. Available at: http://energentgroup.com/west-texas-insandity-continues-permian/. [Accessed 10/5/17].

Cavender, A.B. 2011. ‘Texas Law Requires Disclosure of Hydraulic Fracturing Chemicals as of February 1, 2012.’ Available at: https://www.pillsburylaw.com/images/content/3/9/3903.pdf. [Accessed 6/8/17].

Carr, H. 2017. ‘Wipe Out! - How Will Permian E&Ps Dispose Of All That Produced Water?’ RBN Energy LLC Daily Blog. Available at: https://rbnenergy.com/wipe-out-how-will-permian-eandps-dispose-of-all-that-produced-water. [Accessed 7/18/17].

Chatterji, A.K., and M.W. Toffel. 2010. “How Firms Respond to Being Rated.” *Strategic Management Journal,* 31: 917-945.

Chen, J., R. Knight, H.A. Zebker, and W.A. Shreuder. 2016. “Confined Aquifer Head Measurements and Storage Properties in the San Luis Valley, Colorado, from Spaceborne InSAR Observations.” *Water Resources Research,* 52: 3623-3636.

Conley, T.G., and F. Molinari. 2007. “Spatial Correlation Robust Inference with Errors in Location or Distance.” *Journal of Econometrics,* 14: 76-96.

Currie, J., M. Greenstone, and K. Meckel. 2017. “Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania.” *Science Advances,* 3(12): 1-9.

Doshi, A.R., G.W.S. Dowell, and M.W. Toffel. 2013 “How Firms Respond to Mandatory Information Disclosure.” *Strategic Management Journal,* 34: 1209-1231.

Ellsworth, W.L. 2013. “Injection-Induced Earthquakes.” *Science,* 341(6142): 1225942.

Estaville, L.E., and R.A. Earl. 2008. *Texas Water Atlas,* 1st ed. College Station: Texas A&M University Press.

Farah, N. 2017. “The Effects of Hydraulic Fracturing on Agricultural Productivity.” Working Paper. Available at:

https://drive.google.com/file/d/0Bxp0L418FENjWklucjI3RzFvVkE/view. [Accessed 11/10/17].

Fetter, R.T. 2017. "Fracking, Toxics, and Disclosure." Working Paper. Available at: https://sites.duke.edu/trfetter/files/2016/11/fracking-toxics-disclosure.pdf. [Accessed 3/15/18].

Fetter, R.T., A.L. Steck, C. Timmins, and D.H. Wrenn. 2017. "Learning by Viewing? Social Learning, Regulatory Disclosure, and Firm Productivity in Shale Gas." Working Paper. Available at: https://shared.aces.illinois.edu/sites/shared.aces.illinois.edu/files/learning-viewing-paper_April_2017.pdf. [Accessed 5/20/18].

Feyrer, J., E.T. Mansur, and B. Sacerdote. 2017. "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution." *American Economic Review,* 107(4): 1313-1334.

Fitzgerald, T. 2013. "Frackonomics: Some Economics of Hydraulic Fracturing." *Case Western Reserve Law Review,* 63(4): 1337-1362.

Freyman, M. 2014. "Hydraulic Fracturing & Water Stress: Water Demand by the Numbers." Shareholder, Lender, & Operator Guide to Water Sourcing – A Ceres Report. Available at: https://www.ceres.org/resources/reports/hydraulic-fracturing-water-stress-water-demand-numbers. [Accessed 6/15/17].

Gold, R. 2014. *The Boom*, 1st. ed. New York: Simon & Schuster.

Goldenberg, S. 2013. "A Texas Tragedy: Ample Oil, No Water." The Guardian – Environment. Available at: https://www.expressnews.com/business/eagle-ford-energy/article/Bigger-wells-but-more-water-12512262.php. [Accessed 8/9/18].

Granger, C.W.J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica,* 37(3): 424-438.

Hanson, G. 2011. 'How Are Appropriate Water Sources for Hydraulic Fracturing Determined? Pre-Development Conditions and Management Phase Water Usage.' Louisiana State University Shreveport, as part of the proceedings of the Technical Workshops for the Hydraulic Fracturing Study: Water Resource Management, EPA. Available at: https://www.epa.gov/sites/production/files/documents/01a_Hanson_-_Appropriate_Water_Sources_508.pdf. [Access 10/8/18].

Hardberger, A. 2016. "Texas Groundwater Markets and the Edwards Aquifer: A Case Study for the Political Economy of Water Markets Project." Unpublished, St. Mary's University School of Law. Available at:

https://static1.squarespace.com/static/56d1e36d59827e6585c0b336/t/5805468415d5dbb1ab59a3a9/1476740751543/Texas-Groundwater-Markets-Hardberger.pdf. [Accessed 8/15/17].

Hausman, C., and R. Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas." *Brookings Papers on Economic Activity,* 46(1): 71-139.

He, X., N. Lu, and R.P. Berrens. 2017. "The Case of the Missing Negative Externality? Housing Market Effects of Fracking in the Niobrara Shale Play, Colorado." *Journal of Environmental Economics and Policy,* 7(3): 223-243.

Herrnstadt, E., R. Kellogg, and E. Lewis. 2019. "Information Asymmetry, Drilling Distortions, and Oil and Gas Leases." Working Paper. Available at: https://home.uchicago.edu/~kelloggr/Papers/HKL_primaryterms.pdf. [Accessed 1/22/19].

Hill, E., and L. Ma. 2017. "Shale Development and Drinking Water Quality." *American Economic Review: Papers & Proceedings,* 107(5): 522-525.

Hitaj, C., A. Boslett, and J.G. Weber. 2014. "Shale Development and Agriculture." *Choices,* 29(4): 1-7.

Hitaj, C., A. Boslett, and J.G. Weber. 2017. "Fracking, Farming, and Water." Working Paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3041868. [Accessed 9/28/18].

Holladay, J.S., and J. LaRiviere. 2017. "The Impact of Cheap Natural Gas on Marginal Emissions from Electricity Generation and Implications for Energy Policy." *Journal of Environmental Economics and Management,* 85: 205-227.

Houston & Texas Central Railroad Co. v. W.A. East. 1904. 98 Tex. 146, 81 S.W. 279.

Horner, R.M., C.B. Harto, R.B. Jackson, E.R. Lowry, A.R. Brandt, T.W. Yeskoo, D.J. Murphy, and C.E. Clark. 2016. "Water Use and Management in the Bakken Shale Oil Play in North Dakota." *Environmental Science & Technology,* 50: 3275-3282.

Howarth, R.W., A. Ingraffea, and T. Engelder. 2011. "Natural Gas: Should Fracking Stop?" *Nature,* 477: 271-275.

Hubert, M. and B. Bullock. 1999. "The First Big and Bold Step Toward Meeting Texas's Future Water Needs." *Senate Bill 1.* 30 TEX. TECH. L. REV. 53, 66.

Jackson, R.B., A. Vengosh, J.W. Carey, R.J. Davies, T.H. Darrah, F. O'Sullivan, and G. Petron. 2014. "The Environmental Costs and Benefits of Fracking." *Annual Review of Environment and Resources,* 39: 327-362.

James, A., and B. Smith. 2017. "There Will Be Blood: Crime Rates in Shale-Rich U.S. Counties." *Journal of Environmental Economics and Management,* 84: 125-152.

Knittel, C.R., K. Metaxoglou, and A. Trindade. 2015. "Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets." NBER Working Paper No. 21627 - Oct 2015.

Kondash, A.J., N.E. Lauer, and A. Vengosh. 2018. "The Intensification of the Water Footprint of Hydraulic Fracturing." *Science Advances,* 4(8): 1-8.

Kondash, A., and A. Vengosh. 2015. "Water Footprint of Hydraulic Fracturing." *Environmental Science & Technology,* 2: 276-280.

Konschnik, K., and A. Dayalu. 2015. "Hydraulic Fracturing Chemicals Reporting: Analysis of Available Data and Recommendations for Policymakers." *Energy Policy,* 88: 504-514.

Krupnick, A.J. and H.G. Gordon. 2015. "What Experts Say About the Environmental Risks of Shale Gas Development." *Agricultural and Resource Economics Review,* 44(2): 106-119.

Kusnetz, N. 2012. 'The Bakken Oil Play Spurs a Booming Business – In Water.' High Country News. Available at: http://www.hcn.org/issues/44.13/the-bakken-oil-play-spurs-a-booming-business-in-water. [Accessed 6/4/17].

Kuwayama, Y., S. Olmstead, and A. Krupnick. 2015. "Water Quality and Quantity Impacts of Hydraulic Fracturing." *Current Sustainable Renewable Energy Reports,* 2: 17-24.

Lange, I., and M. Redlinger. 2019. "Effects of Stricter Environmental Regulations on Resource Development." *Journal of Environmental Economics and Management,* 96: 60-87.

Lesikar, B., R. Kaiser, and V. Silvy. 2002. 'Questions About Groundwater Conservations Districts in Texas.' Texas A&M University Report. Available at: http://twri.tamu.edu/reports/2002/2002-036/2002-036_questions-dist.pdf. [Accessed 1/17/18].

Lin, Z., T. Lin, S.H. Lim, M.H. Hove, and W.M. Schuh. 2018. "Impact of Bakken Shale Development on Regional Water Uses and Supply." *Journal of the American Water Resources Association,* 54(1): 225-239.

Lyon, T.P., and J.W. Maxwell. 2004. *Corporate Environmentalism and Public Policy*, Cambridge University Press.

Lyon, T.P., and J.W. Maxwell. 2008. "Corporate Social Responsibility and the Environment: A Theoretical Perspective." *Review of Environmental Economics and Policy,* 2(2): 240-260.

Maniloff, P. 2019. "Can Learning Explain Deterrence? Evidence from Oil and Gas Production." *Journal of the Association of Environmental and Resource Economists*, 6(5): 853-881.

Mason, C.F., L.A. Muehlenbachs, and S.M. Olmstead. 2015. "The Economics of Shale Gas Development." *Annual Review of Resource Economics,* 7: 269-289.

Matisoff, D.C. 2013. "Different Rays of Sunlight: Understanding Information Disclosure and Carbon Transparency." *Energy Policy,* 55: 579-592.

Mielke, E., L.D. Anadon, and V. Narayanamurti. 2010. "Water Consumption of Energy Resource Extraction, Processing, and Conversion." Energy Technology Innovation Policy Research Group. Discussion Paper No. 2010-15. Belfer Center for Science and International Affairs, Harvard Kennedy School, Harvard University.

Mitchell, A. L., M. Small, and E. A. Casman. 2013. "Surface Water Withdrawals for Marcellus Shale Development: Performance of Alternative Regulatory Approaches in the Upper Ohio River Basin." *Environmental Science and Technology,* 47(22): 12669-12678.

Muehlenbachs, L.A. 2012. "Testing for Avoidance of Environmental Obligations." Resources for the Future Discussion Paper 12-12. Available at. https://ideas.repec.org/p/rff/dpaper/dp-12-12.html. [Accessed 6/11/19].

Muehlenbachs, L.A., and A.J. Krupnick. 2013. "Shale Gas Development Linked to Traffic Accidents in Pennsylvania." Resources for the Future Report. Available at: http://common-resources.org/2013/shale-gas-development-linkedto-traffic-accidents-in-pennsylvania. [Accessed 12/15/17].

Muehlenbachs, L.A., and S. Olmstead. 2014. "Hydraulic Fracturing and Water Resources." *Choices,* 29(4): 1-5.

Muehlenbachs, L.A., E. Spiller, C. Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review,* 105(12): 3633-3359.

Muehlenbachs, L.A., S. Staubli, and Z. Chu. 2017. "The Accident Externality from Trucking." Resources for the Future Report. Available at:

http://www.rff.org/research/publications/accident-externality-trucking. [Accessed 12/15/17].

Nicot, J.-P. 2012. "Current and Future Water Demand of the Texas Oil and Gas and Mining Sectors and Potential Impact on Aquifers." *GCAGS Journal,* 1: 145-161.

Nicot, J.-P., and B.R. Scanlon. 2012. "Water Use for Shale-Gas Production in Texas, U.S." *Environmental Science & Technology*, 46(6): 3580-3586.

Nicot, J.-P., R.C. Reedy, R.A. Costley, and Y. Huang. 2012. 'Oil & Gas Water Use in Texas: Update to the 2011 Mining Water Use Report.' Austin, TX: Bureau of Economic Geology, University of Texas at Austin. Available at: http://www.twdb.texas.gov/publications/reports/contracted_reports/doc/0904830939_2012Update_MiningWaterUse.pdf. [Accessed 6/21/17].

Nicot, J.-P., B.R. Scanlon, R.C. Reedy, and C.A. Costley. 2014. "Source and Fate of Hydraulic Fracturing Water in the Barnett Shale: A Historical Perspective." *Environmental Science & Technology,* 48(4): 2464-2471.

Olmstead, S.M., L.A. Muehlenbachs, J.-S. Shih, Z. Chu, and A.J. Krupnick. 2013. "Shale Gas Developments on Surface Water Quality in Pennsylvania." *Proceedings of the National Academy of Sciences,* 110(13): 4962-4967.

Olmstead, S.M., N. Richardson. 2014. "Managing the Risks of Shale Gas Development Using Innovative Legal and Regulatory Approaches." *William & Mary Environmental Law and Policy Review*, 39(1): 177-199.

Provencher, B., and O. Burt. 1993. "The Externalities with the Common Property Exploitation of Groundwater." *Journal of Environmental Economics and Management,* 24: 139-158.

Rahm, D., B. Fields, B., and J.L. Farmer. 2015. "Transportation Impacts of Fracking in the Eagle Ford Shale Development in Rural South Texas: Perceptions of Local Government Officials." *The Journal of Rural and Community Development,* 10(2): 78-99.

Rahm, B.G., and S.J. Riha. 2012. "Toward Strategic Management of Shale Gas Development: Regional, Collective Impacts on Water Resources." *Environmental Science & Policy,* 17: 12-23.

Rahm, B.G., and S.J. Riha. 2014. "Evolving Shale Gas Management: Water Resource Risks, Impacts, and Lessons Learned." *Environmental Science Processes & Impacts,* 16: 1400-1412.

Richey, S.R., B.F. Thomas, M.-H. Lo, J.T. Reager, J.S. Famigliette, K. Voss, S. Swenson, and M. Rodell. 2015. “Quantifying Renewable Groundwater Stress with GRACE.” *Water Resources Research,* 51: 5217-5238.

Scanlon, B.R., I. Duncan, and R.C. Reedy. 2013. “Drought and the Water-Energy Nexus in Texas.” *Environmental Research Letters,* 8: 1-14.

Scanlon, B.R., R.C. Reedy, and J.-P. Nicot. 2014a. “Comparison of Water Use for Hydraulic Fracturing for Unconventional Oil and Gas Versus Conventional Oil.” *Environmental Science & Technology,* 48: 12386-12393.

Scanlon, B.R., R.C. Reedy, and J.-P. Nicot. 2014b. “Will Water Scarcity in Semiarid Regions Limit Hydraulic Fracturing of Shale Plays?” *Environmental Research Letters,* 9(12): 1-14.

Scanlon, B.R., R.C. Reedy, F. Male, and M. Hove. 2016. “Managing the Increasing Water Footprint of Hydraulic Fracturing in the Bakken Play, Unites States.” *Environmental Science & Technology,* 50: 10273-10281.

Scanlon, B.R., R.C. Reedy, F. Male, and M. Welsh. 2017. “Water Issues Related to Transitioning from Conventional to Unconventional Oil Production in the Permian Basin.” *Environmental Science & Policy,* 51: 10903-10912.

Schafft, K.A., Y. Borlu, and L. Glenna. 2013. “The Relationship Between Shale Gas Development in Pennsylvania and Local Perceptions of Risk and Opportunity.” *Rural Sociology,* 78(2): 143-166.

Small, M.J., P.C Stern, E. Bomberg, S.M. Christopherson, B.D. Goldstein, A.L. Israel, R.B. Jackson, A. Krupnick, M.S. Mauter, J. Nash, D.W. North, S.M. Olmstead, A. Prakash, B. Rabe, N. Richardson, S. Tierney, T. Webler, G. Wong-Parodi, and B. Zielinska. 2014. “Risks and Risk Governance in Unconventional Shale Gas Development.” *Environmental Science & Technology,* 48(15): 8289-8297.

Smith, S.M. 2018. “Economics Incentives and Conservation: Crowding-In Social Norms in a Groundwater Commons.” *Journal of Environmental Economics and Management,* 90: 147-174.

Stepan, D.J., R.E. Shockey, B.A. Kurz, N.S. Kalenze, R.M. Cowan, J.J. Ziman, and J.A. Harju. 2010. ‘Bakken Water Opportunities Assessment – Phase 1.’ Northern Great Plains Water Consortium. Prepared by the Energy & Environmental Research Center – University of North Dakota for the Industrial Commission of North Dakota. Available at: https://www.undeerc.org/water/pdf/FracWaterPhaseIreport.pdf. [Accessed 7/1/18].

Stevens, R.B., and G.L. Torell. 2018. "Exceptional Drought and Unconventional Energy Production." *Sustainability,* 10(4): 1-16.

Street, B. 2018. "The Impact of Economic Opportunity on Criminal Behavior: Evidence from the Fracking Boom." Working Paper. Available at: https://brittanystreet.github.io/website/Street_Brittany_jmp.pdf. [Accessed 4/29/19].

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. Available at: https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=332990. [Accessed 6/6/17].

U.S. Geological Survey (USGS). 2016. 'Water Questions & Answers: How much water does the average person use at home per day?' Available at: https://water.usgs.gov/edu/qa-home-percapita.html. [Accessed 11/18/17].

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.

Weber, J.G., and C. Hitaj. 2015. "What Can We Learn About Shale Gas Development from Land values? Opportunities, Challenges, and Evidence from Texas and Pennsylvania." *Agricultural and Resource Economics Review,* 44(2): 40-58.

Weber, J.G., J.W. Burnett, and I.M. Xiarchos. 2016. "Broadening Benefits from Natural Resource Extraction: Housing Values and Taxation of Natural Gas Wells as Property." *Journal of Policy Analysis and Management,* 35(3): 587-614.

3. DO FEDERAL OIL AND GAS ASSESSMENTS AFFECT LEASE ACQUISITION BEHAVIOR? A NATURAL EXPERIMENT IN THE WILLISTON BASIN

3.1. Introduction

In the upstream sector of the oil and gas industry (exploration and production), various stages of information acquisition occur before the drilling of a well, let alone production, begins. Similar to a gold rush, Agerton (2017; 2019) describes how an initial land or mineral rights "rush" occurs during the early development of a new shale play when search costs are lower. However, unconventional oil and gas plays are not homogeneous in their geology; they vary significantly across space and implying that certain areas within a formation will naturally contain more oil and gas and also be more productive than other areas.[91]

How then do firms determine where to locate? Previous literature in this area has trended towards studying how firms learn about geology by drilling new wells, and become more proficient at producing oil and gas with experience and the adoption of new technologies. Although important, these studies do not adequately address a key aspect of the exploration stage, akin to the information firms acquire and use to guide strategic location decisions to acquire acreage and subsequently produce oil and gas.[92] Instead, they implicitly assume, or confirm ex post, that firms had prior knowledge on geological

[91] Source: https://www.glossary.oilfield.slb.com/en/Terms/h/heterogeneity.aspx.
[92] Upstream firms are commonly valued based on oil and gas reserves – their primary assets. Source: https://www.stout.com/en/insights/article/valuation-methodologies-oil-gas-industry/.

quality *before* deciding to locate and acquire the options to drill in specific areas. Drilling exploratory wells and conducting private geologic research is costly, yet in the age of unconventional oil and gas development, the proficiency of oil and gas producers is highly dependent on acquiring large, contiguous amounts of the best acreage.[93] This paper contributes to this literature by studying how firms complement unobserved private knowledge with public sources of information on resource abundance *before* acquiring new acreage, and how this information affects the terms on negotiated leases– two critical stages in the development process.

It is intuitive that in a pre-development stage, firms must develop knowledge on where oil and gas deposits exist and in what abundance. They develop knowledge about an area (such as a basin) by analyzing massive amounts of data collected at a variety of scales, from spatial data on production and well logs spanning tens to hundreds of miles, to granular subsurface data on microscopic cracks in rocks. They also use data from a variety of sources, including previous and ongoing exploration efforts, and public and private repositories (Allison and Mandler 2018). Within a basin, a firm then determines more locally which hydrocarbon deposits, or formation(s), it prefers.[94] Given heterogeneity in geological quality across the spatial extent of formations, it then

[93] A conversation with personnel from Vista Oil & Gas indicated that firms often spend $2 million on exploration-related measures before drilling and completing a new well.
[94] A basin is a *large* depression in the Earth's crust, and can be thought of as the outermost bowl that contains the entirety of all petroleum accumulations within.

determines the most desirable parts to target for future extraction based on the geological quality, extraction costs, expected oil and gas prices, and other considerations.[95,96]

This data gathering and subsequent decision-making (i.e. search) process used to determine where to locate is long, complex, and requires strategic planning. According to Shell, it can take years of research and survey work to amass sufficient geological data that inform where the best places are to drill exploratory wells, and eventually where and how to drill oil and gas production wells.[97] There also appears to be a large market for this information among firms, as innovative data and analytics companies now offer premium features such as DrillingInfo's Graded Acreage tool, designed to guide firms in their search for areas with attractive geological quality and help them learn about best practices from competitors in an area.[98]

After determining potential target areas, firms must acquire mineral rights before drilling and exploration can occur. Commonly thought of as a real option (Kellogg 2014; Herrnstadt et al. 2019), the acquisition of mineral rights provides firms with the option to drill and explore for oil and gas over a guaranteed period. In the U.S., mineral rights are typically acquired via centralized auctions over state-owned minerals, or through informal negotiations with private landowners or those with ownership of the minerals. Unlike in

---

[95] I loosely define this here in terms of abundance and its mapping to expected production.

[96] These may include the ease-of-access (or economic feasibility of production), the costs associated with proximity to drilling and completion inputs (e.g. pipe, water, and sand) and takeaway options such as pipelines, the presence and composition of competitors in the vicinity, and other local conditions such as fiscal and regulatory environments.

[97] Source: https://www.shell.com/energy-and-innovation/overcoming-technology-challenges/finding-oil-and-gas.html.

[98] Source: https://info.drillinginfo.com/blog/uncover-hidden-value-graded-acreage/.

the case of lease auctions where firms more directly compete by attempting to outbid each other (and in this process more accurately reveal their true value of a prospective property), informal negotiations with landowners provide a setting that is more favorable to firms on several dimensions.[99]

First, firms are able to negotiate in relative incognito if they contract a third party to negotiate for them. Second, to acquire large amounts of contiguous acreage without sending a strategic location signal to competing firms; conversations with DrillingInfo indicate that firms routinely negotiate and informally finalize agreements with a collection of private mineral rights owners well before reporting the official lease documents publicly (which they then commonly report all at once). Lastly, in informal negotiations, asymmetric information exists to the extent of knowledge that a firm possesses on the geology and productive potential of the prospective property (or those nearby) that the opposition does not. Hence, when negotiating with mineral rights owners who presumably have less bargaining ability due to limited knowledge on the geology below their property and less resources or sophistication to figure it out, firms can more easily withhold private information and negotiate more favorable terms on leases such as a lower royalty rate, lower bonus payment, and a longer primary term.[100]

---

[99] Covert and Sweeney (2019) estimate that auctioned leases generate 67% larger up-front payments and find that they are significantly more productive than negotiated leases.

[100] Many items can appear on a lease agreement. Of most importance to this paper are:

i. the *royalty rate*, or the percent of production paid to the lessor throughout the life of the well(s) on a lease;
ii. a *per-acre bonus payment* paid to the lessor shortly after signing the lease;
iii. the *primary term*, a fixed number of years or months during which the lessee need do nothing in order to keep the lease in effect. If no exploration or production occurs during this period, a lease will terminate unless an option for a lease extension and subsequent extension bonus payment were

In this paper, I use the results of a geology-based oil and gas assessment conducted by the U.S. Geological Survey (USGS) to study how an exogenous shock of publicly available information on resource abundance affected the location of and the terms on new leases in the Williston Basin. The results, announced during the early development of the basin in April 2008, provided estimates of the total amount of unconventional oil and gas resources within each of six discretely defined areas (assessment units) within the basin. Given that the estimates provided an update to firms' private beliefs on resource stocks, I argue that the results narrowed the spatial extent of uncertainty over resource abundance across this region, and test the hypothesis that after the announcement of the assessment results firms began to negotiate leases and accumulate acreage in the areas estimated to have more resource abundance. In addition to the locational impacts of the assessment, I also study its impacts on a variety of outcomes (i.e. terms) on oil and gas leases negotiated by firms in these areas.

To preview the results, I show that in two of the top three assessment areas (i.e. those with the most estimated resources), the number of new leases post assessment increased significantly relative to the unassessed areas of the basin. In the area with the most estimated oil and gas resources, the effect of the assessment on the average size of a new lease was significantly larger than the effect for all other areas. Intuitively, this means that since lots of unleased acreage was available in this area before the assessment, fast-acting firms acquired larger leases and more contiguous space in a more resource abundant

negotiated in the lease terms. If drilling and or production occur, the lease is held in force so long as production is occurring from the lease.

area. Similarly, the assessed area with the fewest estimated resources had negative changes in both the number of new leases and also the average size of a new lease, consistent with the idea that firms opted for more certain production areas, or those with more potential. These findings either provide an indication that firms previously knew less about resource abundance across the spatial extent of the Williston Basin, or that the assessment results conflicted with prior beliefs. Lastly, I show that the terms negotiated on leases in each area changed significantly after the assessment. Of primary interest are the effects on the average primary term length, royalty rate, bonus payment, the proportion of leases with an option for a term extension, and average extension term length and extension bonus payment. I find that although firms who acquired acreage in the assessed area estimated to contain the most resources gave mineral right owners larger bonus payments, they took a larger entitlement of the lease's expected surplus by negotiating lower royalty rates.

These findings are helpful for USGS to understand how its government-funded work is used and to determine its value. More importantly, they provide insight on the information flows that help to improve proficiency in oil and gas development. The results also suggest an important area for future research as they raise questions over the welfare implications of the assessments, namely their potential contribution to asymmetric information in lease negotiations between firms and mineral rights owners. Hence, at the expense of mineral right owners, fast-acting firms were plausibly able to exploit an arbitrage opportunity to obtain better acreage and negotiate favorable terms, effectively locking in lower royalty rates and bonus payments before the lease market adjusted to reflect the higher true value of the minerals. This is important because the Fort Berthold

Reservation sits within the heart of Williston Basin, and anecdotal evidence from a lawsuit claims that Native Americans were significantly lowballed during lease negotiations, and were "cheated" out of approximately $1 billion in royalty payments when compared to the terms on leases signed by nearby counterparts located just outside of the reservation.[101]

3.2. Literature Review

Although it had been known for decades that abundant oil and gas resources were trapped in shales and tight formations (unconventional sources), prior to the mid-to-late 2000s, most oil and gas production in the U.S. came from conventional sources.[102] In other words, from sources with hydrocarbons located in parts of the geology that did not require techniques beyond drilling a vertical well (and later supplementation with enhanced oil recovery measures) to extract.[103] Yet, it was not until horizontal drilling, hydraulic fracturing, and other complementary exploration technologies were combined in a way that enabled economical extraction from unconventional resources in meaningful quantities.[104]

The "boom" in the unconventional oil and gas industry first occurred in Texas, where early firms targeted natural gas trapped in the Barnett shale. Shortly after, a similar boom for oil occurred in the Bakken formation (within the Williston Basin) of North Dakota. The resources of the Bakken propelled North Dakota to become the second largest oil-producing state in the U.S., after Texas, in 2012. As the industry evolved, so has our

---

[101] Source: https://www.propublica.org/article/land-grab-cheats-north-dakota-tribes-out-of-1-billion-suits-allege.
[102] These formations are known as "source" rocks, from which conventional oil and gas migrated.
[103] Source: https://www.glossary.oilfield.slb.com/en/Terms/s/secondary_recovery.aspx.
[104] Source: https://www.eia.gov/analysis/studies/usshalegas/pdf/usshaleplays.pdf.

understanding of the technologies and techniques that firms use to explore for and extract oil and gas. Previous studies however, are limited in their inferences about how firms acquire this knowledge, especially their knowledge about geological quality *before* locating in a specific area and only later deciding to drill. Instead, many concentrate on how upstream firms learn by doing, a concept spawned from a seminal paper that modeled dynamic production efficiency in the airline industry (Benkard 2000).

Contributing to this literature are Levitt (2009), Kellogg (2011), Covert (2015), Fitzgerald (2015), Seitlheko (2016), Steck (2018), and Fetter et al. (2018), who apply the same principles to study how firms become more proficient at producing oil and gas with individual and joint drilling experience, adoption of and passive experimentation with new technologies, and observing drilling and completion inputs used by competitors. Other important papers have studied how investment in drilling responds to changes in uncertainty (Kellogg 2014), and how oil production does not respond to oil prices but drilling activity and costs respond strongly (Anderson, Kellogg, and Salant 2018).

To highlight the fact that technological advancements and learning are not the only important considerations for well productivity and the future exploitation of unconventional resources, Montgomery and O'Sullivan (2017) find that previous studies that fail to properly disentangle the influence of technology from that of well location and targeted placement into favorable geology significantly overestimate the effect of technology. Using a case study of wells in the Williston Basin, they estimate that only half of the improvement in well productivity can be associated with technology changes, and much of the gains actually come from drilling in geological sweet spots.[105] In a study of the joint leasing and drilling decision made by firms in the Eagle Ford shale when the search for new leases is costly, Agerton (2017) models leasable acreage as a depletable resource, but treats the geology across acreage as perfectly understood— a large assumption. Of the previously mentioned studies, only Covert (2015) incorporates information on observable resource abundance,[106] but implicitly assumes that firms had prior knowledge of the geology before locating in an area, and only after lease acquisition and drilling do they gather further knowledge about the underlying geology and update drilling and completion techniques.

A commonality in all of these studies is the use of major simplifying assumptions about true geological quality and or firm beliefs (or knowledge) about geological quality. Homogeneous spatial distributions of oil and gas deposits was a common assumption that economists previously believed enabled them to infer how new drilling and well stimulation and completion technologies influence the proficiency of extraction. This assumption clearly does not hold when attributing contributions for gains in production proficiency, but it also does not hold during the firm location (i.e. mineral rights acquisition) process. Agerton (2019) builds on this idea and conducts another analysis about learning in the industry. Instead of modeling how firms learn how to drill, he models the process of how they learn *where* to drill. The analysis assumes that firms know the

---

105 Source: https://www.glossary.oilfield.slb.com/en/Terms/s/sweet_spot.aspx.

106 Specifically, he incorporates geological characteristics (i.e. thickness, total organic content, and thermal maturity) of the rock into which each well is drilled, using data from the North Dakota Geological Survey maps and GIS shape files published in 2008.

unconditional distribution of geology (not the realization at each location) and investigates how firms choose drilling locations based on knowledge about geological quality acquired during drilling. The findings echo Montgomery and O'Sullivan (2017), as naïve regressions failing to control for location overestimate the influence of technology by nearly threefold. Agerton (2017) also estimates a $1 billion welfare loss due to imperfect initial information on geology at the outset of drilling, and a correlation of firms' priors with true geological quality of 74%, although this correlation appears inflated due to an incorrect assumption regarding the calculation of the spatial distribution of geological quality used in the study.[107]

The learning-by-doing literature contributes significantly to the knowledgebase on drilling proficiency, and Montgomery and O'Sullivan (2017) and Agerton (2019) are convincing of the importance of disentangling improvements in production proficiency due to technological progress from firms drilling into areas with better geology. However, the findings of the latter two studies come with a few caveats as they do not adequately explain the full spectrum of information acquisition that occurs *before* firms decide to

---

[107] Agerton (2019) defines the geographic extent of Louisiana's Haynesville shale by following a University of Texas Bureau of Economic Geology study on the geological quality of the Haynesville shale. As he mentions, the study used data on the thickness and total organic content of the shale to calculate a measure of estimated "original gas in place" per square mile for a grid of one-mile squares for the whole play. However, he incorrectly states that this spatial distribution of geological quality was created solely using geological data, and not from well production (the study was conducted over 5 years after drilling began to occur and used production data over 2008-12 to create the distribution). Hence, the assumption that these estimates provide exogenous information on the spatial distribution of resource quality likely does not hold since the same production data was likely used by firms in their initial location decision. This implies the estimated 74% correlation between firm beliefs and true geology is inflated due to the model picking up much of the same production data used in the estimation of the distribution of geological quality. Meaning, it remains unclear how much of this information firms acquired during drilling and how much they knew previously.

locate (i.e. acquire mineral rights in an area), and what occurs *afterwards* from drilling. Agerton (2019) comes the closest, but admits the goal of the analysis is not to understand decisions made before drilling—a crucial stage in the process of developing unconventional hydrocarbons with large implications for the need to drill exploratory wells and the proficiency of production wells. Hence, the study is unable to accurately disentangle what firms knew about geology prior to locating in an area from what they learned while drilling. Instead, it only confirms previous notions, that location decisions rest on unobserved *prior* beliefs and private information about the quality of a drilling location, with a stated assumption that the terms on lease agreements and drilling and production activity perfectly reflect a firm's prior beliefs on the geological quality of a location.[108]

This paper contributes to the above literature by studying a new dimension for how upstream firms learn about geology before making location decisions. I show that it is not just private information that is important to firms when locating, but that they also use freely available information from government agencies to aid them in identifying optimal areas for lease acquisitions. These actions are intuitive given costly search and a lack of

[108] As found by Covert and Sweeney (2019), this assumption does not hold when comparing leases acquired via auction to those acquired through informal negotiations with private mineral rights owners. Over otherwise equal geology, firms in the Permian Basin pay larger bonuses for actioned leases with lower royalty rates and longer primary terms, relative to those negotiated informally. The wells drilled on these leases are also found to be more productive, presumably due to firms having more time to learn about geology and optimize drilling and completion inputs, along with the greater incentives to produce more oil and gas on leases with lower royalty rates.

familiarity with the geology during the early development of a new play, especially for smaller firms who have less flexibility to amass acreage and drill exploratory wells.[109]

3.3. Background

3.3.1. Oil and Gas Resource Assessments

3.3.1.1. Historically

Government-funded oil and gas assessments are not a new construct to the energy industry. In response to the oil embargo in 1973-74, the USGS Branch of Oil and Gas Resources (BOGR) was formed in 1974 with the primary objective to conduct scientific studies that provided a framework for objective National Oil and Gas Assessments (NOGAs) conducted at the U.S. geologic province level. Later, in 1984, the proprietary Nehring Significant U.S. Oil and Gas Field database came available and provided more granular play-level assessments for the first time.[110] After this innovation, the play-based approach became the basis for the 1988 and 1995 releases of the NOGA, which primarily focused on conventional oil and gas resources across all 70 geological provinces in the U.S. (Figure 10). The primary idea of NOGAs was to place U.S. resources in a global context alongside the results of the USGS World Oil and Gas Assessments.

In 1994, BOGR made its first steps to define and assess the unconventional resources of the U.S. – including shale oil, shale gas, tight oil, tight gas, and coalbed gas.

---

[109] To an extent, the results of Agerton (2019) are consistent with the land-rush idea, as the strategy of some firms is to acquire a portfolio of acreage and subsequently optimize on which areas to drill based on acquired geological knowledge from drilling. However, even with an initial land grab occurring, the rapid acquisition of acreage does not mean that firms discontinued their search or paid no attention at all to new opportunities to acquire prime acreage. In fact, it is likely that smaller firms needed to be more precise in when targeting acreage, given their financial constraints, and paid significant attention to new geological surveys.
[110] Source: http://www.nehringdatabase.com/index.html.

However, due to changes in program funding that caused staff reductions in 1995, BOGR, the Branch of Coal Geology, and the Branch of Sedimentary Processes were merged to form the USGS Energy Program, later renamed the USGS Energy Resources Program (ERP). In 2000, the new NOGA Project was launched with a focus on unconventional hydrocarbons after a quantitative methodology was developed to assess these resources. With fewer staff on hand, another major change to the program resulted in a rolling approach to the assessments. Instead of waiting until the end of a 3-5 year period to release the results for all 70 provinces, as done with past NOGAs, individual provinces were assessed in a prioritized manner with results announced shortly after.

As this program was underway, the Energy Policy and Conservation Act of 2000 was signed and required USGS (while working with the Bureau of Land Management, the Department of Energy, U.S. Forest Service, and the U.S. Fish and Wildlife Service), to assess the technically recoverable oil and gas resources under Federal lands in the priority provinces of the U.S. The Energy Policy Conservation Act Amendments law was later signed in 2005, and in order to "ensure accurate comparisons of geological resources," required USGS to use the same methodology for all assessments of unconventional resources in all areas.[111] Since the USGS is a bureau of the U.S. Department of the Interior, it also must respond to requests for oil and gas assessments by other agencies and by Congressional request.

[111] Source: https://www.dmr.nd.gov/downloads/USGSOilandGasAssessmentProject.pdf.

3.3.1.2. Current Efforts and Assessment Beneficiaries

Today, the ERP is part of the USGS Energy and Minerals Mission Area that conducts research and assessments on the location, quantity, and quality of mineral and energy resources in order to provide a better understanding of the economic, environmental, and human health impacts of resource extraction and use.[112] The federally funded program is the sole provider of reliable, impartial, and publicly available estimates of geological energy resources for the U.S., assessing resource potential through in-depth studies of the geologic, geophysical, and geochemical framework.[113] It also invests in innovation, from research to enable and improve assessments of current energy resources to understanding and assessing the potential for transformative new energy resources.

Over the past decade, many of the geologic provinces with significant unconventional resources have been re-assessed as new data become available. The results are often used to inform policymakers and support science-based decision-making regarding domestic and foreign energy resources, their responsible use, and to manage energy resources on Federal lands. They are also used by a variety of stakeholders, including local, State and Federal governments, land and resource management bureaus of the Department of the Interior, federal environmental and national security agencies, State geological surveys, the *energy industry*, and the public and environmental community.[114] Notably, the U.S. Energy Information Administration also uses them as the

---

[112] The geologic energy resources that the ERP studies are: oil, natural gas, coal, coalbed methane, gas hydrates, geothermal resources, uranium, oil shale, and bitumen and heavy oil.
[113] The 2019 budget was $25,879,000 (USGS 2019).
[114] Source: https://energy.usgs.gov/GeneralInfo/AbouttheEnergyProgram.aspx.

basis for computing reserves estimates for various basins in the U.S. and globally (USGS 2019). The results for each assessment dating back to 2000 are available on the ERP's Central Energy Resources Science Center's website.[115,116] Of interest to this study is the 2008 assessment of the Bakken shale formation within the Williston Basin.

3.3.2. Williston Basin Geology

The Williston Basin, located in parts of Montana, North Dakota, South Dakota, and Saskatchewan (Canada), is one of the premier unconventional oil basins in the world. It developed nearly two billion years ago and covers approximately 300,000 square miles (Figures 10 and 11) as it formed over millions of years when shallow seas deposited rich organic matter into the sediment when living organisms died. Over time, layers of sedimentary rocks over 16,000 feet deep accumulated near the center of the basin (the town of Williston is located near center, hence the name), and the enormous heat and pressure converted the organic matter into oil and gas.[117] The many layers of rock in the basin are grouped into various formations by geologists and are distinguishable by their age and, therefore, depth (Figure 12).

---

115 Source: https://www.usgs.gov/centers/cersc/science/united-states-assessments-undiscovered-oil-and-gas-resources?qt-science_center_objects=0#qt-science_center_objects.

116 Source: https://www.usgs.gov/energy-and-minerals/energy-resources-program/science.

117 Source: https://www.ndstudies.gov/gr8/content/unit-i-paleocene-1200-ad/lesson-1-changing-landscapes/topic-2-geology/section-2-williston-basin.

The Bakken and Three Forks formations constitute the most prominent oil-producing parts of the Williston Basin and sit roughly 10,000 feet below the surface.[118,119] The full spatial extent of the Bakken covers 200,000 square miles across the U.S. and Canada (Covert 2015). Within the Bakken are three distinct layers or "shale members" as referred to by USGS, which contain hydrocarbons: an upper member, middle member, and a lower member.[120] In descending order, each member is of lesser spatial extent, and it is understood that the upper and lower members (shale layers) are organic-rich and thus are the petroleum "source rocks" of the formation (Pollastro et al. 2008).[121]

The middle member of the Bakken varies more in thickness and, relative to the upper and lower members, has significantly less consistency in its geologic characteristics— lithology, thickness, and petrophysical properties (Pollastro et al. 2008). It is made of sandstone and dolomite, meaning it did not form from organic matter and therefore is not considered a source rock since it does not generate oil on its own. Instead, its porosity, which is conducive to petroleum migration from the upper and lower members, and its natural vertical and horizontal fractures enhance oil production in both conventional and unconventional reservoirs of the Bakken (Pollastro et al. 2008). These

---

[118] The Lodgepole formation is another low-porosity, but lesser oil-producing formation of the basin and overlays the Bakken formation, from which its oil migrates.
[119] They formed at different points in time when the seas expanded and deposited organic material into areas with low dissolved oxygen (anoxic z), but are separated by accumulations of sediment with little or no organic material occurring when the seas receded.
[120] Although it was not included in the 2008 assessment, there is also a fourth layer of the Bakken. The pronghorn member, located in-between the lower member and the Three Forks formation, is geologically and stratigraphically defined as part of the Bakken, and was later assessed with the Three Forks formation (Gaswirth et al. 2013).
[121] The term source rock refers to the rocks from which hydrocarbons have been generated or are capable of being generated. The source rock is the origin of hydrocarbons found in conventional reservoirs, and contains the many unconventional hydrocarbon molecules that remain trapped within the rock.

features make it an important component of the Bakken-Lodgepole total petroleum system (TPS), delineated in the 2008 assessment (Figures 11 and 13).[122]

3.3.3. Bakken Assessment Methodology and Results

USGS resource assessments are fundamentally based on geology, and use probabilistic and assessment methodologies as a means of quantifying geologic hypotheses and uncertainties (Charpentier and Cook 2011). In its 2008 assessment of the Bakken formation, USGS first delineated the Bakken-Lodgepole TPS, which involved mapping the distribution of source rocks and known petroleum accumulations and determining the timing of petroleum generation and its migration. Within the TPS, the area of the oil generation window (i.e. the area of thermally mature oil-generating source rocks) for the unconventional Bakken reservoir was then delineated by contouring both

---

[122] A TPS can be interpreted as a volumetric mapping of geology containing genetically related hydrocarbons. Formally, it includes those generated by a pod or closely related pods of mature source rock, and all components and processes necessary to generate and store hydrocarbons. In other words, within a basin, an individual TPS is defined by the volume of geology (i.e. across space and subsurface depth) containing molecularly related petroleum-generating rock and petroleum accumulations deposited in similar periods from similar sources of organic matter.
Sources: https://www.glossary.oilfield.slb.com/en/Terms/p/petroleum_system.aspx; and https://gsa.confex.com/gsa/2001ESP/finalprogram/abstract_7267.htm.

hydrogen index and well-log resistivity values of the upper shale member— the youngest and of greatest areal extent.[123,124]

Within this window, seven assessment units (AUs), or geological boundaries, were delineated— five unconventional AUs and two conventional AUs (Figure 13).[125] USGS defined the AU boundaries using a geologic model based on the thermal maturity and geographic extent of the source rocks, the matrix porosity (i.e. petrophysical character) of the middle sandstone member, and the natural vertical and horizontal fractures of the basin (Pollastro et al. 2008). After defining each boundary, the total undiscovered oil and gas resources of the six AUs in Figure 13 were quantitatively assessed. As outlined by

---

[123] Thermal maturity is an important phenomenon for the determination of prospective shale oil and gas accumulations at the initial stages of exploration. Based on thermal maturity values, source rocks are categorized into three groups.

i. Thermally *immature* source rocks are those buried at relatively shallower depths and typically have not been exposed to sufficient heat to convert organic matter to hydrocarbons, but may generate small amounts of natural gas as temperature and pressure increases.
ii. Thermally *mature* source rocks are those buried deeper and under a thicker overburden of rock, and have been exposed to temperatures and pressures high enough to generate crude oil and small amounts of natural gas. Source rocks in this state are said to be in the oil generation phase or "window." Rock buried just deeper than this will generate mostly natural gas due to increasingly higher temperatures, and accordingly will produce both oil and natural gas during extraction. At a certain depth and therefore temperature threshold, the "gas window" occurs, where *only* natural gas is generated, and any previously generated oil will also be transformed into natural gas.
iii. *Post-mature* source rocks are those at depths and temperatures that exhaust the hydrocarbon generation potential of the source rock, and hydrocarbon molecules break apart and are non-extractable.

Sources: in-person conversations with Vista Oil & Gas and https://infolupki.pgi.gov.pl/en/gas/thermal-maturity-organic-matter-and-gas-exploration.

[124] Essentially, the hydrogen index is an indicator of the thermal maturity of the rock (Covert 2015). Resistivity logging is another fundamental method used in formation evaluation, used to characterize the electric conductivity of fluids within a formation. Since hydrocarbons do not conduct electricity, resistivity values are used to distinguish whether the fluid within a formation contains hydrocarbons, water, and the proportions within. When used together, data generated by these metrics aid in the delineation of the spatial extent of hydrocarbon presence and abundance.

Source: https://www.glossary.oilfield.slb.com/en/Terms/r/resistivity_log.aspx.

[125] Note: one of the two conventional AUs is not shown in the figure, as it was not assessed. This AU was within the Lodgepole formation.

Pollastro et al. (2008), the methods used were based on the geologic elements of a TPS, including: (1) source rock distribution, thickness, organic richness, maturation, petroleum generation, and migration; (2) reservoir rock type (conventional or unconventional), distribution, and quality; and (3) character of traps and time of formation with respect to petroleum generation and migration. Detailed framework studies in stratigraphy, structural geology and the modeling of petroleum geochemistry, combined with historical exploration and production analyses, were also used to aid in the estimation of resource abundance for each AU.

The results of the assessment, published in April 2008, are provided in Table 8.[126,127] A total mean technically recoverable resource of 3.65 billion barrels of oil, 1.848 trillion cubic feet of natural gas, and 113 million barrels of natural gas liquids were estimated to reside within the area of the Bakken formation's oil generation window (AU1 – AU5). However, there is heterogeneity in the spatial distribution of resources. AU4 was estimated to have the most oil, natural gas, and natural gas liquids per square mile, followed by AU3, AU5, AU1, and AU2. Hence, the easternmost part of the Bakken was estimated to contain the most densely concentrated resources, and the AUs located in the north and the central-southwest portions of the oil-generating window were estimated to have a lower amount of resources per square mile. AU6, the only conventional AU

[126] Note: only mean estimates were provided for the conventional AU6.
[127] Available at: https://pubs.usgs.gov/fs/2008/3021/pdf/FS08-3021_508.pdf.

assessed and whose oil migrated from the oil generation window of the Bakken, was estimated to have relatively few total resources.[128]

Given that discretely defined areas of the Bakken that delineate where the majority of its oil resides, I treat these areas as the units of observation in this study. Meaning, within the Williston Basin, a TPS was defined that provided more certainty about the spatial extent to which Bakken formation oil resources are located. Relative to the unassessed areas of the basin, the assessment results for the six AUs provided a narrower window with more certainty about the spatial distributions of oil deposits within the Bakken TPS. My focus in this study is to understand how the development activity occurring within the AUs changed after the assessment results, relative to activity occurring outside of this area. In the next subsection, I describe my empirical approach used to identify an effect that the USGS assessments had on various oil and gas development outcomes.

3.4. Empirical Strategy

This paper examines the extent that the results of a USGS resource assessment published in April 2008 caused changes in oil and gas leasing activity— namely, the geographic location and the negotiated terms on new leases— in the Williston Basin. As indicated by Smith (2018a), prior to 2009, the development of the basin was still very much in its infancy as few wells had been drilled and firms were not well able to

[128] Note: under all of these areas lie additional hydrocarbons, residing in the Three Forks formation. Given the greater costs to drill and reach the deeper Three Forks formation, firms typically concentrated drilling efforts in the Bakken formation during the early stages of development (Covert 2015).

distinguish high potential from low potential drilling sites.[129] Instead, a conglomerate of small and medium-sized (at the time) independent firms in the region was primarily focused on transitioning drilling and completion inputs and techniques used in unconventional shale gas plays to those containing shale oil.[130] Given the expansive 300,000-square-mile extent of the basin, the publicly available assessment results provided a plausibly exogenous shock of information on the distribution of geology and resource abundance in the region, which I hypothesize firms used to narrow their search radius for acreage and fine-tune their beliefs on the potential "sweet spots."

Given as good as randomly assigned areas of the Bakken, delineated based on geology that was determined naturally over millions of years, I exploit the assessment results as a natural experiment using a research design as follows: I treat the five unconventional AU1-AU5 and the lone conventional AU6 as distinct areas. I also treat the area outside of the six AUs but still within the Williston Basin, as one distinct area. I then create monthly units of observation for a set of leasing activity outcomes in each of these seven areas of the basin. The primary outcomes of interest are a count for the number of new lease acquisitions, the average size of a new lease, the average royalty rate and bonus payment on a new lease, the average primary term length on a new lease, and others, for

[129] Smith (2018a) constructs a supply curve of the remaining shale oil reserves in the Bakken, and estimates that 50% of the remaining technically recoverable resources located in the formation can be economically produced if oil prices remain near $50 per barrel. His model uses historical drilling data and maximum likelihood estimates of the number and productivity of remaining drill sites. However, in estimation he omits wells from before 2009, as the selective sampling postulate used is likely not applicable in the early years, due to the infancy of development in the basin and firms needing more time to learn where the higher quality geology existed.

[130] Source: http://eprinc.org/2011/08/the-bakken-boom-an-introduction-to-north-dakotas-shale-oil/#sthash.GeV0sohX.dpbs.

each month in each of the seven areas (Figure 13). These outcomes were chosen as they enable me to study strategic positioning by firms on the landscape, and include the most commonly negotiated lease terms and those that economic theory would predict to constitute profit-maximizing firms strategies used during lease acquisition.[131] Since upstream firms are often valued based on the quality of their acreage, an intuitive development strategy would include amassing contiguous acreage over better geology, and negotiating longer primary terms or an option to extend this term, and lower royalty rates and bonus payments in order to maximize the expected value of leases.[132]

To analyze the effects of the assessment on leasing activity within the Williston Basin, I implement a difference-in-differences methodology using the design above. My estimating equation is:

$$Y_{it} = \beta_0 + \sum^{6}{}_{i=1} \delta_k \mathrm{Post_t} * AU_i + \boldsymbol{\beta X}_{kt} + \lambda_t + \gamma_i + \varepsilon_{it}, \tag{3.1}$$

where Y_{it} is one of the value for development outcome *k* of area *i* in month *t*, and $Post_t \ * \ 2008AU_k$ is an interaction term between an indicator for the time the assessment results were announced (April 2008) and an indicator for assessment unit *i* (AU_i). The parameters

[131] McFarland (2014) describes in detail the basic or "essential" lease terms and others that should be considered by a mineral rights owner in order to negotiate a reasonable and fair lease. The online appendix for Vissing (2018) also outlines these terms and many other clauses, or addendums that appear on leases and have important implications for profitability. It is apparent that there is a positive correlation between the occurrence of these clauses (i.e. more lessor-favorable lease agreements), bargaining experience, lessors that do their due diligence before signing a lease, and or lessors that consult with an oil and gas attorney prior to agreeing to terms, as recommended by McFarland (2014).

[132] Upstream oil and gas firms are commonly valued based on their primary assets— proven oil and gas *reserves* (Gold 2014, pp. 129), which are hydrocarbon deposits that can be extracted under current prices and existing technologies. Reserves are different from *resources*, defined as all hydrocarbons that can technically be recovered at *any* price. These terms are commonly misinterpreted, and in this paper, I assume that the amount of resources in an area has a close correlation with the amount of reserves, or at least provides a signal of resource abundance to firms.

on the interaction terms are of primary interest, and estimate how development outcome k in AU i changed relative to the areas of the Williston Basin that were not assessed. I also include a set of controls for each area, $\boldsymbol{X}_{kt}$, and year-month and area fixed effects, λ_t and γ_k, which are included to control for time trends common to all areas (e.g. oil price) and trends that are specific to each area (e.g. average differences in leasing activity, local geology, and other area-specific differences), respectively. Post assessment, I predict leasing activity to increase in AUs that were estimated to contain more oil and gas, as geology and resource availability became more certain, with the largest increases in activity occurring in the areas estimated to contain the most oil and gas.[133] The identifying assumption is that absent the assessment, leasing activity outcomes in the assessed area would have trended similarly to the activity in unassessed areas. Figures showing these trends for each outcome are available at the end of the paper, and are discussed in the next subsection.

3.5. Data

3.5.1. Data Hurdles and Manipulations

I use leasing data from DrillingInfo, which records the universe of mineral rights transactions filed in county registries of deeds. I first restrict this data set to include lease transactions observed in Montana, North Dakota, and South Dakota over 2000-2017. In ArcMap, I then layer the grid coordinates for each lease with shape files for the Williston

[133] Although I do not analyze it in this paper, theory might also predict increases in the number of permit applications and drilling activity, as firms who already had acreage under asset in these areas and saw the results as positive news may have shifted forward their development plans to explore a new opportunity.

Basin and each AU to create indicators for whether the lease was located within the Williston Basin, and whether it was located within each of the six AU areas (or none at all) from the 2008 assessment.[134] For the purposes of this analysis, I also restrict the sample to only leases located within the Williston Basin, where the assessment occurred and where the majority of development activity was occurring in the region.

After dropping all observations located outside of the Williston Basin, roughly 500,000 observations of lease instruments remained, although many of these were duplicate records, occurring for a variety of reasons including fractioned leases (in such cases a record existed for each fractioned owner of mineral rights), but mostly due to a lack of specificity with the lease location.[135] Given that I only had access to a basic DrillingInfo subscription at the time of this analysis, I did not have a unique lease identifier that enabled me to easily remove duplicates.[136] To circumvent this issue, I sort the leasing data based on lease instrument type, lessor name, the number of acres, and year-month, and use this combination of characteristics to generate a (still fuzzy) identifier of each lease transaction, and then remove duplicate records for the same lease. This dataset also contained records for eight different types of lease instruments, including (in a descending order of frequency): lease; lease memorandum, lease extension, lease amendment, unknown transaction type, lease ratification, lease option, release of lease, and seismic

---

[134] Shape files available for download at https://certmapper.cr.usgs.gov/data/apps/noga-data/?provcode=5031.

[135] The DrillingInfo subscription I had access to at the time of this analysis only enabled me to observe the centroid of the abstract in which the lease was located. I did not have address information of the lessee, so I could only differentiate leases within the same abstract using the name of the lessee, instrument date, and other lease terms and characteristics.

[136] In future research I will look to obtain a higher-tiered DrillingInfo subscription.

memorandum. Since this study concerns leases and the terms on leases, I then restrict the sample by removing all non-lease transactions, leaving 207,820 unique leases over 2000-2017.[137]

Lastly, the basin was assessed again in April 2013, and included a re-assessment of the Bakken and an assessment of the Three Forks formation.[138] Since my goal is to study the changes in leasing activity occurring due to the 2008 assessment, I omit lease observations from after March 2013 because in the time between assessments, over 4,000 wells were drilled causing some depletion in the Bakken and enabling additional learning about geology (Gaswirth et al. 2013). Hence, leasing and other development behavior likely changed after this period due to many other reasons that complicate this analysis.

3.5.2. Outcomes of Interest on Leases

Typical oil and gas leases contain several items determined during negotiation between lessor (mineral rights owner) and lessee (firm). The first is the royalty rate, which pays the lessor a percentage of production throughout the life of the producing well(s) located on the lease. This rate is typically between a minimum of 12.5% and up to as high as 25%, although there is no upper limit. Another item is the primary term, which specifies

[137] One caveat here is that over time, firms have learned how to disclose less information during leasing and other processes. Evidence is apparent in the increasing number of firms who engage in private negotiations and acquire mineral rights, but submit a lease memorandum instead of an official lease. This is advantageous for firms because in a lease memorandum, they are only required to disclose the location of the lease (lessor and lessee and their addresses), the primary term, lease extension terms (if any), and a description of the leased premises. It is a bare bones document that allows them (legally) to be strategic since they can keep private the financial terms of the lease. This is strategic since they can be used as leverage in subsequent negotiations if obtained nearby mineral rights owners, but also so firms do not have to deal with angry mineral rights owners who previously signed leases with less favorable terms. Since these lease memorandums do not have many of the outcomes I am interested in, I do not include them in the analysis. Source: https://www.daily-jeff.com/article/20140805/OPINION/308059394.

[138] Details on the 2013 assessment results are available in Gaswirth et al. (2013).

a fixed amount of time (years or months) during which the lessee need do nothing in order to keep the lease in effect. If no exploration or production occurs during this period, the lease will terminate unless the lessee had also negotiated an option for a lease extension and subsequent extension bonus payment. If production occurs, the lease is held in force so long as production is occurring.[139,140]

Other terms include a bonus payment, which is effectively a front-end lump-sum payment made to the mineral rights owner for signing the lease, and additional terms regarding the extension of the primary term on a lease if the lessee fails to explore during the first term. DrillingInfo records whether an option to extend the primary term was negotiated for each lease and, for leases with this option; it records the extension term length as well as the extension term bonus payment (another lump-sum payment). Although not recorded by DrillingInfo, other addendums on leases have also begun to appear more frequently on leases, including non-development clauses, freshwater supply and wastewater disposal provision options, and others (see online appendix for Vissing 2018). I do not have access to any of these in my data set.

Lastly, I create several explanatory variables. As Gold (2014) describes, a common strategy firms use during development stages is to observe the locations and drilling activity of competing firms, especially the larger firms who have more resources, better

[139] According to conversations with a Ron Kaiser, at Texas A&M University, there is sometimes a gray area between what was specified in the lease and what constitutes production "activity." Many leases specify that production must occur for a lease to remain in effect, but firms have been known to maneuver around this, sometimes claiming drilling a well, preparing a well pad, or constructing a road to a proposed well pad location constitute a sufficient production-related activity to hold the lease.
[140] See Smith (2018b) and Herrnstadt et al. (2019) for more information on firm incentives related to expiring primary terms.

technology, and who are presumed to be more sophisticated explorers and better at finding the best places to drill. Paying close attention to publicly reported data, as well as using spies that are frequently sent out, helps firms with fewer resources gather inexpensive intel about the behaviors of larger firms and provides opportunities to 'piggy-back' on their exploration efforts by acquiring mineral rights nearby (Gold 2014). Echoing this empirically is Levitt (2016), who finds that firms pay close attention to drilling histories of other firms, both in terms of making the initial location decision and subsequent drilling decisions. In my model, it is important to control for this type of activity. I classify large firms based on the total acres acquired in all leases negotiated during my sample period. Of the 1,668 firms with lease records in my data set, *five* of them acquired 31.5% of all acreage, ten of them acquired 42.6% of all acreage, and *fifteen* acquired 50.2% of all acreage in leases signed. I then create a variable for the total number of leases acquired by the top ten largest operators in each area in each month, and five lagged terms to use as controls for this type of leader-follower behavior.

3.5.3. Intuition on Development and Negotiation Strategies of Firms

Amassing contiguous acreage over better geology is valuable to firms due to the economies of scale associated with spatially proximate acreage (Vissing 2018), and the capital at stake over such acreage is large.[141] Negotiating a longer primary term, or

---

[141] In the recent bidding war between Chevron and Occidental Petroleum for Anadarko Petroleum, each firm is interested in the prime, contiguous acreage that Anadarko has in the Permian Basin in Texas. It has led to bids reaching nearly $60 billion, which would constitute the fourth largest acquisition in the history of the oil and gas industry.
Source: https://www.forbes.com/sites/davidblackmon/2019/04/25/7-more-things-you-need-to-know-about-oxys-new-bid-for-anadarko/#75425d2c2ba8.

alternatively negotiating an option to extend the length of the primary term, is also desirable for a firm as it provides more time to learn about the geology in the area and optimize inputs before making a drilling decision (Kellogg 2014; Agerton 2019; Herrnstadt et al. 2019). Lower royalty rates and bonus payments are desirable, although negotiating favorable outcomes on both of these terms on the same lease is more difficult, and there are unique tradeoffs that occur between the primary term, royalty rate, and bonus payment on lease agreements. Hence, intuitively a firm would prefer to pay a larger upfront bonus to the mineral rights owner (many of which may have a high discount rate and value more the upfront money), and negotiate a lower royalty rate— entitling it to a larger share of the expected surplus over the life of a productive lease. Recent evidence of this and the value firms place on longer primary terms and lower royalty rates comes from the Federal Bureau of Land Management's third-quarter auction of oil and gas leases in September 2018, which broke several records for bonus payments.[142]

3.5.4. Final Data Set and Trends

I collapse the above dataset to construct monthly values for each variable in each of the seven defined areas within the Williston Basin over 2000 through March 2013. Trend figures for these variables are available in the Appendix A, but note they cover the

---

[142] The two-day sale saw Matador Resources pay a record $95,001 per-acre for drilling rights, and total sales from the auction generated a record gross revenue of nearly $1 billion for leases on the New Mexico side of the Permian Basin, which are granted with a low 12.5% royalty rate and a long ten-year primary term. Sources: https://www.bloomberg.com/news/articles/2018-09-13/matador-revealed-as-mystery-buyer-of-record-permian-acreage-sale.; https://opportune.com/Energy-Sector-Insights-Events/Insights/What-Does-Record-BLM-Oil-Gas-Lease-Auction-Mean-For-Valuation/; and https://www.forbes.com/sites/davidblackmon/2018/09/10/the-rest-of-the-story-on-new-mexicos-record-permian-basin-lease-auction/#52cc22f0786d.

full sample period through 2017 to show all changes occurring. Figure 14 shows how the number of new leases per month changed over time. There is heterogeneity in the pre-assessment trends across areas, along with some pre-assessment period divergence in AU3, AU4, and the unassessed area. However, as expected, a spike occurs in 2008 for all areas (including the non-assessed area). Figure 15 shows how the average size of a new lease changed over time for each area. There is significant heterogeneity in the magnitudes, but the pre-assessment trends for all AUs track the unassessed area reasonably well after 2005 when there was a large spike. Hence, the parallel trends and the spike in 2008 for all AUs (except AU2 and AU6) provide important visual evidence that firms responded to the assessment by acquiring new and larger leases in areas with better geology and more certain resource abundance.[143]

Figures 16, 17, and 18 show time trends that reflect lease terms growing increasingly in favor of the lessor.[144] Figure 16 shows similar pre-assessment trends for the average primary term length in all AUs; and they track the trend in the unassessed area well. The figure also shows a declining average term length for all areas over the sample period (at a faster rate for the unconventional AUs after 2008), but term lengths were

[143] The increases in the average lease size in certain areas indicate that large plots of unleased acreage were available before the assessment. A downward trend thereafter is expected, as unleased acreage is a finite resource that is depleted as more leases are acquired and subsequently held by production as drilling occurs.
[144] A caveat of Figures 17 and 18 is that they do not reflect the terms from all mineral rights acquisitions over this sample period. As it became common to report lease memorandums instead of regular lease documents, over time an increasingly large portion of new acreage acquisitions did not have financial terms available, meaning that the monthly averages are biased to the extent of selection between firms that report lease memorandums and those that report regular lease documents. I do not yet have statistics on the proportion of lease transactions that were leases vs. lease memorandums, but this may be an important variable to control for in my model.

persistently the shortest for the two AUs estimated to have the most resources (AU3 and AU4).[145] Figure 17 shows that the average royalty rate increased over time for all areas (at a slower rate for AU6 and the unassessed area after 2008)— likely reflecting increasing competition among firms in the leasing market, although some could be due to better bargaining on behalf of mineral rights owners. The pre-assessment trends all reasonably track the unassessed area except for AU1 and AU5. Figure 18 shows the average bonus payment over time in each area, and shows parallel, albeit near-zero pre-assessment trends for all areas. It also shows a large spike for four of the areas in 2008, with two of the largest increases occurring in AU4 and AU5 (each within the top three of estimated resource abundance).

Figures 19, 20, and 21 show trends for terms on leases with an option to extend the primary term. Although the average primary term length decreased over time, it appears that the shorter primary terms were complemented by an option to extend this term for a portion of leases. This is apparent in Figure 19, as the pre-assessment trends for the proportion of leases in AUs that contained an extension option trended similarly to the proportion in the unassessed area, but increased in 2008 and several years thereafter. Figures 20 and 21 show trends for the average extension term length and extension bonus payment for leases with an extension option. Each figure shows relatively similar pre-assessment trends between AUs and the unassessed area, and all areas saw increases of

[145] This indicates that firms' expectations over production may have become more certain in these areas as they were willing to negotiate leases with terms and drill and produce hydrocarbons sooner in order to hold a lease.

different magnitudes in the average extension term length and bonus payment after the assessment.[146]

3.6. Results

I estimate the effect of the 2008 USGS assessment on eight different leasing activity outcomes in the Williston Basin. One of these outcomes indicates the direct effect of the assessment on demand: the number of new leases per month in each area. The other seven arise out of the bargaining process. The royalty rate, bonus payment and extension bonus payment all indicate the amount of money (and potential amounts of money) transferred from the firm to the mineral rights owner. Holding all else equal, the landowner prefers higher amounts in each of these. However, there are also trade-offs, as a landowner may prefer a lower bonus payment if offered a higher royalty rate or vice versa.

For several of the remaining four terms of the contract, it is less obvious whether the landowner would prefer a higher value or not. The size of a new lease indicates the number of acres below which the mineral rights owner has agreed to allow exploration and or production of hydrocarbons by the lessor. In general, agreeing to lease a larger plot of acres is beneficial to the landowner as firms may be willing to pay larger bonuses (and or royalty rates), but also due to a quantity effect from the contribution to royalty revenues of production on additional acreage. However, landowners also have the discretion to negotiate exploration restrictions, for example if they were to wish not to have horizontal

[146] One caveat of these trend figures is that they are at an annual frequency, which I use for better visual clarity in each figure.

wellbores drilled under certain parts of their property, in which case would reduce the total number of acres on the lease and may depress the amount of money transfers they receive.

The option to extend a lease beyond the primary term and the length of extension term are each likely to be correlated with the bargaining ability of the mineral rights owner. An option to extend the primary term might be preferable to in cases where a large extension bonus payment is offered and grants the firm additional time to drill a more proficient well, thereby increasing royalty revenues relative to what would accrue from a less proficient well. Alternatively, the option to extend the lease and a longer negotiated term of the extension might be less desirable if the lessor holds the drilling rights for the duration of the term without drilling, and there was another firm willing to drill and provide royalty payments sooner.

In Tables 9 through 12, the results are presented for each of the leasing activity outcomes. Each table includes the results for two outcomes under two different specifications, and each specification includes six interaction terms (Post x AU): one for each AU dummy multiplied by a post period dummy equal to one for the period April 2008 through the end of the sample, and zero before. The coefficients on these terms represent the effects of the assessment on the leasing activity outcome in each AU, relative to the non-assessed area. The amount of oil resources (MMBO per square mile) estimated for each assessed area is also included in parentheses next to each interaction term. Each specification also includes two sets of fixed effects, one for the areas of the research design and one for all year-months in the sample. The only difference between the two specifications is that the first estimates equation (3.1) without controls; the second includes five lags for the total number of acres acquired by the top ten largest firms. The following paragraphs describe the results for each outcome in order of relevance.

Table 9 presents the results showing the effect of the assessments on the total number of leases and the average size per new lease. These results provide intuition about where changes in leasing activity occurred in the Williston Basin after the assessment, and insight on how firms in the region appeared to believe the results. The first and third columns show the estimates for the specifications without controls. For brevity, I only discuss the results for the specifications that include controls for the lagged number of acres acquired by large operators (columns two and four). In column two, the assessment had a statistically significant effect on the number of new leases acquired in all areas, yet the effects were negative for three of the AUs, particularly conventional AU6, which had the fewest estimated resources, and AU1 and AU4, the latter of which was estimated to have the greatest resource abundance.

The result for AU4 is intuitive when combined with the result in column four, which shows a highly significant and large increase in the average size of a new lease post assessment in this area. When taken together, it appears that fewer leases were acquired in AU4, but the leases that were acquired were significantly larger than the leases acquired elsewhere, indicating that lots of unleased acreage existed in this area prior to the results of the assessment, which changed quickly after the assessment results were announced. To interpret this parameter, the assessment caused an increase in the average size of a new lease by 161.85 acres (nearly 40% when estimated with a natural log functional form), relative to the average size of a new lease in the unassessed area. This increase is more than three times as large as the second largest effect in AU5, which experienced a 50.99-acre increase in the average size of new leases. AU3, the area estimated to have the second greatest resource abundance, saw the largest increase in the number of new leases acquired and the third largest effect on the average size of a new lease, followed by AU1 and AU2. Conventional AU6 had a negative, but insignificant change in lease size.

Table 10 shows the effects of the assessment on the average primary term length and the average royalty rate. In column two, the assessment caused relatively small increases of roughly one to four months in AU1, AU2, AU3, and AU5, no statistically significant change in the term length for AU4, and a negative change in the length for AU6. For comparison, the average primary term length across all areas was just over 62 months prior to the assessment. The signs and magnitudes of these effects are somewhat unintuitive as I anticipated seeing a decline in the average term length after the assessment, since firms had more certainty relative to the unassessed area. Although firms may have been able to negotiate better terms during the early development of the basin before learning about leasing occurred by mineral rights owners, and providing them more time to study the geology in this area before drilling.

In column four, the assessment had a highly statistically significant effect on the average royalty rate in all AUs except AU5. For comparison, the average royalty rate was 17.56% for all assessed areas prior to the assessment, and 15.55% for the unassessed area. The effect on the average royalty rate in AU1 is negative, meaning that the assessment nudged two lease terms in this area in a direction favorable to the industry. Two of the AUs with the most estimated resources saw the largest change in the average royalty rate— AU3 and AU4, the former of which also saw the largest increase in the number of leases transacted (shown in column two of Table 9). The large effects here are likely due to firms willing to pay a slight premium for the additional resource certainty from the assessment, but these findings of pass-through of geological quality into royalty rates to mineral rights owners runs counterintuitive to the findings of Brown et al. (2016), who find no such evidence of pass through. AU6 and AU2 had the next highest changes in average royalty rate, but it is unintuitive why AU6 had a positive change.

Table 11 shows the effects of the assessment on the average bonus payment on new leases (across all areas, the average bonus payment was $346 per acre prior to the assessment), and the proportion of leases that contain the option to extend the primary term (roughly, 11% of leases had an extension option before the assessment). In column two, I find statistically significant and negative changes in the average bonus payment in AU1 and AU6, but large and meaningful increases in each of the other four AUs. AU4 had the largest increase at $5,828.02 per acre post assessment, followed by AU3 at $3,420.01 per acre, AU2 at $2,453.84 per acre, and AU5 at $562.89 per acre. The magnitudes of these changes are consistent with the estimated resource abundance in AU4 and AU3, but less so for AU2, estimated to have the second lowest resource density. The results are also consistent with those in Table 9, as a premium is paid for larger leases (i.e. those with more acreage), although the larger leases in AU5 only saw larger bonus payments, but not royalty rates. In column four, there was a decline of between four and eight percent in the proportion of leases with an extension option after the assessment in three of the AUs, an increase in two AUs, and no statistically significant change in AU1.

With the exception of the large increase in AU6, these results are somewhat surprising, although firms may have become less concerned with extension terms if they acquired acreage in areas perceived to have more certain production and or they negotiated longer primary terms.

Table 12 presents the effects of the assessment on the average extension term length and extension bonus payment for only leases containing an extension option. Prior to the assessment, the average extension term was just over eighteen months and the average extension bonus payment was $30.44 per acre. In column two, the average extension term declined by just by between 1.5 to four months for AUs 1-4, but increases by 2.28 and 6.02 months in AU5 and AU6, respectively. In column four, the average extension bonus payment increased in five areas after the assessment, with increases ranging from just under $23.87 and $25.85 per acre in AU1 and AU5 to nearly $140 per acre in AU4. Although they are not of greatest priority in this paper, changes in the extension terms are helpful to understand how additional dimensions of the bargaining process evolved as the leasing market matured.

3.7. Discussion

An increasingly large literature in economics has studied how information contributes to oil and gas development outcomes. Regarding exploration and production, previous research has concentrated on how information is acquired with experience and is used to inform future drilling location and input decisions (e.g. Levitt 2009, Kellogg 2011, Covert 2015, Levitt 2016, Steck 2018, and Agerton 2019).[147] Vissing (2017; 2018) has studied how firms have higher valuations for contiguous acreage and finds that those with geographic concentrations use it as bargaining power to obtain more outcomes during leasing activities. Other works in this area has found that firms are able to negotiate lease terms entitling them to more of the surplus generated from production when negotiating with mineral rights owners of lower socio-economic status, and across racial and ethnic groups (Vissing 2015; Timmins and Vissing 2017).

This paper contributes to the above studies on the importance of knowledge in the oil and gas industry in two primary ways. First, I show that it is not only private information that firms pay attention to, but freely available public information as well, which they used to acquire strategic acreage during early development of the Williston Basin. Second, the findings suggest several important welfare and bargaining implications. The results show that mineral rights owners in two of the AUs estimated to contain the most abundant resources were compensated with better royalty rates (AU4 and AU3), which were also accompanied by higher lump-sum bonus payments per acre (Table 11). However, this was not the case for all areas assessed in 2008. AU5 saw large increases in the number of new leases and the size of new leases after the assessment, yet mineral right owners in this area only saw larger bones payments, but not higher royalty rates (see

[147] Other seminal research prior to these studies investigated how information and peer effects from competing firms in the oil and gas industry aid strategic development activity, particularly in offshore lease acquisition and drilling activity (e.g. Hendricks et al. 1987; Paddock et al. 1988; Hendricks and Porter 1988; Hendricks and Kovenock 1989; Hendricks and Porter 1992; Hendricks and Porter 1993; Hendricks et al. 1993; Hurn and Wright 1994; Porter 1995; Hendricks and Porter 1996; and Lin 2009; 2013).

column four in Table 10 and column two in Table 11).[148] Similarly, the assessment led to more industry favorable terms in AU1— longer average primary terms, lower royalty rates, and lower bonus payments— yet AU1 was estimated to have the fourth best resource abundance of the six areas that were assessed. Hence, mineral rights owners were less well off than those in the other areas after the assessment.

This paper has several shortcomings, including a need for more resolution in the data and more robustness checks.[149] One direction for future research is to implement a spatial regression discontinuity design to estimate a local average treatment effect using the outcomes of leases near the borders of assessed areas. This approach would enable a comparison of leases just across a threshold where the geology should still be constant, which appears to be true given the 2013 assessment of the Williston Basin involved a re-assignment (i.e. slight changes) of the borders of the AUs. Additionally, there are large firm valuation and financial market implications associated with USGS assessments, as they plausibly enable firms to increase value and improve profitability by doing little more

---

[148] Note: it could be the case that the Fort Berthold Reservation is located within AU5. If this is the case, these findings provide empirical evidence that the leases they signed were less favorable than those in other areas were. Alternatively, more state lands could be in this area than other areas. As Leonard and Parker (2019) show, private ownership generated more oil production than government ownership over 2005-2015 in the Bakken shale. They also find that scattered government holdings within private areas further reduced production. Although I have not yet connect production to my findings, it could be the case that government-owned leases generated less lease revenue, or vice versa as found in Texas by Covert and Sweeney (2019).

[149] A significant caveat is that there are some discrepancies across outcomes when testing the parallel trends assumption. In other words, when testing for pre-divergence in outcomes before the assessment results were announced in April 2008, there is a statistically significant difference between assessed and unassessed areas for several leading terms (or months) for several leasing outcomes. This may indicate that firms had knowledge of the assessment results prior to their announcement to the rest of the public, and or that the leasing market was responding to other signals before the assessment. Future research will consider quarterly or annual frequencies, since lease negotiations may take longer than one month to finalize.

than acquiring better acreage. Lastly, the findings both echo and conflict with those of Brown et al. (2016), who study the relationship between royalty rates and the geological quality below the property for which they were signed.[150] Future research in this area should aim to better understand the bargaining implications hinted at in this paper, and how firms and mineral rights owners learn to negotiate (e.g. Brehm and Lewis 2019). This is because firms who are earliest to negotiate and sign leases in the areas with more resources are more likely to be able to negotiate better terms on their leases, as the leasing market may not have yet fully responded or adjusted to the new information on the resources in these areas. As shown in the regressions regarding the financial terms on leases, some pass-through appears to have occurred in the development of the Williston Basin. Although it is apparent that asymmetric information exists in this market as well, and due to the USGS assessment, firms were able to negotiate leases entitling them to a larger share of the surplus on new development in certain areas after the assessment.

3.8. Conclusion

USGS assessments of oil and gas resources are highly relevant to energy policy, especially as recovery-enhancing techniques such as horizontal drilling and hydraulic fracturing have become widespread in the U.S. and increased the potential exposure for a variety of spillovers from oil and gas activities by expanding the area of development. These technologies, combined with better information supplied by the assessments likely

[150] They find no evidence of Ricardian rents– i.e. larger rents for mineral rights owners in resource-abundant areas than those in less abundant areas. Instead, they find limited pass-through of resource abundance into royalty rates, and conclude that mineral right owners benefit from resource abundance through a production quantity effect, not through negotiating better lease terms.

enable more efficient wells to be drilled that have less production uncertainty. The results of this paper indicate that during the early stages of the development of a new play; firms may have less knowledge on geology than previously believed and external information helps them to strategically shift and or narrow the radius of their search for favorable geology. This is intuitive for at least two reasons. First, the acquisition of accurate geologic knowledge and information is time intensive and costly, and there had been minimal exploration occurring in the Williston Basin prior to the boom— meaning, there was relatively little information available to base location decisions on without spending on exploration. Second, the key driver of the boom in this region was a collection of many small, independent firms who had less sophistication and fewer resources than larger firms who entered late and essentially needed to use all external information available.

Evidence of this comes from a 2008 Executive Summary by the North Dakota Geological Survey on the Bakken Formation Reserve Estimates, which describes the lack of information, resources, and ill familiarity with the resource that firms had who entered early into the development of the Bakken:

> *"Drilling results with the increase in oil prices have attracted new companies into the Williston Basin. Existing and new companies commonly have staff with limited or no knowledge of the play. These companies rely on information from the current literature and the state government to quickly answer their questions."* (pp. 3).[151]

[151] Source: https://www.dmr.nd.gov/ndgs/bakken/newpostings/07272006_BakkenReserveEstimates.pdf.

Once firms understand the broader areas, where resources are more abundant and geology is more favorable, additional private exploration, akin to what is studied by Agerton (2019), is useful to understand the geology more locally. Due to the competitive nature of the upstream oil and gas industry, it places a large precedence on finding competitive advantages and maintaining the proprietary nature of its development activities, and for good reason. In the flavor of a first-mover game, Gold (2014) informally describes this behavior. Out of a desire to avoid sub-optimal completion of a well, he discusses how firms attempt to acquire every bit of information about a prospective play and its geology, before making a drilling decision. Historically, this involved the standard industry practice of studying the characteristics of legacy wells, including production logs, gathering knowledge on targeted formations and associated depths.

Although not studied in this paper, firms with new or existing leases in these areas are also plausibly more likely to begin the next stage of development, which might entail applying for permits to drill new wells and conducting additional exploratory research. In an extension of this work, I would expect areas with larger resource estimates to see an increase in the number of permit applications in the immediate periods following the publication of the survey results. Assuming that the permit application process is quick and successful, I also expect to see an increase in the number of wells drilled in these areas shortly after, as well as an increase in production. The new production could be due to more wells drilled, but also from an increase in the productivity of wells if they are drilled in areas with a higher concentration of oil and gas, or those areas consistent with the results of the USGS assessments.

References

Agerton, M. 2017. "Anatomy of a Shale Boom: Optimal Leasing and Drilling with Costly Search." Working Paper. Available at: http://markagerton.com/files/Agerton_AnatomyOfAShaleBoom.pdf. [Accessed 7/18/18].

Agerton, M. 2019. "Learning Where to Drill: Drilling Decisions and Geological Quality in the Haynesville Shale." Working Paper. Available at: http://markagerton.com/files/Agerton_LearningWhereToDrill.pdf. [Accessed 4/14/19].

Allison, E., and B. Mandler. 2018. "Subsurface Data in the Oil and Gas Industry." American Geosciences Institute Critical Issues Program - Petroleum and the Environment, Part 23/24. Available at: https://www.americangeosciences.org/sites/default/files/AGI_PE_Subsurface_web_final.pdf. [Accessed 5/1/19].

Anderson, S.T., R. Kellogg, and S.T. Salant. 2018. "Hotelling Under Pressure." *Journal of Political Economy*, 126(3): 984-1026.

Benkard, C.L. 2000. "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review*, 90(4): 1034-1054.

Brehm, P.A., and E. Lewis. 2019. "To Trade or Not to Trade: Oil Leases, Information Asymmetry, and Coase." Working Paper. Available at: http://www2.oberlin.edu/faculty/pbrehm/BrehmLewis_WY_Leasing.pdf. [Accessed 5/7/19].

Charpentier, R.R., and T.A. Cook. 2011, "USGS Methodology for Assessing Continuous Petroleum Resources." U.S. Geological Survey Open File Report 2011–1167. Available at: https://pubs.usgs.gov/of/2011/1167/. [Accessed 10/13/18].

Covert, T.R. 2015. "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale." SSRN Scholarly Paper ID 2481321. Rochester, NY: Social Science Research Network. Available at: http://papers.ssrn.com/abstract=2481321 [Accessed: 7/10/18].

Covert, T.R., and R.L. Sweeney. 2019. "Relinquishing Riches: Auctions vs Informal Negotiations in Texas Oil and Gas Leasing." Energy Policy Institute at the University of Chicago (EPIC) Working Paper NO. 2019-52. Available at: https://home.uchicago.edu/~tcovert/webfiles/cs_texas.pdf. [Accessed 3/29/19].

Fetter, T.R., A.L. Steck, C. Timmins, and D. Wrenn. 2018. "Learning by Viewing? Social Learning, Regulatory Disclosure, and Firm Productivity in Shale Gas." NBER Working Paper No. 25401. Available at: https://www.nber.org/papers/w25401. [Accessed 4/10/19].

Fitzgerald, T. 2015. "Experiential Gains with a New Technology: An Empirical Investigation of Hydraulic Fracturing." *Agricultural and Resource Economics Review*, 44(2): 83-105.

Gaswirth, S.B., K.R., Marra, T.A., Cook, R.R., Charpentier, D.L., Gautier, D.K., Higley, T.R., Klett, M.D., Lewan, P.G., Lillis, C.J., Schenk, M.E., Tennyson, and K.J. Whidden. 2013. "Assessment of Undiscovered Oil Resources in the Bakken and Three Forks Formations, Williston Basin Province, Montana, North Dakota, and South Dakota, 2013." U.S. Geological Survey Fact Sheet 2013–3013, 4 p. Available at: https://pubs.usgs.gov/fs/2013/3013/. [Accessed 4/15/18].

Gold, R. 2014. *The Boom*, 1st. ed. New York: Simon & Schuster.

Hendricks, K. and D. Kovenock 1989. "Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration." *RAND Journal of Economics* 20 (1989), 164-182.

Hendricks, K., and R.H. Porter. 1988. "An Empirical Study of an Auction with Asymmetric Information." *American Economic Review*, 78(5): 865-883.

Hendricks, K., and R.H. Porter. 1992. Hendricks, K., and R. Porter, "Joint Bidding in Federal OCS Auctions." *American Economic Review Papers and Proceedings*, 82: 506-511.

Hendricks, K., and R.H. Porter. 1993. "Determinants of the Timing and Incidence of Exploratory Drilling on Offshore Wildcat Tracts." NBER Working Paper 4605. Available at: https://www.nber.org/papers/w4605. [Accessed 6/13/19].

Hendricks, K., and R.H. Porter. 1996. "The Timing and Incidence of Exploratory Drilling on Offshore Wildcat Tracts." *American Economic Review*, 86(3): 388-407.

Hendricks, K., R.H. Porter, and B. Boudreau. 1987. "Information, Returns, and Bidding Behavior in OCS Auctions: 1954–1969." *Journal of Industrial Economics*, 35: 517-542.

Hendricks, K., R.H. Porter, and G. Tan. 1993. "Optimal Selling Strategies for Oil and Gas Leases with an Informed Buyer." *American Economic Review: Papers and Proceedings*, 83(2): 234-239.

Herrnstadt, E., R. Kellogg, and E. Lewis. 2019. "Royalties and Deadlines in Oil and Gas Leasing: Theory and Evidence." Working Paper. Available at: https://home.uchicago.edu/~kelloggr/Papers/HKL_primaryterms.pdf. [Accessed 4/7/19].

Hurn, A. and R. Wright. 1994. "Geology or Economics? Testing Models of Irreversible Investment Using North Sea Oil Data." *Economic Journal*, 104: 363-371.

Kellogg, R. 2011. "Learning by Drilling: Interfirm Learning and Relationship Persistence in the Texas Oilpatch." *The Quarterly Journal of Economics*, 126(4): 1961-2004.

Kellogg, R. 2014. "The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling." *American Economic Review*, 104 (6): 1698-1734.

Leonard, B., and D. Parker. 2019. "Private vs. Government Ownership of Natural Resources: Evidence from the Bakken." Working Paper. Available at: https://aae.wisc.edu/dparker/wp-content/uploads/sites/12/2019/02/Leonard-and-Parker-Full-2-28-19-1.pdf. [Accessed 3/17/19].

Levitt, C.J. 2009. "Learning Through Oil and Gas Exploration." Working Paper. Available at: https://pdfs.semanticscholar.org/3ef7/128b17edc465e1104281dbf4efa9b11a1a39.pdf. [Accessed 6/11/19].

Levitt, C.J. 2016. "Information Spillovers in Onshore Oil and Gas Production." *Resource and Energy Economics*, 45: 80-98.

Lin, C.-Y.C. 2009. "Estimating Strategic Interactions in Petroleum Exploration." *Energy Economics*, 31: 586-594.

Lin, C.-Y.C. 2013. "Strategic Decision-Making with Information and Extraction Externalities: A Structural Model of the Multistage Investment Timing Game in Offshore Petroleum Production." *The Review of Economics and Statistics,* 95(5): 1601-1621.

McFarland, J.B. 2014. "Checklist for Negotiating an Oil and Gas Lease." Graves, Dougherty, Hearon, and Moody, P.C., Austin, TX. Available at https://www.gdhm.com/wp-content/uploads/2016/08/jbm-ogleasechecklist.pdf. [Accessed 3/2/18].

Montgomery, J.B., and F.M. O'Sullivan. 2017. "Spatial Variability of Tight Oil Well Productivity and the Impact of Technology." *Applied Energy*, 195: 344-355.

Paddock, J.L., D.R. Siegel, and J.L. Smith. 1988. "Option Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases." *Quarterly Journal of Economics,* 103: 479-508.

Pollastro, R.M., T.A. Cook, L.N.R., Roberts, C.J., Schenk, M.D., Lewan, L.O., Anna, S.B., Gaswirth, P.G., Lillis, T.R., Klett, and R.R. Charpentier. 2008. "Assessment of Undiscovered Oil Resources in the Devonian-Mississippian Bakken Formation, Williston Basin Province, Montana and North Dakota, 2008." U.S. Geological Survey Fact Sheet 2008–3021, 2 p. Available at: https://pubs.usgs.gov/fs/2008/3021/pdf/FS08-3021_508.pdf. [Accessed 4/15/18].

Porter, R.H. 1995. "The Role of Information in U.S. Offshore Oil and Gas Lease Auctions." *Econometrica,* 63: 1-27.

Smith, J.L. 2018a. "Estimating the Future Supply of Shale Oil: A Bakken Case Study." *Energy Economics,* 69: 395-403.

Smith, J.L. 2018b. "How and Why Petroleum Leases Are Held by Production." *Land Economics,* 94(1): 52-72.

Steck, A.L. 2018. "Industry Dynamics with Social Learning: Evidence from Hydraulic Fracturing." Available at: https://sites.duke.edu/asteck/files/2017/11/steck_jmp.pdf. [Accessed 7/10/18].

Timmins, C., and A. Vissing. 2017. "Environmental Justice and Coasian Bargaining: The Role of Race and Income in Lease Negotiations for Shale Gas." Working Paper. Available at: https://sites.duke.edu/christophertimmins/files/2017/02/environmental-justice-coasian.pdf. [Accessed 5/6/19].

USGS. 2019. "Budget Justifications and Performance Information Fiscal Year 2019." The United States Department of the Interior – U.S. Geological Survey. Available at: https://www.doi.gov/sites/doi.gov/files/uploads/fy2019_usgs_budget_justification.pdf. [Accessed 4/12/19].

Vissing, A. 2015. "Private Contracts as Regulation: A Study of Private Lease Negotiations Using the Texas Natural Gas Industry." *Agricultural and Resource Economics Review,* 44(2): 120–137.

Vissing, A. 2017. "One-to-Many Matching with Complementary Preferences: An Empirical Study of Market Power in Natural Gas Leasing." Working Paper. Available at: https://pdfs.semanticscholar.org/5014/ca081949118b77a24e5c0b75f050280ce9c6.pdf. [Accessed 4/2/19].

Vissing, A. 2018. "One-to-Many Matching with Complementary Preferences: An Empirical Study of Market Power in Natural Gas Leasing." Working Paper. Available at: https://pdfs.semanticscholar.org/5014/ca081949118b77a24e5c0b75f050280ce9c6.pdf. [Accessed 4/2/19].

4. USING SITE CHARACTERISTICS IN MARINE RECREATIONAL FISHING MODELS: A NEW SITE AGGREGATION APPROACH

4.1. Introduction

Marine recreational angling contributes significantly to the economy in the Gulf of Mexico as it attracts anglers from a variety of areas across the U.S., and bringing in money to coastal economies via the purchase of fishing gear, bait, boat, food, and other revenue-generating trip expenditures. Although these flows are direct services to society, they are dependent on indirect flows such as the abundance of a particular fish species, among other available amenities that contribute to a fishing experience, and influence both the fishing trip and site choice decisions of anglers. To model the indirect flows, revealed preference methods such as the travel cost random utility maximization (RUM) model are often used. The travel cost RUM model is valuable because of its ability to capture and reveal angler demands and behavior towards a conventionally nonmarket good— a particular fishery and the costs anglers incur to access it. Inferences can then be made about the value an angler places on a fishery indirectly in the form of his or her site selection, as a function of the cost of taking the trip, and the site-specific characteristics the angler bases his or her decision on.

The Marine Recreational Information Program (MRIP) of the National Oceanic Atmospheric Administration (NOAA) is the primary program in the U.S. responsible for counting and reporting catch and effort data for public fishing sites along coastal U.S. waters. In this paper, we use MRIP data in recreation demand models to study of the behavior of single-day trip anglers (≤300-mile round trip) who targeted Spotted Seatrout in the Gulf of Mexico over 2013-2015. In previous recreation studies using MRIP data, fishing sites have been aggregated to the county or region level, and models incorporated only a small number of site-specific characteristics, usually including the cost associated with traveling to a fishing site and some form of expected catch as the only factors explaining angler site choice (e.g. Whitehead and Haab 1999; Gentner 2007; and Lovell and Carter 2014).[152]

Information on other site characteristics has largely been limited for several reasons. First, extensive information on site characteristics is often prohibitively expensive to obtain via traditional survey methods and was unavailable through MRIP until 2013. Second, the estimation of models with large choice sets can introduce a large computational burden (Parsons and Needlemen et al. 1992; Feather 1994; Kanaroglou and Ferguson 1996; Haab and Hicks 1999; Haener et al. 2004; Lemp and Kockelman 2012; von Haefen and Domanski 2013; and Alvarez et al. 2014). It also appears that many site characteristics are overlooked due to a limited amount of high-resolution catch (or fish population abundance) data available at the site level; aggregating to the county level helps to alleviate this concern and ease estimation. Site aggregation and the omission of relevant site characteristics however, can cause significant bias to parameter estimates on travel distance and expected catch if these terms are correlated with unobserved site characteristics through their contribution to angler utility (i.e. site choice).

---

[152] Dundas et al. (2017) is one exception who estimate a site-level model, yet they do not include expected catch rates in their site choice model.

In this paper, we estimate site choice models at the individual site level, making use of a dataset that allows us to incorporate a variety of site characteristics not previously analyzed when modeling marine recreational fishing behavior.[153] These data, published by the National Marine Fisheries Service (NMFS) supplementary to the MRIP data, are time-varying, but given the short sample period there is only a small amount of temporal variation, since physical site changes commonly require significant time and funding.[154] In a variety of specifications, we explain angler fishing site choice using both time-variant and time-invariant versions of the characteristics data, travel distance, and expected catch rates for each site— predicted using correlations of catch at nearby sites and over time. These specifications allow us to investigate how the inclusion of previously unobserved site characteristics, such as the proximity to bait and tackle shops, the number of boat ramps, ample parking space, fishing site pressure (congestion), and other amenities, influence site choice.[155]

The modeling framework of Murdock (2006) has become the standard in the recreation demand literature.[156] Her approach incorporates a set of alternative specific constants (ASCs) for all fishing site alternatives into site choice models to address the problem of unobserved site characteristics. Hence, we compare our specifications to a

---

[153] The data can be obtained from the MRIP Public Access Fishing Site Register at: https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp.

[154] The time-varying characteristics data were obtained via email from NOAA Fisheries.

[155] Site pressure data from the MRIP site register are not included in the analysis in this paper and are left for future work.

[156] The modeling framework of Murdock (2006) aims to address the problem of unobserved site characteristics, and is grounded in the modeling framework of Berry (1994) and Berry et al. (1995), who present a model to analyze a differentiated goods market with unobserved product characteristics, and apply this technique in a discrete choice model of automobile demand, respectively.

benchmark model that includes only the travel distance and expected catch rate as explanatory variables, and a set of ASCs for all fishing site alternatives. We use these specifications to test the hypothesis that an extensive set of characteristics can substitute for the use of ASCs, whether they can mitigate endogeneity and resulting biased parameter estimates on the catch and travel distance terms, and to understand how anglers value particular site characteristics. Since the effective management of fish stocks is dependent on an understanding of angler behavior, this information is useful to fishery managers as it provides insight on the changes in the behavior of anglers in response to improving the available amenities in fishing sites (Fujiwara et al. 2018).

In addition to estimating models with complete choice sets, we use a partial site aggregation method that is similar to the approaches used in Parsons and Hauber (1998) and Parsons, Plantinga, and Boyle (2000), but new to the marine recreation demand literature and allows us to study site aggregation bias in a new dimension. In particular, our approach makes use of an outside option fishing site choice that is determined using a distance-based ranking threshold. The approach is based on the assumption that anglers will tend to focus on a subset of nearby sights when making their site-choice decision, so we lump all sites other than the *n* that are nearest to each angler into one option. Our approach builds on the findings from Parsons and Hauber (1998) that sites that are more distant are significantly less relevant to an angler's choice set. This consideration is especially relevant for day-trip anglers who choose a fishing site while considering an allocation of their time between traveling and time spent fishing.

To preview the results, as expected we find that the use of time-invariant site characteristics instead of ASCs provides a nearly identical parameter estimate for expected catch, and a qualitatively similar, but smaller (in magnitude) estimate on the travel distance term. We find that including the time-varying characteristics improves model fit when included with the ASCs. The travel distance parameter remains the same, but the magnitude on the expected catch term is attenuated significantly. The parameter estimates on several of the site characteristics are important to recreational anglers. Namely, we find that private boat anglers are more inclined to visit sites with a bait shop, where major fishing tournaments are held, and with more boat ramps and trailer parking spaces. On the other hand, we find a disutility associated with fishing sites located near restaurants and lodging, which we presume is due to more congestion created by non-anglers. Lastly, we find that estimating the same models after reducing the angler choice set to include a subset of more relevant fishing sites and an outside fishing option provides qualitatively similar results, and improves model run time by 50-75%.

4.2. Literature Review

Choice behavior is one of the fundamental concerns in economics. RUM models are commonly used to study how humans and firms make choices among discrete alternatives, and have been applied in many different settings. McFadden (1973) developed the original RUM model, and used it in an application to study a series of transportation decisions made by individuals. RUM models have since been applied in many other contexts such as the purchasing behavior of humans, choice of college attended, occupation choice, strategic decisions made by firms, and many others. The recreation demand literature has also recognized these models as a useful tool to study how recreationists make a single choice among a specific set of discrete alternatives, including whether to take a fishing or other recreation trip, and where. Appropriately specifying a recreationist's choice set, however, is of great importance due to various tradeoffs and or shortcomings associated with arbitrarily specifying a choice set, including aggregating choices (or sites), including too few or too many alternatives, and others (Haab and Hicks 1999).

The tradeoff between site aggregation and estimation efficiency was investigated in seminal studies of recreation demand in the 1990s, where site aggregation was predominantly studied in settings other than marine fishing. Kaoru and Smith (1990) were the first to investigate the implications of site aggregation in a RUM. In an application with a much larger degree of aggregation, Parsons and Needleman (1992) show that although aggregation is often used to reduce the choice set to a manageable size, it comes with losses in estimation efficiency and caution against it with evidence from their study of recreational fishing on lakes in Wisconsin. Schuhmann (1998) and Train (1998) recognize this shortcoming but to account for potential bias when aggregating fishing sites at the county level, they follow the suggestion of Ben-Akiva and Lerman (1985) by including the log of the number of fishing sites as a quality variable to obtain parameters that are equivalent across scales. Heaner et al. (2004) also find that when developing models of recreational hunting, the size of the choice alternative in spatial context matters.

There are several reasons, both stated and implicit, for why sites have frequently been aggregated in site choice models. Although McFadden (1978) proved that a sampling-of-alternatives design will provide consistent model parameters, it is becoming more common for researchers to use more behaviorally realistic discrete choice models where McFadden’s proof does not hold (Nerella and Bhat 2004). Still, the sampling of alternatives can reduce computational run time substantially. In their study of residential location, Nerella and Bhat (2004) advise against the use of sample sizes that are too small. For conditional or multinomial logit models, they recommending that using an eighth of the size of the full choice set can be used at a minimum and suggest that a fourth of the size is a desirable target. In a recreational fishing study investigating the tradeoff between model run-time and the efficiency (or precision) of welfare estimates using samples with different sized choice sets, von Haefen and Domanski (2013) find that there is modest efficiency loss and significant time savings for conditional logit models with a small number of alternatives. They recommend researchers use best judgement based on their specific need to determine an acceptable amount efficiency loss when reducing choice set sizes.

In another early attempt to a find a better mechanism than site aggregation, Parsons and Kealy (1992) modeled site choices of recreationists using Wisconsin lakes, and analyzed the effect of using randomly drawn subsets of alternatives to estimate the model parameters. Their results suggest that using draws from the complete set of alternatives might be more effective for reducing choice set sizes and computational burden. Feather (1994) finds results that echo these findings in his model of recreational angling in Minnesota. Parsons and Hauber (1998) investigate the sensitivity of RUM parameter and welfare estimates to changes in the spatial boundary of the choice set. Their results imply that the definition of the spatial boundary of the choice set is less important than one might expect. The eighty-fifth percentile of the trips taken have a one-way travel time of less than one hour and a median one-way travel time of approximately twenty minutes, and parameter and welfare estimates show little change as the choice of boundary is increased from 1.6 to 4.0 hours. In their study, the average number of sites in an angler's choice set increased from approximately 350 to more than 1,500 sites over this range, and it was shown that the distant sites are, for the most part, irrelevant to the angler's choice set.

Other early work developed new methods to reduce choice set sizes by accounting for familiar vs. unfamiliar vs. favorite sites (Parsons, Masey, and Tomasi 1999), treated nearby sites as close substitutes and more distant sites as aggregate alternatives in a nested framework (Parsons, Plantinga, and Boyle 2000), used ways to confine the number of choices modeled such as the efficiency approach (Scrogin et al. 2010), aggregated unimportant sites (Lupi and Feather 1998), predicted choice sets (Ben-Akiva and Boccora 1995), or eliminated sites completely (Peters, Adamowicz, and Boxall 1995; and Hicks and Strand 2000). Phaneuf and Herriges (1999) implemented a similar choice set aggregation approach to Parsons and Hauber (1998), but in a Kuhn-Tucker framework, and found similar results that make modeled sites appear unique and increased the magnitudes of estimated welfare effects.

Site aggregation is a particularly important consideration in studies using MRIP data because the complete choice set can contain from several hundred to thousands of possible sites, and it is plausible that some or many of these sites (especially the more distant ones) have a near-zero probability of being chosen. Although most studies that use

MRIP data have aggregated sites to the county level or higher, some have taken one or more additional steps to reduce the size of choice sets, mitigate site aggregation bias, or reduce computational burden. Most MRIP studies that aggregate sites have included the log of the number of choices within each county (Gentner 2007; Haab et al. 2012; Lovell and Carter 2014; and Alvarez et al. 2014). Hindsley, Landry, and Gentner (2011) use Marine Recreational Fisheries Statistics Survey (MRFSS) data, the MRIP parent, and develop a RUM model to investigate the problems inherent with on-site sampling and propose a correction method based on propensity scores that mitigates the sample selection bias. Whitehead and Haab (1999) eliminate sites by using distance and historical catch to determine whether a site is too far or unproductive to be considered a viable choice. Their results were found not to be significantly affected by the elimination of non-viable choices.

One unfortunate consequences of site aggregation is that it makes the use of site characteristics in the model more difficult. It may make sense to assume that the expected catch is roughly the same for all sites within a county, but it makes substantially less sense to assume that a trip to a county is driven by the average number of parking spaces in sites within the county– that is a characteristic that affects only one site. We believe that studies using MRIP have aggregated sites to reduce computational burden, but also due to catch data limitations— a common issue in recreational fishing demand modeling (Morey and Waldman 1998). Given the nature of recreational fishing data and how it is generated, catch data are commonly unavailable due to either no surveys conducted or no anglers present during surveys, even though catch rates may not have changed since the last sampling period. When estimating a site-level model, this shortcoming implies a need to obtain more data, use historical catch rates, and or use catch data from nearby sites to predict catch for a site in a given period. Aggregating sites, say to the county level, is one remedy that has been used and can be viewed as empirically convenient. This is because it is unlikely that all sites within a given county are not surveyed in a period (i.e. significantly fewer missing data issues), and because the observed catch of one angler is not likely to significantly affect the county-level catch rate in that period. Such site aggregation however, is not as useful when the goal is to provide more local policy prescriptions, so alternative workarounds are needed to estimate a site-level model.

Murdock (2006) describes how researchers cannot expect to observe all site characteristics affecting site choice. Using a large sample of 3,581 choice occasions, she studies the demand for recreational fishing on lakes in Wisconsin, and implements an approach that enables her to recover policy-relevant travel cost parameters that differ from traditional models by a factor of four. Her two-stage approach controls for unobserved site characteristics using ASCs and enables her to relax the common assumption that all relevant site characteristics are observable in the available data. She compares the results to specifications with time-invariant site characteristics, including the catch rates for various fish species (although admittedly outdated), and dummy variables for physical lake and lake access characteristics, regulations, fish management, urban area, and residential and industrial development around fishing sites. Also included in her model are interaction terms between individual characteristics such as gender, age, income, and boat ownership to allow for taste heterogeneity.

This paper contributes to the above literature in several ways. First, we make an empirical contribution by incorporating a new and extensive dataset containing time-invariant and time-varying site characteristics. The use of these characteristics is made possible by estimating an individual site-level model, which helps us to circumvent site aggregation bias concerns, and enables us to make an important contribution, as ACSs are unable to control for the temporal component of angler demand for a site due to changes in its amenities. Hence, we contribute by improving on previous models that only include ASCs and a select few explanatory variables, and providing insight on other amenities valued by anglers. Second, we predict expected catch rates in a new way at the site level, making use of spatial and temporal correlations in catch rates from the same and nearby sites. Lastly, we implement a new partial aggregation approach to estimate a site-level model with less computational burden. Future studies using MRIP data would benefit from considering each of these items.

4.3. Data

4.3.1. MRIP

The complete set of MRIP data consists of hundreds of thousands of point-intercept surveys conducted in six two-month waves throughout each year, at coastal sites around the U.S. Sites are sampled (or surveyed) according to a stratified design— the most popular sites are surveyed the most.[157] Recreational anglers are surveyed at the close of

[157] Given the stratified survey design that MRIP uses, it also makes available sampling weights. In theory, these weights should be used in estimation; however, since we are not conducting welfare analysis in this dissertation, we have omitted them. By omitting the survey weights, the qualitative nature of our results remain the same, although their precision should be interpreted with some caution as anglers that are more avid are likely to be surveyed more than a representative angler (i.e. avidity bias).

their fishing day to obtain detailed information on their trip, as well as any catch and associated fish size information if available. In this paper, we make use of the trip and catch data, as well as site distances data, which are published supplementary to the MRIP data and include the estimated driving distance (in miles) and travel time (in hours) from each intercept site to each surrounding zip code.[158] Since we study anglers fishing on what constitutes a single-day trip, we follow Lovell and Carter (2014) and exclude fishing sites located more than 150 miles from an angler's home zip code from all choice sets— a threshold deemed feasible for an angler to travel to and from the site within a day and allowing ample time to fish. We also restrict our sample to include anglers fishing for Spotted Seatrout as either a primary or secondary target species, and restrict their choice sets to include sites where catch data for Spotted Seatrout were available on site or if they were available at nearby sites.[159] Our final data set includes 6,996 unique single-day fishing trips taken by private boat anglers to 432 different sites along the coasts of Mississippi, Alabama, and western Florida over 2013-2015.[160] We focus on anglers

[158] We use travel distance, instead of cost, as distance makes up the largest component of travel cost. In future research, we will calculate an actual travel cost for each site alternative for each angler, making use of the time variable when multiplying by the wage rate for each angler.

[159] Note: we chose anglers targeting Spotted Seatrout, the second most frequently targeted species in the Gulf of Mexico over 2004-15 according to MRIP data. In future research, we will expand the analysis to include Red Drum as a target species, as it was another highly-targeted species in the Gulf, and many anglers reporting Spotted Seatrout as a primary target also reported Red Drum as a secondary target, or vice versa.

[160] Note: in future research, we will also add trips from 2016. However, our sample period will be restricted to 2013-16 for two reasons. First, the site characteristics data are limited prior to 2013, as MRIP transitioned to the use of a new online site register, which also incorporated a new way of reporting fishing pressures. Conversations with MRIP indicated that the characteristics are not consistent over time before 2013 since they were not validated or required to be surveyed. Hence, the pre and post-2013 characteristics are not comparable. Second, post-2016, MRIP halted to provide home zip code data with angler trip records (it now only includes angler's home county code), meaning that we would be using different travel distances for trips over 2013-16 than 2017-18. Zip code information in all previous years has also been stripped from the MRIP data currently available online.

fishing in these states because Texas does not participate in MRIP and Louisiana stopped after 2013.

Summary statistics of the data used in our analysis are available in Tables 13 and 14. Table 13 presents angler and trip attributes. Anglers in our sample come from ten different states, but anglers whose legal residence is in a distant state are included because they reported a home zip code within 150 miles of the fishing site, possibly because of seasonal residency away from their home state. Seventy-nine percent of trips occurred on weekend days, and the average fishing party was between two and three anglers. As seen in Table 14, on average, anglers fishing in Florida travel the furthest (36.23 miles) and those fishing in Mississippi travel the least (16.04 miles). The average size of choice sets for anglers fishing in Florida was the largest (over 150 sites within 150 miles), and was the smallest for anglers fishing in Mississippi (over 90 sites within 150 miles). The average duration of fishing was roughly equal for all states at around 4.5 hours per trip. The most avid anglers were those that were fishing in Mississippi, and the least avid were those that were fishing in Florida, where the average reported number of days fished over the last 12 months were 59 and 50 days in each state, respectively.

4.3.2. Expected Catch

Site choice models of marine recreational angling can suffer from several forms of endogeneity from the expected catch variable. First, catch rates are not the only observable component determining angler site choice. Previous studies using MRIP data have omitted important site characteristics that help characterize the quality of a site or improve the fishing experience. When omitted, these characteristics bias the parameter estimate on the expected catch term to the extent of their correlations with catch rates and their contribution to angler utility.[161] Second, selection (or avidity) bias is a concern in RUM models. Anglers that are more avid have different preferences and expectations about how many fish they might catch, and they also fish more regularly than the rest of the population and are disproportionately more likely to be surveyed. They also commonly move, or sort, closer to fishing sites.[162] Third, an angler's site choice affects the catch rate at that site after visiting and catching fish, yet it is also likely that anglers choose a particular site because of a good catch rate at that site. Hence, some simultaneity would occur if we used observed angler catch for the given two-month wave in the calculation of the catch rate for that site in the same wave. Since we are estimating a site-level model, and given missing data for certain sites in certain periods attributable to the survey design, a unique approach is needed to predict expected catch rates for each site in each wave.

Most studies using MRIP data have calculated expected catch by aggregating sites to the county level and then used some form of lagged average catch rates at all sites within each county as a proxy for expected catch. We correct for the missing data issue and for the simultaneity concern by proxying for current wave catch rates with a weighted spatial-temporal form of lagged (or past) catch rates that predicts deviations of expected catch

[161] Ideally, an instrument for expected catch would be used, but at the time of this analysis, data on an instrument such as fish population was unavailable.
[162] MRIP sample weights may be used to correct for selection bias introduced by the survey design, and additional models can be specified to control for avidity using interaction terms between angler-specific variables that proxy for avidity, such as the number of days fished in the last year. Hindsley et al. (2011) introduce an approach to address this issue, and bias associated with sample choices may not reflect the site choices of the true population. We leave this for future research.

rates from the county average one wave (or two-month period) ahead for all sites over the whole sample. We believe this approach is reasonable, as there is plausible correlation between nearby fishing sites and over time since fish migrate and because marine anglers commonly travel several miles on the water while in their boats.

Our approach predicts the expected catch at site j in wave t, $C_{j,t}$, based on observed catch data at surrounding sites and at the same site in previous periods. Specifically, the prediction of $C_{j,t}$ starts with the average catch over the full period observed (2012-15) in the county in the j^{th} site is located, $\bar{C}_j$.[163] We then estimate the deviation of $C_{j,t}$ from $\bar{C}_j$ using catch observations in surrounding sites in the current wave and in all sites in the previous waves, with the weight given to a particular observation declining as the distance from j and t increases in both space and time. Let C_k be the catch reported for the k^{th} observation of the data, i.e. for an individual angler at a specific site. We proxy expected catch for each site in each wave by using all observations using a weighted sum:

$$E\left(C_{j,t}\right)=\frac{\Sigma_k \omega^{dist}\left(k,j\right)\omega^{wv}\left(k,t\right)\omega^{yr}\left(k,t\right)\left(C_k-\bar{C}_k\right)}{\Sigma_k \omega^{dist}\left(k,j\right)\omega^{wv}\left(k,t\right)\omega^{yr}\left(k,t\right)}+\bar{C}_j \tag{4.1}$$

where $\bar{C}_k$ is the average catch in the county of the k^{th} observation, and $\omega^{dist}\left(k,j\right)$, $\omega^{wv}\left(k,t\right)$, and $\omega^{yr}\left(k,t\right)$ are weights given for the k^{th} observation. We assume that the weights decline geometrically as follows: $\omega^{dist}\left(k,j\right)=exp\left(-\alpha_d \cdot dist\left(k,j\right)\right)$ where $dist\left(k,j\right)$ is the linear distance between site j and the site of the k^{th} observation,

---

[163] Note: although our trip data used in estimation cover 2013-15, we also use catch data from 2012 in order to predict catch rates in 2013.

$\omega^{wv}(k,t) = exp\left(-\alpha_w \cdot \left(t - w(k)\right)\right)$ where $w(k)$ is the wave of the k^{th} observation for $t - w(k) \leq 5$ and $\omega^{y}(k,t) = \alpha_y$ for $t - w(k) = 6$. We then choose α_d, α_w, and α_y, to minimize the sum of squared differences between predicted expected and actual catch. When these parameters model were estimated in an unconstrained model, the estimated coefficients essentially give equal weight to all observations in the data. To ensure sufficient reasonable diffusion in the weights over time and space, the coefficients were constrained as follows: $0.01 \leq \alpha_d \leq 0.3$; $0.05 \leq \alpha_w \leq 7$; and $0 \leq \alpha_y \leq 1$. For the catch rates in each wave in 2013, we use data from the previous six waves in 2012, starting with wave one. Summary statistics for expected catch rates in each state are available in Table 15.

4.3.3. Public Access Fishing Site Register

The site characteristics data come from the MRIP Public Access Fishing Site Register, a tool maintained by NOAA Fisheries that serves two primary purposes.[164] First, it acts as a public resource for anglers and others, providing detailed information about available amenities and current fishing pressures at publicly accessible marine fishing sites from Maine to Louisiana (see Figure 22 for an example of a characteristics page). Second, it is used by NOAA to gather important site-level information about the frequency of angler fishing trips (effort) throughout the year, what types of fishing occur, and which site amenities attract anglers to visit. Keeping record of this information is important to help ensure that an accurate representation of fishing site pressure (see Table 18 for a

---

[164] See https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp.

description of these data) is gathered so that NOAA can determine properly where to send survey enumerators and in appropriate proportions.[165]

The characteristics list (see Table 19 for a complete list with descriptions) includes an extensive set of site amenities that contribute to an angler's experience before, during, and after fishing.[166] Two versions of the characteristics dataset are used in this analysis. The first is a time-invariant dataset, and represents a snapshot of the characteristics at each site in early 2017.[167] The second contains a time varying set of characteristics, which includes a snapshot of characteristics in each of six two-month waves per year over 2013-2015. However, given the short sample period of this study, the newness of the NOAA program that tracks site characteristics, and the time and funding-intensive nature for changes in characteristics to occur, there is only a small amount of variation present in this dataset for a small subset of the characteristics list. Hence, we focus on the time-invariant characteristics and, where available, those that are time varying. We discuss the characteristics selection process in more detail in the following subsections and provide summary statistics for characteristics of sites in each state in Table 15.

---

[165] Source: https://www.st.nmfs.noaa.gov/recreational-fisheries/data-and-documentation/site-register.

[166] The characteristics of primary interest for this paper include the number of boat slips, car parking spaces, boat ramps, and trailer parking spaces, as well as indicators for whether there is a fuel dock, bait shop, and fish cleaning stations on site. Others include whether fishing is affected by the tide, if there is a fee required for site access or if the site has private access, if there are major tournaments held at the site, and whether there are restaurants and lodging nearby.

[167] A snapshot of the 2017 characteristics is used as a proxy for characteristics over 2013-15 for two reasons. First, we initially believed that more recent values may be more reliable. Second, there was not very much variation in site characteristics over time. A total of 15 of the characteristics had changes at a select number of sites during the sample period, but the number of sites with observed changes was limited. 14.75% of sites had changes in the number of trailer parking spaces. 9.22% had changes in the number of car parking spaces. 8.2% had changes in the number of boat slips. 6.56% had changes in available night lighting onsite. 2.46% had changes in the number of boat ramps, and 1.23% had changes in the availability of nearby lodging. The rest of the characteristics changed at less than 1% of sites or not at all over 2013-2015.

4.4. Empirical Strategy

4.4.1. Standard Approach

Drawing on the econometric models of Morey et al. (1993 and 2002), Parsons and Hauber (1998), Whitehead and Haab (1999), Parsons, Plantinga, and Boyle (2000), Murdock (2006), Haab et al. (2009, 2012), Alvarez et al. (2014), Lovell and Carter (2014), and others, our approach implements site choice models of marine recreational fishing demand. A standard conditional logit model including all fishing sites within 150 miles of an angler's home zip serves as the benchmark for later comparison to a specification using the partial site aggregation approach. As developed by McFadden (1973) and Manski (1977), this model is specified as follows:

$$P_{ij} = P\left(y_i = j\right) = \frac{e^{V_{ij}}}{\sum_{k=1}^{J} e^{V_{ik}}}, \tag{4.2}$$

where $V_{ij} = \beta_d d_{i,j} + \beta_c c_{j,t} + \delta_j + \varepsilon_{ij}$, P_{ij} is the probability that angler i will chose site j, d_{ij} is the travel distance (a portion of the trip cost) for angler i to get to site j, $c_{j,t}$ is the expected catch rate for all anglers at site j, and ε_{ij} is a random error term with a Type-I extreme value distribution.[168] Following Murdock (2006), our benchmark model also includes ASCs, δ_j, to control for unobserved site characteristics.[169] The parameters to be estimated are β_d and β_c, and the log-likelihood equation is:

[168] Note: we include travel distance since it constitutes the largest component of travel cost, which we believe is okay since the goal of this paper is not welfare analysis.

[169] Note: in our specifications that include time-invariant site characteristics, the indirect utility function will not include ASCs and will take the form $V_{ij} = \beta_d d_{i,j} + \beta_c c_{j,t} + \boldsymbol{\beta_q q}_j + \varepsilon_{ij}$, where $\boldsymbol{q_j}$ is a vector of site characteristics.

$$LL(\beta) = \sum_{i=1}^{I} \sum_{j=1}^{J} d_{ij} \cdot log(P_{ij}) , \quad (4.3)$$

where I is the number of anglers in the sample, and d_{ij} = 1 if angler i chooses site j and 0 otherwise. Although the conditional logit remains the workhorse of discrete-choice models, one assumption that underpins the model is the independence of irrelevant alternatives (IIA) condition, which states that any sites added to the choice set will decrease the choice likelihood of all sites by an equal fraction.

4.4.2. Partial Site Aggregation Approach

Often limited by data availability, it is commonly up to the analyst to determine a reasonable method for determining appropriate choice set sizes and compositions. For this reason, researchers have estimated models using choice sets that have been narrowed in a variety of ways, sometimes relying on surveys responses to obtain information on actual choice set perception or awareness of anglers (Parson, Massey, and Tomasi 1999). More routinely, researchers have aggregated sites in of ad hoc ways, such as to the county or region level, which is not sufficient when a goal is to determine the value of local and site specific characteristics that influence an angler's choice decision. Inappropriate assignment of the angler's feasible and realistic choice set has been shown to be counterproductive in recreation demand studies. If sufficient care is not taken when determining both the resolution and composition of choice sets, parameter and welfare estimates can be significantly biased with respect to the true underlying choice set (Haab and Hicks 1999; Hicks and Strand 2000; and Parsons, Plantinga, and Boyle 2000).

Our dataset includes 432 recreational fishing sites in the Gulf of Mexico, spanning from Mississippi to Collier County in southwestern Florida. The average choice set size per angler in our dataset is 144 sites (the largest choice set has 224 sites), feasible for a one day fishing trip. Hence, the comprehensive dataset includes 6,996 angler trip records, and over one million choice observations. We propose a partial site aggregation mechanism and implement alternative specifications of equation (4.1) that make use of an outside option in each angler's choice, for two reasons. First, the complete data set including all choices is nontrivially large and computationally burdensome dataset to estimate RUM models. Second, 144 potential fishing sites to choose from is a significantly large number for the average angler, let alone 224 for the angler with the largest choice set. Although feasible for a one-day trip, it is likely that anglers do not consider the majority of these sites when making fishing site decisions. Parsons and Hauber (1998) provide evidence of this nature and discuss a distant spatial boundary at which point fishing sites become less relevant to anglers, and find that there exists some threshold distance beyond which adding more sites to the choice set has negligible effects on the estimation results. Lupi and Feather (1998) and Parsons, Plantinga, and Boyle (2000) offer similar strategies for doing so, which involve the inclusion of popular individual sites in the choice set, and then including aggregations of other remaining groups of similar sites.

Our approach builds on the fact that anglers tend to visit one of the closest sites. Although an angler could take a day trip to any site within 150 miles of their home zip code, we presume that the most relevant fishing sites are those that are the nearest, possibly due to more familiarity or experience with these sites, among other factors. Hence, anglers trade off travel distance to fish at a site with better catch rates or other amenities. To study the most relevant sites in an angler's choice set, we first rank each angler's set of feasible sites by distance and then aggregate all sites ranked worse than an arbitrarily-chosen ranking.[170]

Figures 23 and 24 provide visual evidence of the relevance of fishing sites nearest an angler's home zip code. Figure 23 provides a histogram of the distance-based distribution (ranking) of angler site selection. Following our prediction, the figure shows that anglers select better-ranked sites with the largest frequency and the frequency decreases with ranking, or as travel distance increases (i.e. downward sloping demand). Figure 24 provides a CDF of the distribution of ranked sites, and shows that 92% of angler trips in our dataset were to one of the nearest 50 sites in an angler's choice set.[171] Building on these findings, we alter the choice set in equation (4.1). Instead of J_i alternatives, we use a modified choice set, J_i^*, consisting of the 50 nearest sites to angler *k*, plus one alternative (or outside option) consisting of all the less relevant options. To model this specification, we rewrite equation (4.1) as:

$$P\left(y_i = j \in J^*\right) = \frac{e^{V_{ij}}}{1+\sum_{k=1}^{n} e^{V_{ik}}} \text{ or } P\left(y_i = j \not\in J^*\right) = \frac{1}{1+\sum_{k=1}^{n} e^{V_{ik}}}. \qquad (4.4)$$

The numerator in (4) remains the same as in the full model if $j \in J^*$, i.e. for one of the n nearest sites. If the outside option is chosen, the numerator collapses to one and is

[170] The distance-based ranking gives a better ranking (i.e. a ranking that is lower in magnitude) to closer sites, and a worse ranking (i.e. a ranking that is higher in magnitude) to more distant sites.
[171] 80.9% of trips are to one of the nearest 25 sites, 57.4% are to one of the nearest 10 sites, and 39% are to one of the nearest 5 sites in an angler's choice set.

estimated as a constant. The denominator in this probability is what changes the most under the new specification, as it is smaller due to fewer choices.

4.4.3. Choosing Site Characteristics

One of the goals of this paper is to determine how anglers value particular site characteristics that contribute to a fishing experience. This subsection introduces our strategy to choose relevant characteristics without introducing multicollinearity issues associated with the inclusion of highly correlated site characteristics. We first constructed a correlation table of site characteristics and independent variables based on the set of chosen sites from our sample of 6,996 anglers. We used this table to identify characteristics that would be potentially collinear if included together in a regression model and to identify a set of preliminary characteristics to include.

Using a threshold of 0.60, we found several groupings of characteristics that were highly correlated with one another, particularly the service-type characteristics such as boat storage, boat maintenance and repair, and fuel dock. This is intuitive since anglers desiring to store a boat on site are likely to also prefer a site with onsite maintenance for convenience purposes. Other convenience-type characteristics such as an onsite fuel dock, bait shop, and tackle shop were highly correlated. Additionally, there was a high degree of correlation between the number of car parking spaces and the number of boat slips, and the number of boat slips and the presence of a fuel dock. This third grouping likely reflects the nature of sites that are more accommodating to higher income anglers who have larger boats and typically leave them docked onsite when not in use. That is, when anglers have a boat that is larger and more difficult to transport (or if they simply like one particular site), they may purchase or rent a boat slip, which reserves a docking space and enables them drive to the site in a car and fill the boat with fuel within the marina. Fishing sites at large marinas with many boat slips also commonly have bait and tackle shops onsite.

Since high correlations were present among combinations of variables within each of these groupings, (i.e. a correlation coefficient greater than 0.60), we decided to include only one variable from each groupings, with the exception of the convenience characteristics.[172] We decided to keep both fuel dock and bait shop since their correlation was barely above the 0.60 threshold, and because they offer different types of convenience for fishing trips, meaning they contribute to angler utility in different ways that are interesting to analysts. This left us with 15 site characteristics. We first specify a model containing this full set of characteristics and no ASCs, and then use several model selection criteria when comparing specifications with site characteristics to the benchmark model containing only the travel distance and expected catch terms.

4.5. Results

The results for the specifications using the complete and partially aggregated angler choice sets are presented in the next two subsections. Identical model selection processes were used before reporting each set of results under the two choice set

[172] In future research, we will also try using first and second principal components for groupings of characteristics, which would create an index for each grouping and enable us to study how a one-unit in the index changes affects an angler's site choice decision, which would be less ad hoc than our current approach of choosing characteristics.

specifications. Initially, a benchmark model was estimated that only included travel distance, expected catch, and ASCs.[173]

4.5.1. Standard Approach

The results for the specifications with the full choice set are available in Table 16. The choice set for each angler includes all fishing sites within 150 miles of the home zip code, and each model was estimated using 6,996 angler trip decisions from a total of 1,007,925 fishing site alternatives. Columns 1 and 2 provide the results for the models with travel distance and expected catch as the only explanatory variables, with column 2 including ASCs. While distance and catch were was statistically significant at the 1% level in both specifications, excluding ASCs (column 1) results in large underestimates of the impact of distance and expected catch, by roughly one-third, respectively, relative to the benchmark in column 2. Column 3 provides the results for the specification that included the same benchmark characteristics as in columns 1 and 2, but used a set of time-invariant site characteristics instead of ASCs. The signs on the parameters for travel distance and expected catch were again in the predicted direction and statistically significant at the 1% level. Interestingly, the magnitude on the expected catch parameter was nearly the same as that in column 2, but the distance coefficient is closer to that in column 1.

Except for the presence of a fuel dock, the coefficients on the other 14 site characteristics were all statistically significant at the 1% level. The signs on each parameter are interesting and rather intuitive. The results showed a positive correlation

[173] Note that a benchmark model without ASCs was also estimated for each choice set specification, which is used for comparison as well.

between the number of boat trailer parking spaces, the number of car parking spaces, the number of boat ramps, and the probability that an angler visits that site, indicative that anglers value not having to fight for parking space or congestion at the ramp when trying to get on the water. The results also showed the presence of night lighting to have a negative correlation with the probability of an angler choosing a site, which we do not have an intuitive explanation for.

Intuitively, the effects of whether access to a site is private or restricted, whether there is an access fee (an increase in the price of a trip), and whether the ocean tide affects fishing activity are all negative— indicating a lower likelihood of an angler choosing a site with these characteristics. The presence of a bait shop and whether major fishing tournaments are held onsite both provide significant utility to anglers. The presence of a fish cleaning station onsite and nearby restaurants and lodging were revealed as disamenities to private boat anglers. Normally, we would consider fish cleaning stations as a characteristic as a positive influence on angler utility. The negative sign could be attributed to certain anglers using them and spending more time onsite and creating congestion at the end of a fishing day when other anglers are trying to maneuver the site, trailer their boat, and depart. The negative signs on the restaurant and lodging coefficients are also intuitive since we are modeling single-day trip anglers who likely do not need lodging, and may associate traffic brought in by local restaurants and lodging with increased site congestion. Lastly, we find a positive sign on the coefficient for the number of charter boats and a negative sign on the coefficient for the number of head boats at a site. These signs are also intuitive as charter boats attract a different population of anglers than head boats, and contribute less to site congestion since they hold significantly fewer people, whereas head boats provide more of a large-group or party fishing experience.[174] Hence, sites with a large number of head boats may send signals of congestion from more novice and or lower income anglers.

Columns 4-6 provide the results for the specifications that included time-varying characteristics and a set of ASCs as in column 2. Unsurprisingly, the coefficients on travel distance and expected catch are nearly identical to the benchmark model with ASCs in column 2. However, with little temporal variation in the majority of the site characteristics in the site registry data, we were limited in the characteristics that could be included and still obtain convergence in estimation. In column 6, we include five site characteristics with some temporal variation: the number of trailer parking spaces, the number of car parking spaces, the availability of night lighting, and the number of boat slips. The interpretation of these characteristics however, is slightly different, as the time-varying site characteristics only capture the extent to which changes in these variables increase site choice. Since there is not much variation provided by the changes across time or across sites, the time-varying variables are only capturing the response to a small amount of change, and only the number of car parking spaces was statistically significant (5% level). We do believe however, that the specification in column 5 is an improvement over the benchmark model in column 2 for two reasons. First, the model fit is better according to

[174] Charter boats routinely hold 6-15 people depending on the size of the boat, and are more expensive since they offer a more personalized and guided experience and fewer people than head boats, which can hold 40 to over 100 people, depending on the size of the boat. Source: https://www.charterfishingdestin.com/charters/.

the lower AIC, but also because the model run time was quicker by nearly 40%. When comparing the results of column 5 with the time-variant characteristics to those in column 3 with the time-invariant characteristics, the latter appears to provide a consistent estimate of the contribution to utility from an increase in the expected catch rate, but the effect of travel distance in this model is biased towards zero.

4.5.2. Partial Site Aggregation Approach

The results for the specifications with the partial choice set aggregation are available in Table 17. The layout of this table is identical to that of Table 16, with the only difference being the composition of angler choice sets, which still include fishing sites within 150 miles of the home zip code, but only the 50 closest sites to each angler are included, plus an outside option. Each model was estimated using 6,996 angler trip decisions from a total of 353,379 fishing site options— roughly one-third the number of options in the specification above.

Columns 1 and 2 again provide the results for the models with travel distance and expected catch as the only explanatory variables, with column 2 including ASCs. The coefficients on travel distance and expected catch were statistically significant at the 1% level and had the correct sign in each column, although relative to column 2 which included ASCs, this time the specification in column 1 overestimates the effect of an additional fish increase in the expected catch rate.[175] Column 3 reports the results for the

[175] Note: in the specifications with the outside option in Table 17, all sites with a ranking worse than 50 are dropped for each individual angler, yet the model is estimated with a full set of ASCs for all possible sites across all angler choice sets. Hence, the comprehensive site list for the partial site aggregation specifications contains 433 site alternatives including the outside option, relative to the 432 in the specifications without the outside option in Table 16.

specification including the same time-invariant site characteristics and no ASCs as in Table 16. The coefficients on travel distance and expected catch characteristics are again statistically significant at the 1% level in column 1, but the magnitudes are smaller than those in column 2. The coefficient estimates on the site characteristics are all similar in sign and statistical significance to those in Table 16, except the presence of a fuel dock was positive and statistically significant at the 10% level in this specification. Except for lodging, the number of charter boats, and the number of head boats, the coefficients are larger in absolute value than their counterpart model in Table 16.

Columns 4-6 provide the estimates for the specifications including time-varying site characteristics and ASCs. Again, the results and conclusions are qualitatively the same as those in the specifications using complete angler choice sets, including statistical significance at the 1% level for each of the same terms as in Table 16. However, the tradeoffs between the benchmark model in column 2 and the specification in column 5 are slimmer. Again, column 5 provides nearly identical parameter estimates on the travel distance and expected catch terms as in column 2, along with a slightly better model fit according to the AIC.

The specifications with the partial site aggregation also provided a signification improvement in model run time relative to the specifications using complete angler choice sets in Table 16. Each of the benchmark models in columns 1 and 2 saw an improved model run time by nearly 75%, with the benchmark model with the ASCs in column 2 converging roughly two hours sooner than its counterpart in Table 16. Similar, although less extreme improvements were found in columns 3-6, the model run time of the specification in column 3 improved by roughly 50%, and that of the preferred specification in column 5 improved by just over 50% (or roughly 47 minutes).

4.6. Discussion

The main results in this paper are mostly intuitive and as expected, yet there is still some additional work needed to fill a few gaps. It is clear that the inclusion of ASCs is important in site choice models of recreation demand, particularly when sufficient site characteristics are unavailable. When comparing our specifications that include ASCs to those that do not, we find that those omitting them persistently estimate lower values on the travel distance parameter, but there is not a significant effect on other coefficients. Based on the AIC, the models with ASCs also have the best overall fit and, given our limited dataset, model fit improves when time-varying site characteristics are included in the model. These findings hold for the specifications using complete angler choice sets in estimation and for those using partially aggregated choice sets, although additional time-varying changes would be useful in future work.[176]

We also contribute to the literature by introducing a new partial site aggregation approach. In previous studies using MRIP data, the lack of granularity in angler choice sets is a major shortcoming. The inclusion of too few or too many sites can result in incorrect or inefficient parameter estimates on the travel distance (or cost) and catch

[176] Note: MRIP has stopped providing home zip code information for anglers, so future research looking to use these data may be limited. However, we do have all MRIP data including site characteristics available through 2016, which were obtained before the angler zip code information was removed from publicly available datasets. Hence, we can expand this analysis by one year to cover trips over 2013-2016, which may provide slightly more variation in site characteristics over time, but would also increase computational burden due to more observations.

variables, and computational burden increases with the level of resolution in choice sets. We overcome the previous lack of resolution by estimating site-level models incorporating all feasible day-trip fishing sites into each angler's choice set. Since it is also reasonable to assume that most anglers do not consider more than a few sites, we estimate alternative specifications that use a smaller choice set that we deem is more relevant to anglers. When comparing the results across the two choice set specifications, we find that the magnitude of the effects for the distance and catch terms are smaller under the specifications estimated using partially aggregated choice sets. Coinciding with the smaller magnitudes on these terms, we find that the magnitudes on the site characteristics increase under the specification using partially aggregated choice sets.

4.7. Conclusion

In this paper, we study the marine fishing behavior of private boat anglers in Alabama, Mississippi, and Western Florida, and make several contributions to the recreation demand literature. To our knowledge, we are the first to use an extensive set of site characteristics in a site-level RUM model using Marine Recreational Information Program data. We also estimate expected catch in a new way, and contribute to the site aggregation literature by introducing a new partial site aggregation approach that places more emphasis on more relevant fishing sites and reduces computational burden by one-half to three-fourths.

The findings associated with site characteristics provide some unique results, and confirm priors that marine fishing trips are influenced by more than just travel distance and expected catch rates alone. We find that many site attributes contribute to angler utility through a variety of pathways. These include convenience and ease of access, evidenced by the impact of characteristics such as an onsite bait shop and fuel dock as well as the number of boat ramps and the number of trailer parking spaces. They also include pathways that anglers potentially perceive as signals of site quality, such as whether major fishing tournaments are held onsite, and congestion-related pathways, evidenced by the disutility that day-trip anglers were revealed to receive from a fishing site's proximity to restaurants and lodging, both of which potentially add non-fishing related traffic that increases congestion. Although it is clear that ASCs should be included in models when the goal is welfare analysis, the specifications in this paper that substitute time-invariant site characteristics for ASCs enable analysts and fisheries managers to qualitatively understand other factors that influence angler site selection and can improve the quality of a fishing experience.

References

Alvarez, S., S. Larkin, J. Whitehead, and T. Haab. 2014. "A Revealed Preference Approach to Valuing Non-Market Recreational Fishing Losses from the Deepwater Horizon Oil Spill." *Journal of Environmental Management*, 145: 199-209.

Ben-Akiva, M., and B. Boccara. 1995. "Discrete Choice Models with Latent Choice Sets." *International Journal of Research in Marketing,* 12: 9-24.

Ben-Akiva, M, and S.R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand,* 1st ed. Boston, MA: MIT Press.

Berry, S. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics,* 25(2): 242–262.

Berry, S., J. Levinsohn, and A. Pakes, 1995. "Automobile Prices in Market Equilibrium." *Econometrica,* 63(4): 841-890.

Dundas, S.J., R.H. von Haefen, and C. Mansfield. 2018. "Recreation Costs of Endangered Species Protection: Evidence from Cape Hatteras National Seashore." *Marine Resource Economics,* 33(1): 1-25.

Feather, P. 1994. "Sampling and Aggregation Issues in Random Utility Model Estimation." *American Journal of Agricultural Economics,* 76(4): 772-780.

Fujiwara, M. J.D. Backstrom, and R.T. Woodward. 2018. "A Coupled Recreational Anglers' Decision and Fish Population Dynamics Model." *PloS One,* 13(10): e0206537.

Gentner, B. 2007. "Sensitivity of Angler Benefit Estimates from a Model of Recreation Demand to the Definition of the Substitute Sites Considered by the Angler." *Fishery Bulletin,* 105: 161-167.

Haab, T.C., and R.L. Hicks. 1999. "Choice Set Considerations in Models of Recreation Demand: History and Current State of the Art." *Marine Resource Economics,* 14(4): 271-281.

Haab, T., R. Hicks, K. Schnier, and J.C. Whitehead. 2012. "Angler Heterogeneity and Species-Specific Demand for Recreational Fishing in the Southeast United States." *Marine Resource Economics,* 27(3): 229-251.

Haener, M.K., P.C. Boxall, W.L. Adamowicz, and D.H. Kuhnke. 2004. "Aggregation Bias in Recreation Site Choice Models: Resolving the Resolution Problem." *Land Economics,* 80(4): 561-574.

Hicks, R.L., and I.E. Strand. 2000. "The Extent of Information: Its Relevance for Random Utility Models." *Land Economics,* 76(3): 374-385.

Hindsley, P., C.E. Landry, and B. Gentner. 2011. "Addressing Onsite Sampling in Recreation Site Choice Models." *Journal of Environmental Economics and Management,* 62: 95-110.

Kanaroglou, P.S., and M.R. Ferguson. 1996. "Discrete Spatial Choice Models for Aggregate Destinations." *Journal of Regional Science,* 36(2): 271-290.

Kaoru, Y., and V.K. Smith. 1990. "'Black Mayonnaise' and Marine Recreation: Methodological Issues in Valuing a Cleanup." *RFF Discussion Paper QE91-02.* Resources for the Future, Washington, DC.

Lovell, S.J., and D.W. Carter. 2014. "The Use of Sampling Weights in Regression Models of Recreational Fishing-site Choices." *Fishery Bulletin,* 112: 243-252.

Lupi, F., and P.M. Feather. 1998. "Using Partial Aggregation to Reduce Bias in Random Utility Travel Cost Models." *Water Resources Research,* 34(12): 3595-3603.

Manski, C.F. 1977. "The Structure of Random Utility Models." *Theory and Decisions,* 8: 229-245.

McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka, ed. *Frontiers in Econometrics.* New York: Academic Press, pp. 105-142.

McFadden, D. 1978. "Modeling the Choice of Residential Location." *Transportation Research Record,* 673: 72-77.

Morey, E.R., R.D. Rowe, and M. Watson. 1993. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *American Journal of Agricultural Economics,* 75: 578-592.

Morey, E.R., W.S. Breffle, R.D. Rowe, and D.M. Waldman. 2002. "Estimating Recreational Trout Fishing Damages in Montana's Clark Fork River Basin: Summary of a Natural Resource Damage Assessment." *Journal of Environmental Management,* 66: 159-170.

Morey, E.R., and D.M. Waldman. 1998. "Measurement Error in Recreation Demand Models: the Joint Estimation of Participation, Site Choice, and Site Characteristics." *Journal of Environmental Economics and Management,* 35(3): 262-276.

Murdock, J. 2006. "Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand." *Journal of Environmental Economics and Management,* 51: 1-25.

Nerella, S., and C.R. Bhat. 2004. "Numerical Analysis of Effect of Sampling Alternatives in Discrete Choice Models." *Transportation Research Record,* 1894: 11-19.

Parsons, G.R., and A.B. Hauber. 1998. "Spatial Boundaries and Choice Set Definition in a Random Utility Model of Recreation Demand." *Land Economics,* 74(1): 32-48.

Parsons, G.R., and M.J. Kealy. 1992. "Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation." *Land Economics,* 68(1): 93-106.

Parsons, G.R., D.M. Massey, and T. Tomasi. 1999. "Familiar and Favorite Sites in a Random Utility Model of Beach Recreation." *Marine Resource Economics,* 14(4): 299-315.

Parsons, G.R., and M.S. Needleman. 1992. "Site Aggregation in a Random Utility Model of Recreation." *Land Economics,* 68(4): 418-433.

Parsons, G.R., A.J. Plantinga, and K.J. Boyle. 2000. "Narrow Choice Sets in a Random Utility Framework of Recreation Demand." *Land Economics,* 76(1): 86-99.

Peters, T., W.L. Adamowicz, and P.C. Boxall. 1995. "Influence of Choice Set Considerations in Modeling the Benefits from Improved Water Quality." *Water Policy and Economics,* 31(7): 1781-1787.

Phaneuf, D.J., and J.A. Herriges. 1999. "Choice Set Definition Issues in a Kuhn-Tucker Model of Recreation Demand." *Marine Resource Economics,* 14(4): 343-355.

Schuhmann. P.W. 1998. "Deriving Species-Specific Benefits Measures for Expected Catch Improvements in a Random Utility Framework." *Marine Resource Economics,* 13(1): 1-21.

Scrogin, D., R. Hofler, K. Boyle, and J.W. Milon. 2010. "An Efficiency Approach to Choice Set Formation: Theory and Application to Recreational Destination Choice." *Applied Economics,* 42: 333-350.

Train, K.E. 1998. "Recreation Demand Models with Taste Differences Over People." *Land Economics,* 74(2): 230-239.

Von Haefen, R., and A. Domanski. 2013. "Estimating Mixed Logit Models with Large Choice Sets. Working Paper. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1027.9768&rep=rep1&type=pdf. [Accessed 3/10/2017].

Whitehead, J.C., and T.C. Haab. 1999. "Southeast Marine Recreational Fishery Statistical Survey: Distance and Catch Based Choice Sets." *Marine Resource Economics,* 14(4): 283-298.

5. CONCLUSIONS

In this dissertation, I study various behavioral and information-related concerns in two economically important industries: upstream oil and gas and marine recreational fishing. The emphasis of the essays on oil and gas-related topics is primarily related to strategic decision-making by firms in this industry, and related spillovers occurring due to oil and gas development activities. The emphasis of the marine recreational fishing essay is to study how the presence of particular fishing site characteristics and amenities influences angler site choice, and to improve tractability for researchers when defining angler choice sets, which could prove helpful to improve computational efficiency during the estimation of large discrete choice models.

In the first essay, I study how oil and gas producers in Texas report their water use differently if a well is located in a groundwater conservation district, and how their water use affects local groundwater levels. Previous research has studied the reporting of chemical additives and other inputs used in hydraulic fracturing fluids, investigated trends in the historical water use of the industry, and qualitatively characterized potential impacts of its water use. This essay is the first to study reporting in the context of water, and the first to show a causal link between the industry's water use and local groundwater availability. Although my primary policy recommendations involve expanding the reporting requirements of firms, the findings in this essay have broader implications as water scarcity is one of the biggest constraints imposed on economic development. Hence, they have direct policy implications for areas new to unconventional oil and gas development, and abroad for nations such as Argentina and China who are looking to develop shale resources located in arid parts of their respective countries.

In the second essay, I study another dimension of upstream oil and gas development. Specifically, I analyze how it is not only private information that firms use to inform location decisions when targeting new acreage to acquire drilling rights, but freely available public information as well. I find that after the results of a US Geological Survey assessment of oil and gas resources in the Williston Basin were announced, firms used the estimates of resource abundance as an aid to acquire large leases in areas with better geology. These findings indicate that during the early stages of the development of a new play, firms may have less knowledge on local geology than previously believed. Other findings in the essay suggest several important welfare and bargaining implications, as I find that pass-through of geological quality to royalty rates occurred in some of the assessed areas estimated to have abundant resources, but in other areas, it did not.

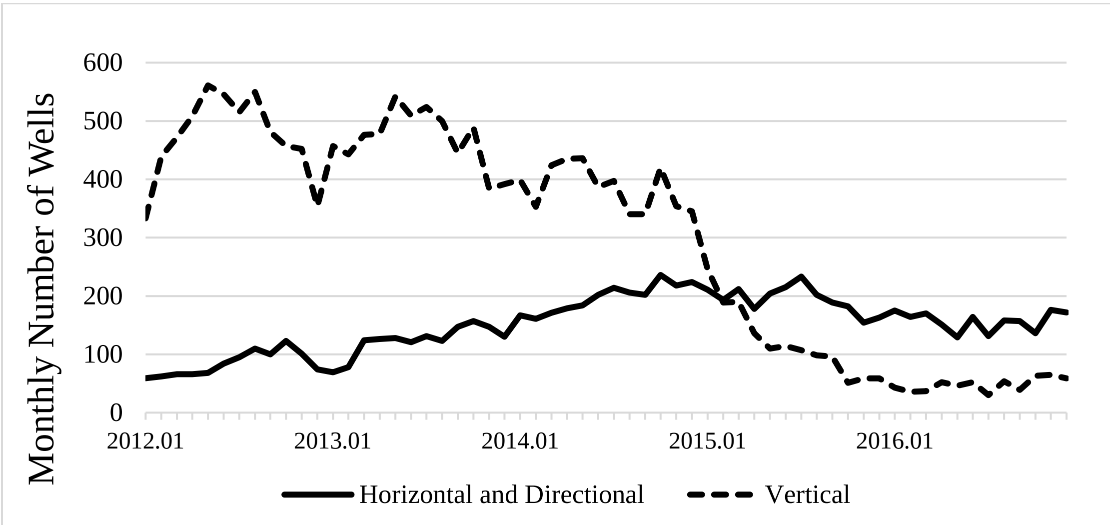
Finally, in the third essay, I examine the influence of an extensive set of fishing site characteristics and amenities on the site choices of private boat anglers fishing in three states in the Gulf of Mexico. I find that a variety of attributes contribute to angler utility through a variety of pathways. These include convenience and ease of access, evidenced by the impact of characteristics such as an onsite bait shop and fuel dock as well as the number of boat ramps and the number of trailer parking spaces. They also include pathways that anglers potentially perceive as signaling site quality, such as whether major fishing tournaments are held onsite, and congestion-related pathways, evidenced by the disutility day-trip anglers were revealed to receive from a fishing site’s proximity to restaurants and lodging, both of which potentially add non-fishing related traffic that increases congestion. I also estimate models using a new site aggregation approach, which makes use of an outside option-fishing site that represents my characterization of all less relevant fishing sites to anglers, but those still within a day's trip. These models provide qualitatively similar results to those that use a full angler choice set in estimation, and improve model run time by nearly one-half to three-fourths across various specifications.

Overall, this dissertation fits into the broader natural resource economics literature that studies the behavior of economic agents. In-depth analyses of specific firm tendencies, strategies, and other behavioral nuances are especially important given the many spillovers that occur due to oil and gas development. Understanding the incentives of both firms and related entities is an important component to guide policymaking regarding new development. They are also important in studies of recreation demand as they enable fisheries managers to understand important factors that influence angler site selection and can improve the quality of a fishing experience. This dissertation attempts to add to the understanding of unconventional oil and gas development and marine recreational fishing on several new dimensions, and using a variety of empirical methods. I recommend future research in these areas to study the value of water to the hydraulic fracturing industry, the many welfare and bargaining implications of asymmetric information that occur during oil and gas lease negotiations, and improvements to the calculation of catch rates and or the generation of high-resolution fish population data, among others.

APPENDIX A

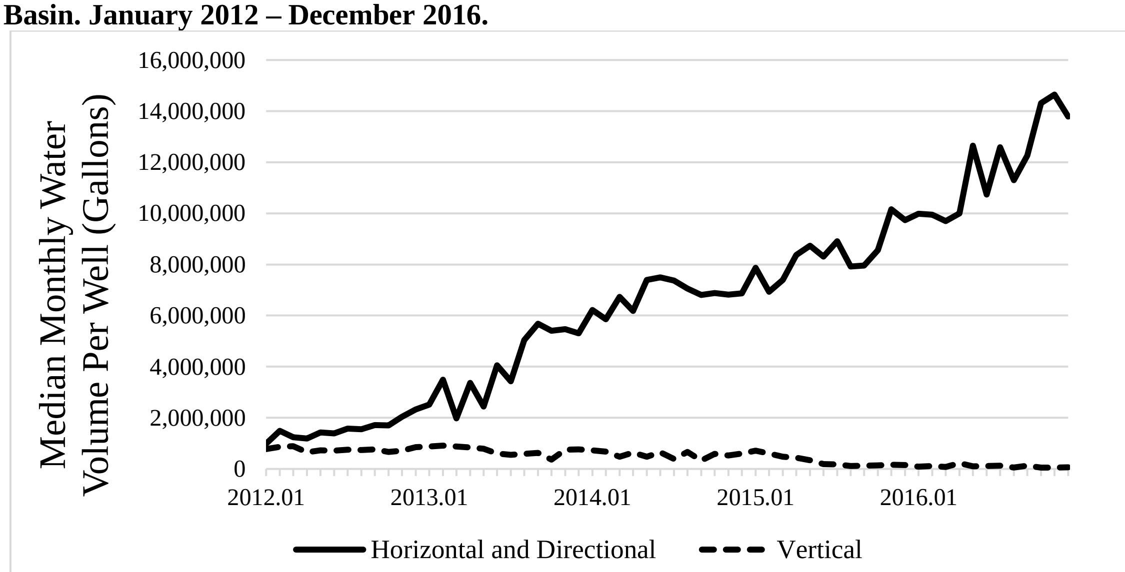
FIGURES AND TABLES

Figure 1. Number of Hydraulically Fractured Wells in the Texas Permian Basin. January 2012 – December 2016.



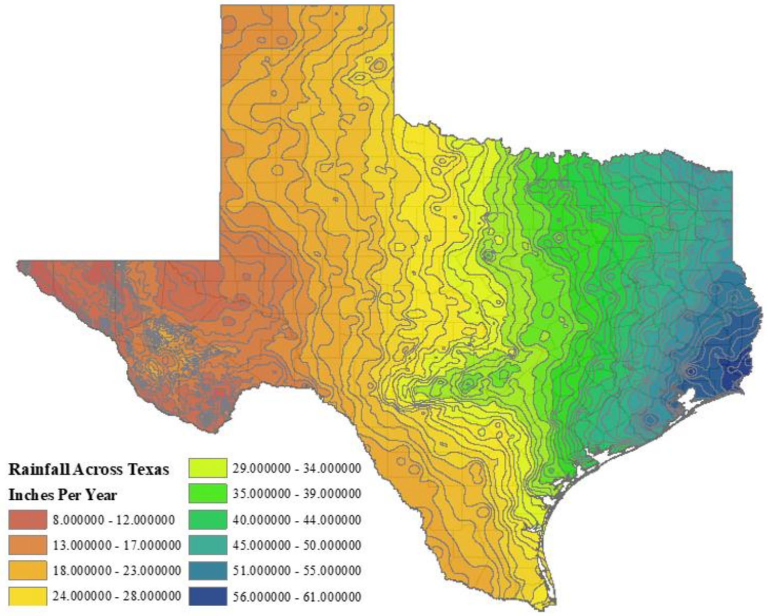
Source: Data from Primary Vision.

Figure 2. Median *Reported* Water Use of Hydraulically Fractured Wells in the Permian Basin. January 2012 – December 2016.



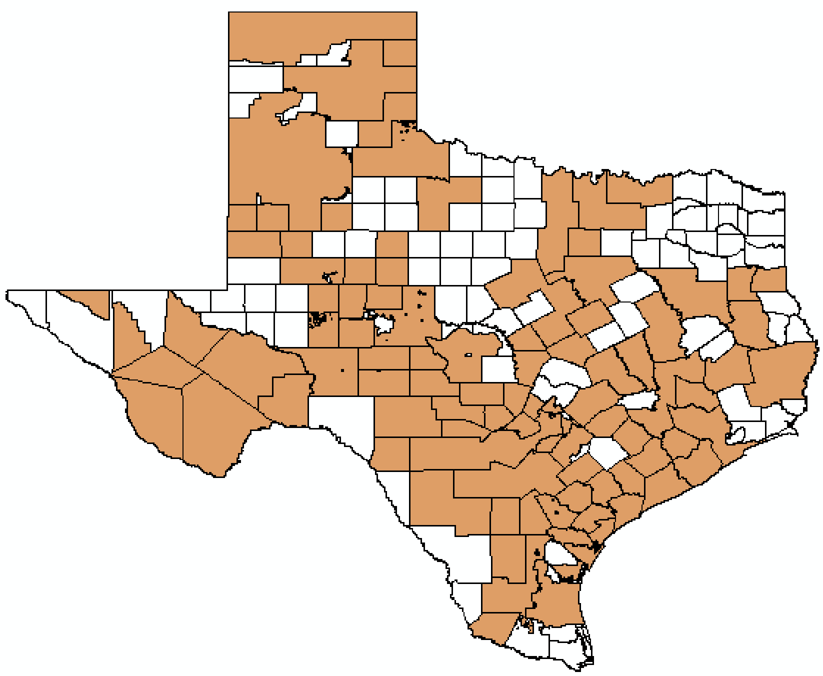
Source: Data from Primary Vision.

Figure 3. Mean Annual Precipitation in Texas. Figure Created by Author in ArcMap.



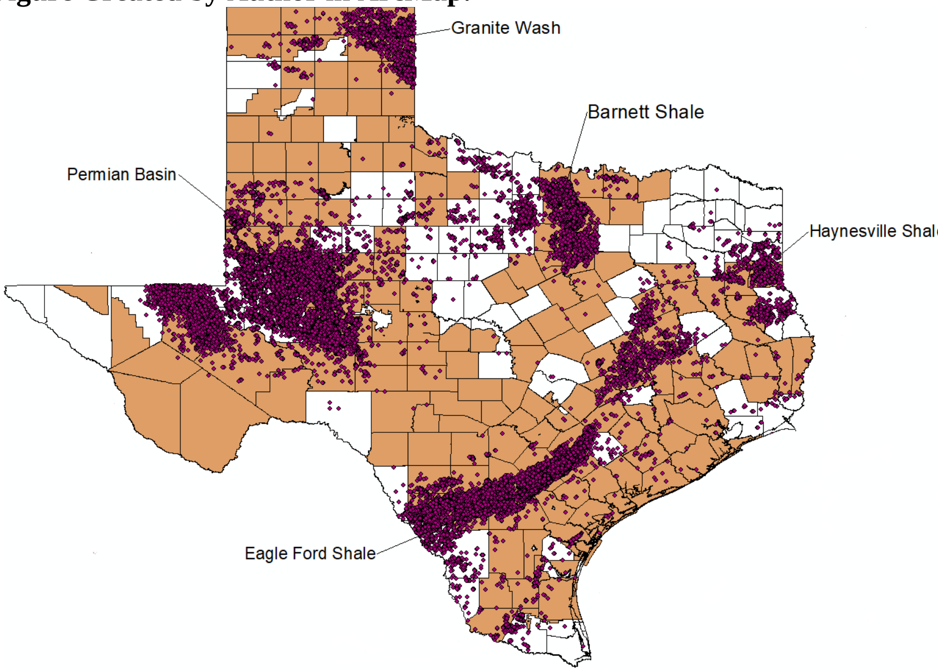
Source: Texas Rainfall Shape File from http://www.twdb.texas.gov/mapping/gisdata.asp.

Figure 4. Map of GCDs in Texas. GCDs in Orange. Figure Created by Author in ArcMap. As of January 2019, There Are 100 Confirmed and 2 Unconfirmed (Pending Election) GCDs in the State. The 100 Confirmed GCDs Cover 180 of the 254 Counties in Texas.



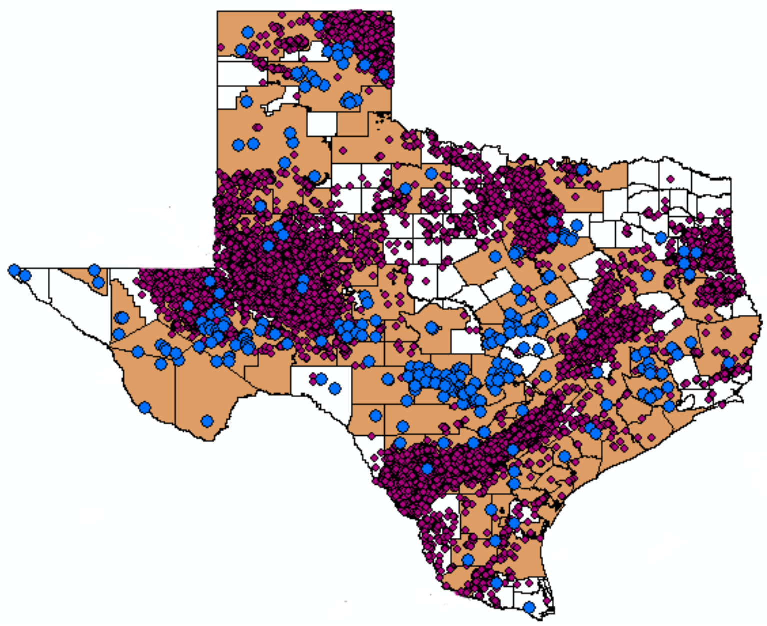
Source: GCD Shape File from https://www.tceq.texas.gov/gis/download-tceq-gis-data.

Figure 5. Map of Texas Counties, GCDs, and Oil and Gas Well Locations. Oil and Gas Wells in Purple. Figure Created by Author in ArcMap.



Source: Oil and Gas Well Location Data from Primary Vision.

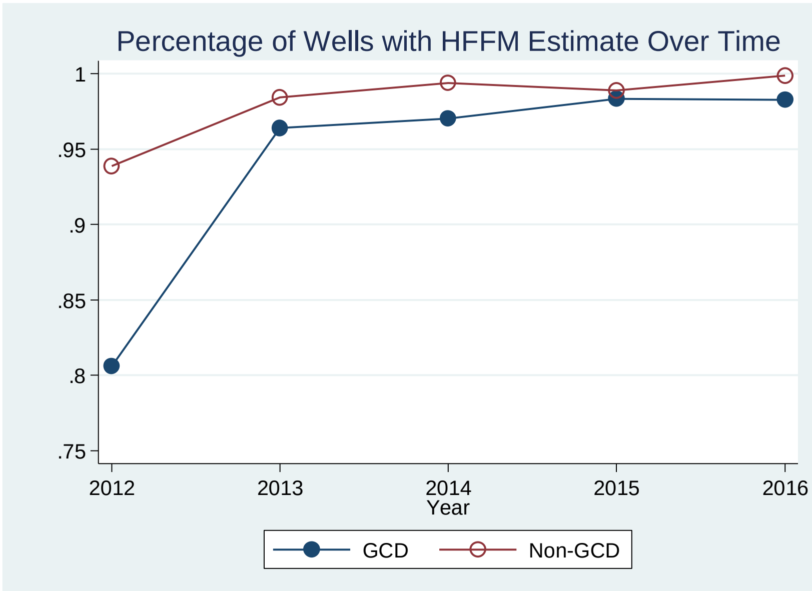
Figure 6. Groundwater Level Monitoring Stations in Texas. Monitoring Stations (Blue Dots) Overlaying Texas Counties, GCD areas, and Oil and Gas Well Locations (Purple Dots). Figure Created by Author in ArcMap.



Source: Groundwater Monitoring Station Locations Obtained from the Texas Water Development Board at https://www.waterdatafortexas.org/groundwater.

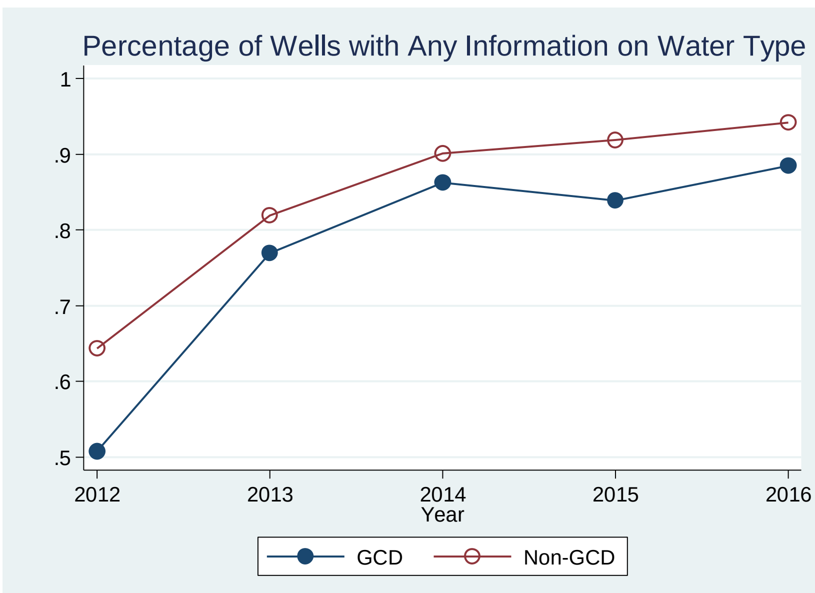
Figure 7. Reporting Over Time by Area.

(a) The Percentage of Well Records for which Primary Vision Estimated a Hydraulic Fracturing Fluid Mass in GCD and Non-GCD Areas.



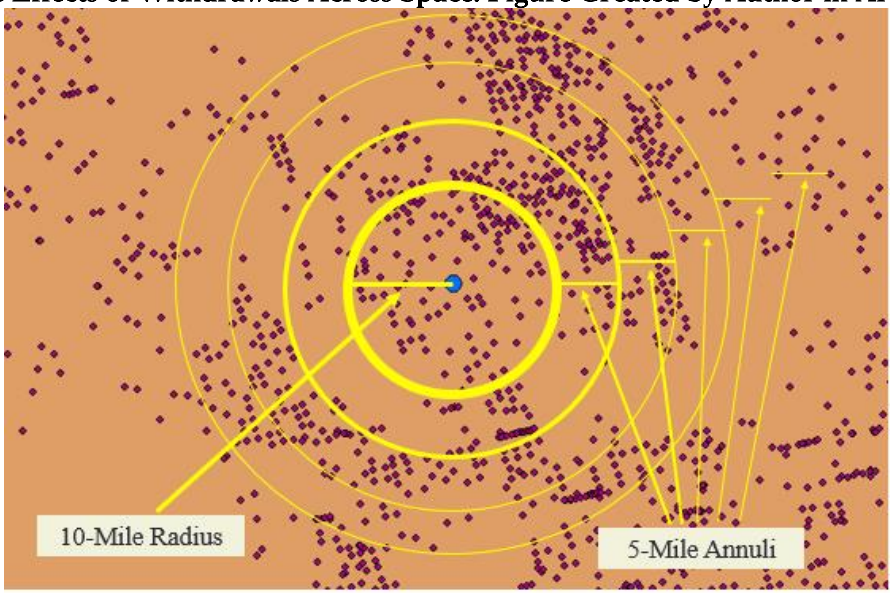
Source: Data from Primary Vision.

(b) The Percentage of Well Records That Contain at Least Some Information on Water Type in GCD and Non-GCD Areas.



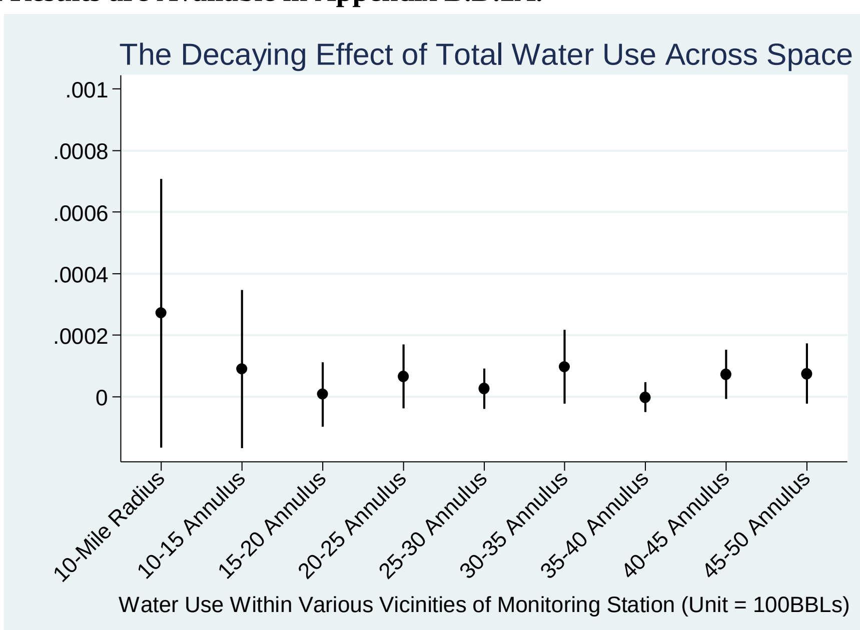
Source: Data from Primary Vision.

Figure 8. Total Water Volumes Used in Hydraulic Fracturing within the Vicinity of a Groundwater Level Monitoring Station. The Thickness of the Lines Reflects the Expected Size of the Effects of Withdrawals Across Space. Figure Created by Author in ArcMap.



Source: Oil and Gas Well Location Data from Primary Vision.

Figure 9. Coefficient Plot of the Effects of Water Use in Hydraulic Fracturing Across Space. Estimates from Equation (2.2). Each Point is Presented with its 95% Confidence Interval. Regression Results are Available in Appendix B.D.2A.



Source: Data from Primary Vision.

Table 1. FracFocus Disclosure Requirements for Texas.

(1) Operator name. (2) Date of completion and hydraulic fracturing treatment(s). (3) County of well. (4) API number. (5) Well name and number. (6) Latitude and longitude of wellhead. (7) Total vertical depth of well. (8) Total volume of water used in the hydraulic fracturing treatment(s) of the well or the type and total volume of the base fluid used in the treatment (if something other than water). (9) Each additive used in the hydraulic fracturing treatments and the trade name, supplier, and a brief description of the intended use or function of each additive in the hydraulic fracturing treatment(s).	(10) Each chemical ingredient used in the hydraulic fracturing treatment(s) of the well that is subject to the requirements of 29 Code of Federal Regulations §1910.1200(g)(2), as provided by the chemical supplier or service company or by the operator, if the operator provides its own chemical ingredients. (11) The actual or maximum concentration of each chemical ingredient in percent by mass. (12) The CAS number for each chemical ingredient listed, if applicable. (13) A supplemental list of all chemicals and their respective CAS numbers, not subject to the requirements of 29 Code of Federal Regulations §1910.1200(g)(2), that were intentionally included in and used for the purpose of creating the hydraulic fracturing treatments for the well.

Source: List Available from House Bill 3328, Texas Legislature.

Table 2. Descriptive Statistics for Wells in GCD and Non-GCD Areas. Median *Reported* Water Volume (Gallons) for All Well Records Over 2012-May 2017 (*Including* Wells from January 2012).

			2012	2013	2014	2015	2016	Through May 2017
GCD Areas	**Horizontal and Directional**	**Water Vol./Well**	4,027,783	4,439,589	5,531,274	9,908,328	8,348,620	>10,900,000
		No. Wells	4,543	5,486	5,909	3,918	2,271	834
	Vertical	**Water Vol./Well**	732,285	737,877	855,418	294,588	69,586	62,106
		No. Wells	3,300	3,290	2,498	885	362	125
	Total No. Wells		7,843	8,776	8,407	4,803	2,633	959
Non-GCD Areas	**Horizontal and Directional**	**Water Vol./Well**	2,540,573	3,442,774	5,567,493	7,131,810	>11,000,000	>11,700,000
		No. Wells	1,332	1,567	2,125	1,759	1,271	417
	Vertical	**Water Vol./Well**	405,501	268,942	195,928	163,873	73,038	50,099
		No. Wells	3,374	3,702	3,233	1,118	457	201
	Total No. Wells		4,706	5,269	5,358	2,877	1,728	618

Source: Data from Primary Vision.

Table 3. Linear Probability Model Results for Reporting Metric 1. Outcome: Hydraulic Fracturing Fluid Mass Calculated (0 or 1)? Each Specification Was Estimated Using Wells in Texas Over February 2012 through May 2017, but *Omitting* Observations for Operators Who Did Not Have Wells in Both GCD and Non-GCD Areas.

	(1)	(2)	(3)	(4)
GCD (0 or 1)	-0.0143*	-0.0098*	-0.0127*	-0.0129*
	(0.00772)	(0.00576)	(0.00719)	(0.00742)
Well Orientation (0 or 1)		-0.0433***	-0.0427***	-0.0414***
		(0.00890)	(0.00940)	(0.00856)
TWV (One Unit = 100k Gallons)		-0.00014***	-0.00013***	-0.00013***
		(0.000033)	(0.000032)	(0.000031)
Refrac (0 or 1)		-0.0031	-0.0023	-0.0019
		(0.00524)	(0.00509)	(0.00508)
#Wells by Operator in County-Month			0.0019***	0.0019***
			(0.00062)	(0.00062)
Operator's 1st Well			0.0346*	0.0342*
			(0.01809)	(0.01791)
Operator's 1st Well in County			0.0144*	0.0146*
			(0.00807)	(0.00811)
Cumulative #Wells in County	No	No	Yes	Yes
Cumulative #Wells by Largest 5 in County	No	No	Yes	Yes
Cumulative #Wells by Operator in County	No	No	Yes	Yes
#Wells by Largest 5 in County-Month	No	No	Yes	Yes
Total #Wells by Operator in County	No	No	Yes	Yes
Drought Controls	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes
N	47521	47521	47521	47521
R-Squared	0.070	0.091	0.095	0.097
Operator FEs	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on county.* p<0.10, ** p<0.05, *** p<0.01

Table 4. Linear Probability Model Results for Reporting Metric 2. Outcome: *Any* Information on Water Type in Completion Report (0 or 1)? Each Specification Was Estimated Using Wells in Texas Over February 2012 Through May 2017, but *Omitting* Observations for Operators Who Did Not Have Wells in Both GCD and Non-GCD Areas.

	(1)	(2)	(3)	(4)
GCD (0 or 1)	-0.0366**	-0.0290*	-0.0304*	-0.0280*
	(0.01779)	(0.01471)	(0.01611)	(0.01590)
Well Orientation (0 or 1)		-0.0842***	-0.0875***	-0.0881***
		(0.01598)	(0.01565)	(0.01520)
TWV (One Unit = 100k Gallons)		0.00009	0.00009	0.00009
		(0.000057)	(0.000055)	(0.000055)
Refrac (0 or 1)		0.0025	0.0005	0.0007
		(0.01599)	(0.01514)	(0.01520)
#Wells by Operator in County-Month			0.0060	0.0060
			(0.00436)	(0.00437)
Operator's 1st Well			-0.0229	-0.0235
			(0.02632)	(0.02630)
Operator's 1st Well in County			0.0219	0.0216
			(0.01363)	(0.01373)
Cumulative #Wells in County	No	No	Yes	Yes
Cumulative #Wells by Largest 5 in County	No	No	Yes	Yes
Cumulative #Wells by Operator in County	No	No	Yes	Yes
#Wells by Largest 5 in County-Month	No	No	Yes	Yes
Total #Wells by Operator in County	No	No	Yes	Yes
Drought Controls	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes
N	47521	47521	47521	47521
R-Squared	0.119	0.133	0.137	0.139
Operator FEs	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on county. * p<0.10, ** p<0.05, *** p<0.01

Table 5. Naïve Fixed Effects Model Results. Outcome: Distance to Groundwater Level (Feet). Units of TWV are in 100BBLs (4,200 gallons).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00040	0.00041	0.00149***	0.00040	
	(0.000251)	(0.000249)	(0.000182)	(0.000251)	
TWV10 x Permian					0.00012***
					(0.000040)
TWV10 x Eagle Ford					0.00111**
					(0.000487)
Drought	No	Yes	Yes	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	15301	15014	5805	15014	15014
# Stations	267	267	106	267	267
R-Squared	0.118	0.128	0.186	0.138	0.147
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 6. Fixed Effects Model Results with Lags for the Short-Term Effects of Water Use in Hydraulic Fracturing. Lags Indicated by (-). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00005	0.00006	0.00031***	0.00005	0.00005
	(0.000048)	(0.000050)	(0.000060)	(0.000049)	(0.000049)
TWV10 (-1)	0.00006	0.00007*	0.00029***	0.00006	0.00006
	(0.000041)	(0.000040)	(0.000034)	(0.000042)	(0.000042)
TWV10 (-2)	0.00008*	0.00009*	0.00036***	0.00008	0.00008
	(0.000048)	(0.000047)	(0.000036)	(0.000048)	(0.000048)
TWV10 (-3)	0.00009***	0.00009***	0.00019***	0.00008***	0.00008***
	(0.000031)	(0.000029)	(0.000047)	(0.000028)	(0.000027)
TWV10 (-4)	0.00012***	0.00012***	0.00027***	0.00012***	0.00011***
	(0.000035)	(0.000035)	(0.000023)	(0.000034)	(0.000034)
TWV10 (-5)	0.00014***	0.00014***	0.00033***	0.00014***	0.00013**
	(0.000048)	(0.000050)	(0.000049)	(0.000049)	(0.000050)
TWV10 (-6)	0.00018	0.00018	0.00052***	0.00018	0.00017
	(0.000119)	(0.000118)	(0.000096)	(0.000119)	(0.000120)
Drought Index	No	Yes	Yes	Yes	Yes
Drought Index Lags - 6	No	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Rain Lags - 6	No	No	No	No	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	13706	13445	5174	13445	13349
# Stations	258	258	103	258	254
R-Squared	0.121	0.134	0.200	0.160	0.162
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

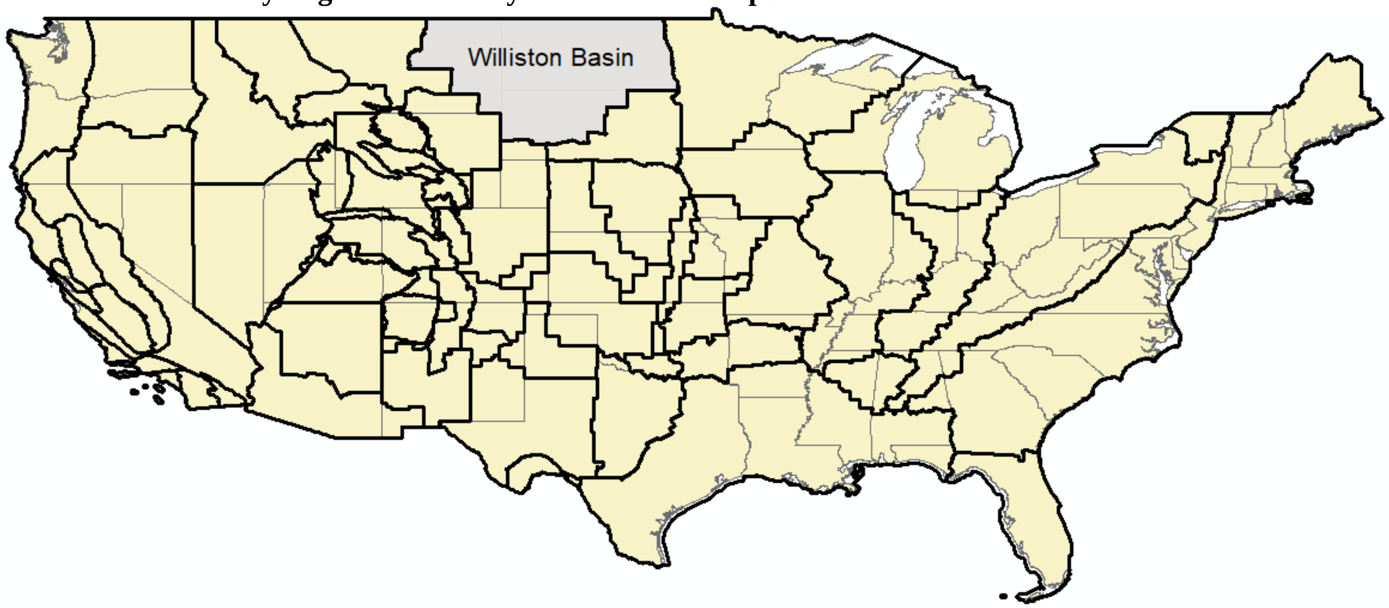
Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

Table 7. Fixed Effects Model Results for the Cumulative Effects of Water Use in Hydraulic Fracturing. Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)
CTWV10 (Unit = 100BBLs)	0.0000185***		
	(0.00000543)		
CTWV10 x GCD		0.0000183***	
		(0.00000531)	
CTWV10 x Permian			0.0000129***
			(0.00000389)
CTWV10 x Eagle Ford			0.0000239***
			(0.00000688)
Drought Index	Yes	Yes	Yes
Drought Index Lags - 6	Yes	Yes	Yes
Rain	Yes	Yes	Yes
Rain Lags - 6	Yes	Yes	Yes
Population	Yes	Yes	Yes
Total Corn Acres	Yes	Yes	Yes
Irrigated Corn Acres	Yes	Yes	Yes
Total Cotton Acres	Yes	Yes	Yes
Cotton Acres Irrigated	Yes	Yes	Yes
Total Sorghum Acres	Yes	Yes	Yes
Sorghum Acres Irrigated	Yes	Yes	Yes
Total Wheat Acres	Yes	Yes	Yes
Wheat Acres Irrigated	Yes	Yes	Yes
Total Rice Acres	Yes	Yes	Yes
Constant	Yes	Yes	Yes
N	13349	13349	13349
# Stations	254	254	254
R-Squared	0.163	0.163	0.164
Station FEs	YES	YES	YES
Year-Month FEs	YES	YES	YES

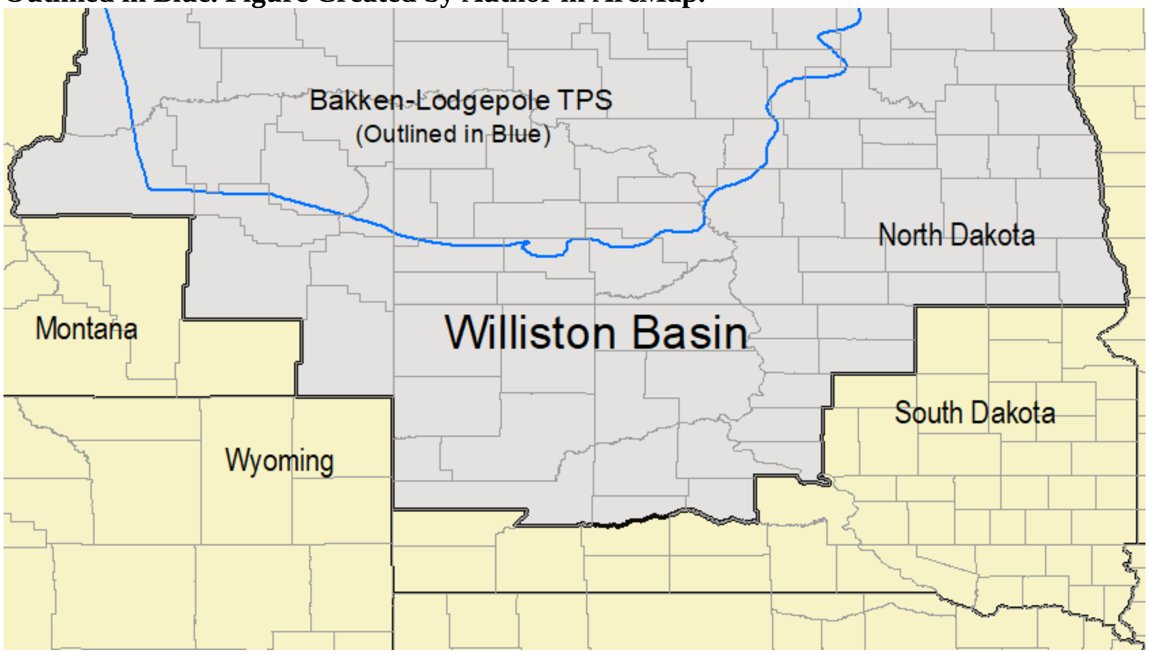
Standard errors in parentheses. Clustered on monitoring station.* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Figure 10. The Geological Provinces of the Unites States. The Provinces (Excluding those Alaska) are Outlined in Black, with the Williston Basin Shaded in Gray. Figure Created by Author in ArcMap.



Sources: U.S. Geological Province Boundaries Shape File Obtained from https://catalog.data.gov/dataset/national-oil-and-gas-assessment-province-boundaries-through-2012. See https://certmapper.cr.usgs.gov/data/apps/noga-drupal/ for Additional Information on the NOGA Provinces. U.S. State Boundaries Shape File Obtained from https://catalog.data.gov/dataset/tiger-line-shapefile-2017-nation-u-s-current-state-and-equivalent-national.

Figure 11. A Map of the Williston Basin Province Boundary. The Basin Covers parts of Montana, North Dakota, and South Dakota. Bakken-Lodgepole Total Petroleum System Outlined in Blue. Figure Created by Author in ArcMap.



Sources: U.S. County Boundaries Shape File Obtained from https://catalog.data.gov/dataset/tiger-line-shapefile-2016-nation-u-s-current-county-and-equivalent-national-shapefile. U.S. State Boundaries Shape File Obtained from https://catalog.data.gov/dataset/tiger-line-shapefile-2017-nation-u-s-current-state-and-equivalent-national. Williston Basin and Bakken-Lodgepole TPS Boundaries Shape Files Obtained from https://certmapper.cr.usgs.gov/data/apps/noga-data/?provcode=5031.

Figure 12. The “Cake-Layer” Nature of the Williston Basin. This Figure Illustrates How Horizontal Wells at Various Depths Are Needed to Produce Oil and Gas from All Formations and *Within* Each Formation. Figure Created by the Author in Microsoft Office.

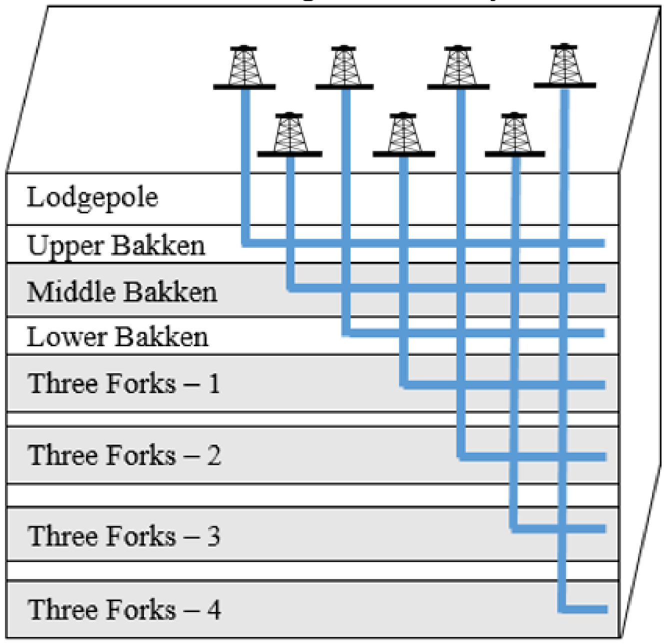
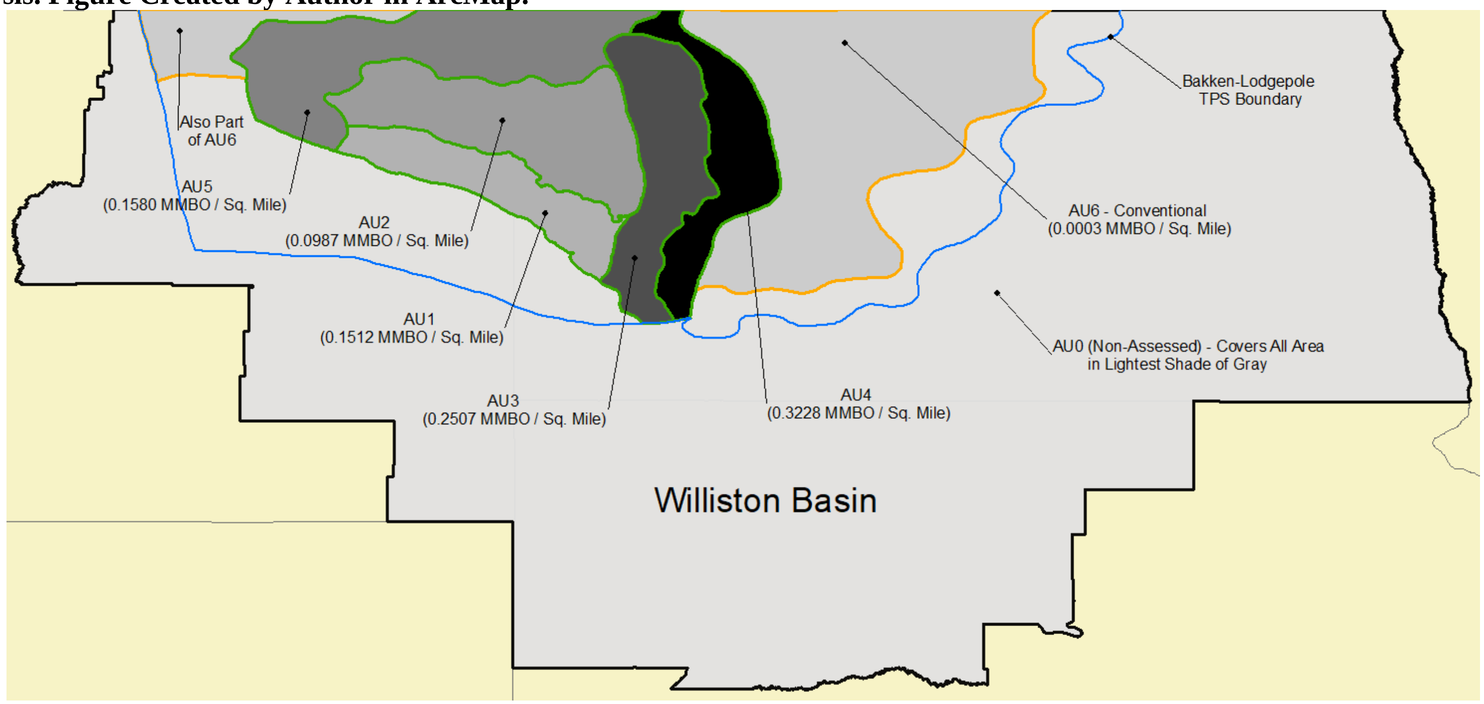
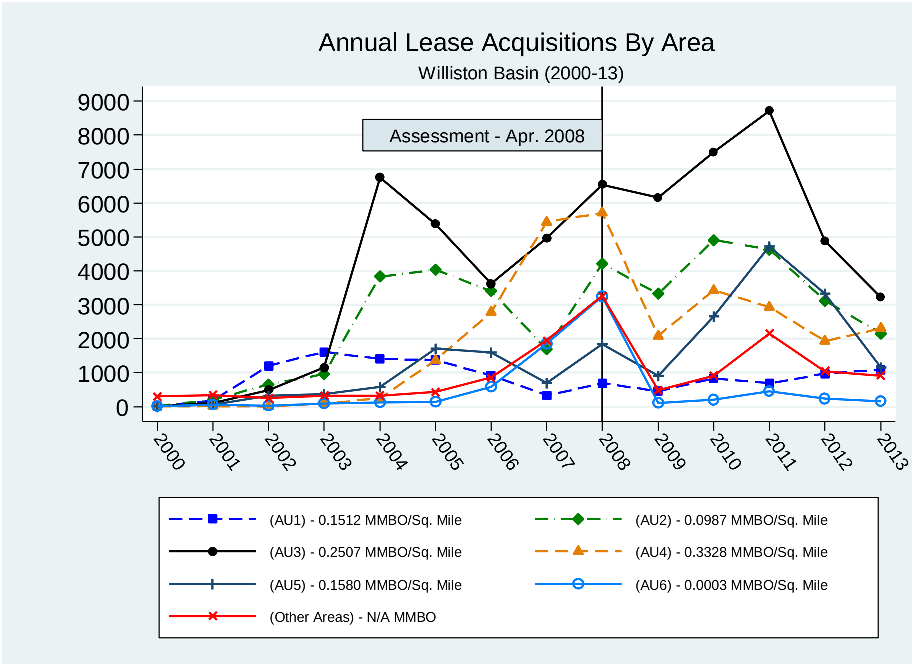


Figure 13. A Visual for the Quasi-Experimental Design. This Figure Shows the Boundaries of the Bakken-Lodgepole TPS (Outlined in Faint Blue) within the Williston Basin, the Five Unconventional AUs (Outlined in Green), and One Conventional AU (Outlined in Orange) Defined for the 2008 Assessment of Undiscovered Oil and Gas Resources in the U.S. Portion of the Bakken-Lodgepole TPS. The Outermost Green Line Defines the Oil Generation Window for the Upper Shale Member of the Formation (the Extent of the Lower Members Fits Within this Boundary). The Conventional AU Shown Was Defined External to the Oil Generation Window. The Second Conventional AU (Within the Lodgepole Formation) is Not Shown Since it Was Not Assessed. The Six Assessed Areas (AU1-AU6) Are Shaded Based on Resource Abundance (Darker Shades Indicate More Estimated Resources). AU0, the Area Outside of this Window in the Lightest Shade of Gray (but Within the Basin), Was Not Assessed and Constitutes the Counterfactual for My Analysis. Figure Created by Author in ArcMap.



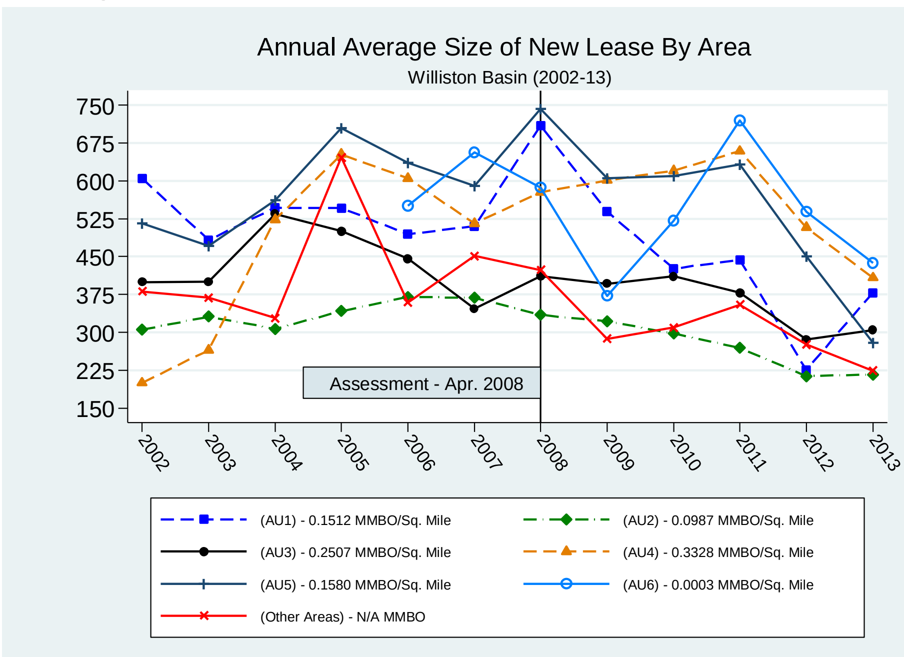
Sources: Numbers in the Figure Based on Estimates from Pollastro et al. (2008). U.S. State Boundaries Shape File Obtained from https://catalog.data.gov/dataset/tiger-line-shapefile-2017-nation-u-s-current-state-and-equivalent-national. Williston Basin, Bakken-Lodgepole TPS, and AU boundaries Shape Files Obtained from https://certmapper.cr.usgs.gov/data/apps/noga-data/?provcode=5031.

Figure 14. Number of Lease Transactions by Area.



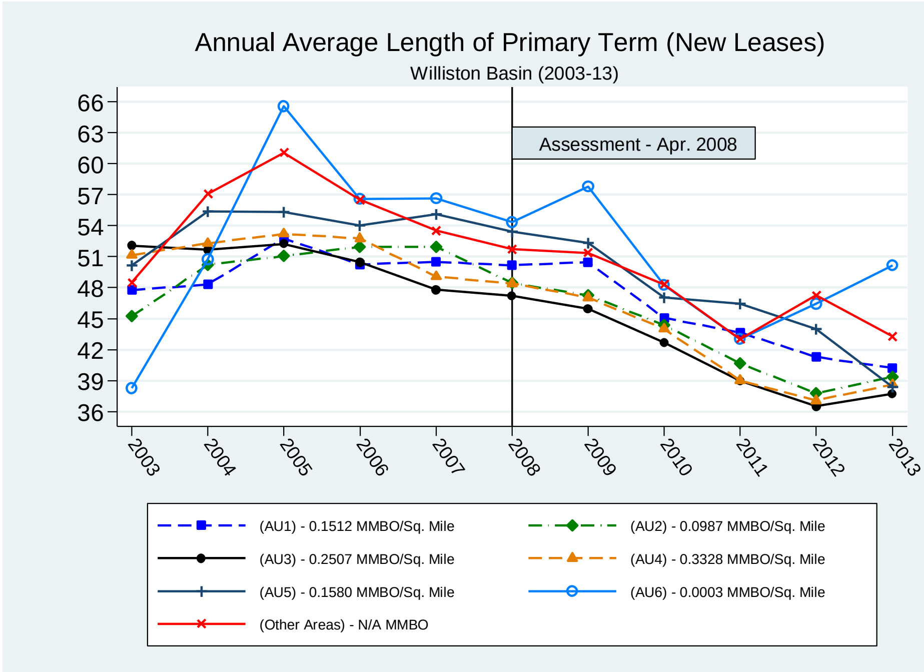
Source: Data from DrillingInfo.

Figure 15. Average Size of New Lease by Area. Lease Size in Acres.



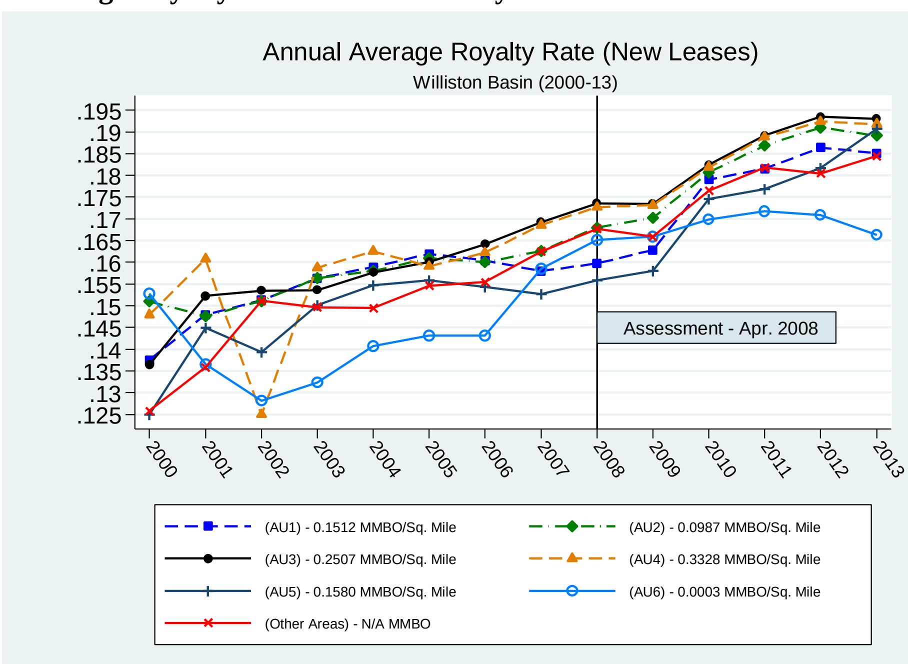
Source: Data from DrillingInfo.

Figure 16. Average Primary Term Length of New Lease by Area. Term Length in Months.



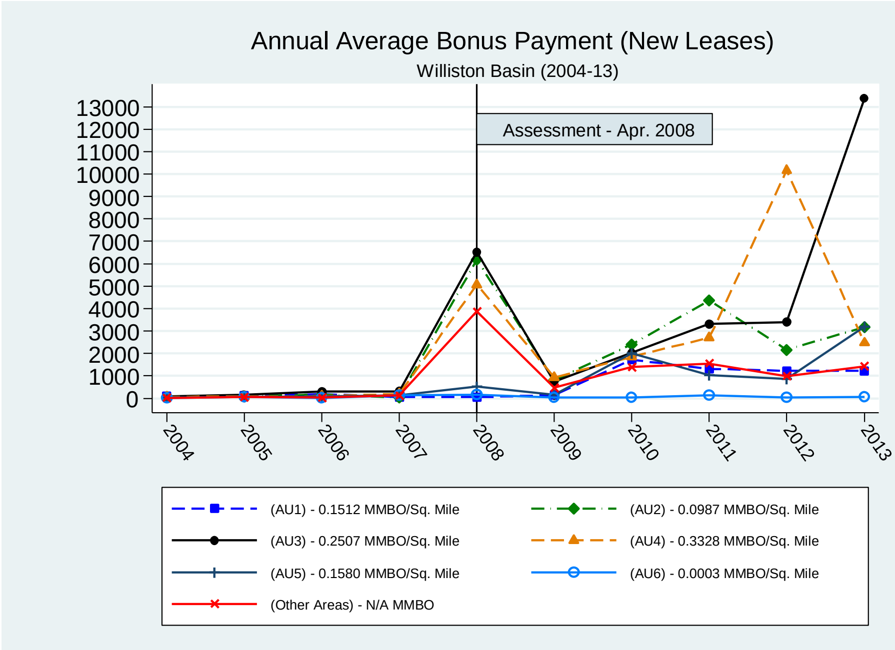
Source: Data from DrillingInfo.

Figure 17. Average Royalty Rate of New Lease by Area.



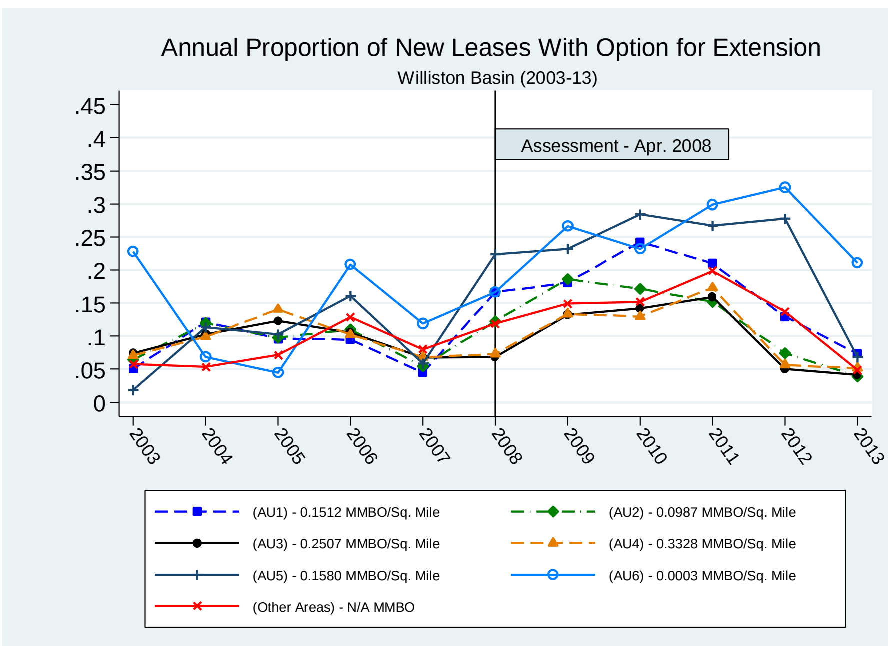
Source: Data from DrillingInfo.

Figure 18. Average Bonus Payment of New Lease by Area. Payment in Dollars Per Acre.



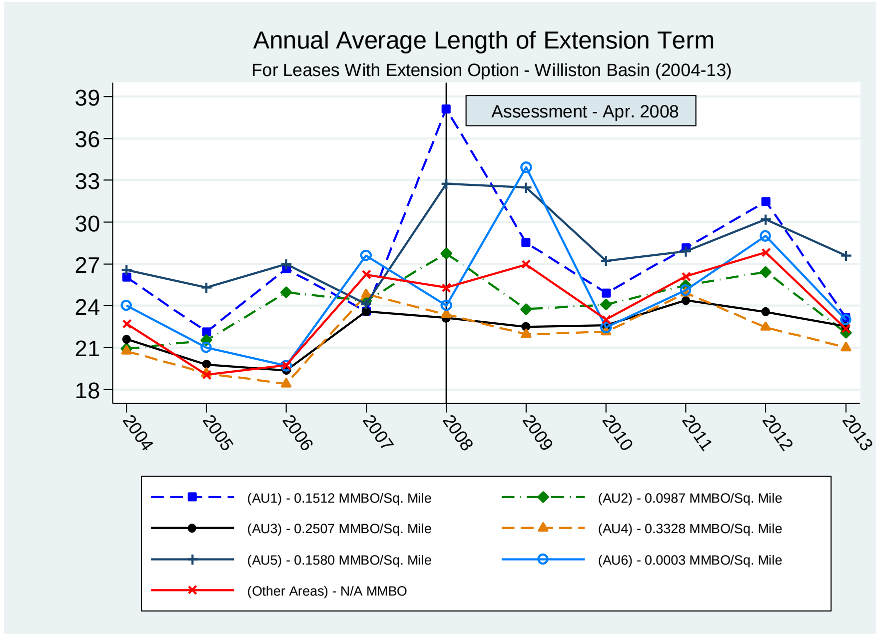
Source: Data from DrillingInfo.

Figure 19. Proportion of New Leases Containing the Option to Extend the Lease Beyond the Primary Term.



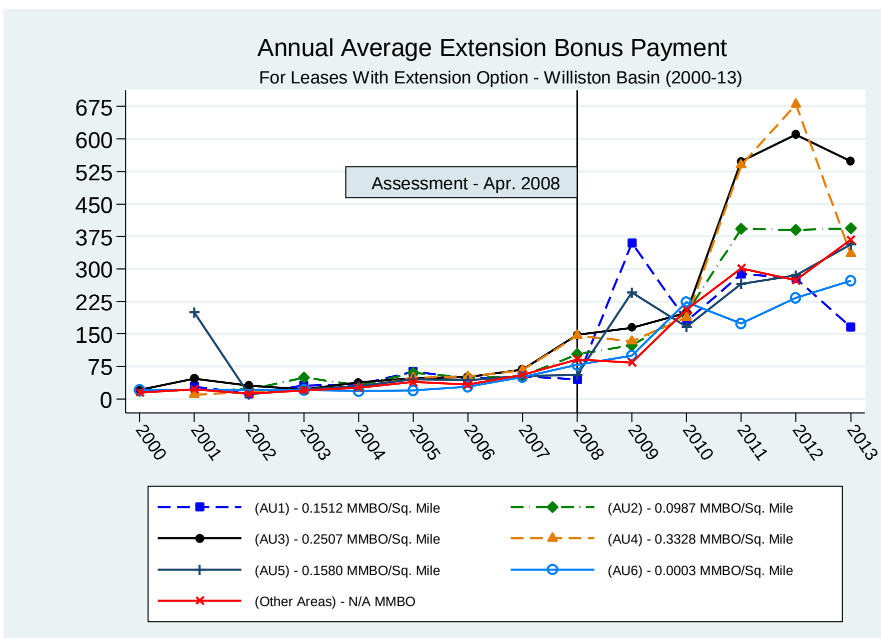
Source: Data from DrillingInfo.

Figure 20. Average Length of Extension Term for Leases with an Extension Option. Term Length in Months.



Source: Data from DrillingInfo.

Figure 21. Average Extension Bonus Payment for Leases with an Extension Option. Payment in Dollars Per Acre.



Source: Data from DrillingInfo.

Table 8. Results of the 2008 Assessment of Undiscovered Oil and Gas Resources in the Bakken-Lodgepole TPS of the Williston Basin. [Note: MMBO, Million Barrels of Oil; BCFG, Billion Cubic Feet of Natural Gas; MMBNGL, Million Barrels of Natural Gas Liquids. F95 Represents a 95-Percent Chance of at Least the Amount Tabulated; Other Fractiles are Defined similarly.].

AU ID and Name	Area (Mi.^{2})	Total Undiscovered Resources												
		Oil (MMBO)					*Gas (BCFG)*				*NGL (MMBNGL)*			
		F95	**F50**	**F5**	**Mean**	**Mean/ Sq. Mile**	**F95**	**F50**	**F5**	**Mean**	**F95**	**F50**	**F5**	**Mean**
AU1 – Elm Coulee-Billings Nose	2,711.09	374	410	450	***410***	**0.1512**	118	198	332	***208***	8	16	29	***17***
AU2 – Central Basin-Poplar Dome	4,915.90	394	482	589	***485***	**0.0987**	134	233	403	***246***	10	18	35	***20***
AU3 – Nesson-Little Knife Structural	3,626.36	818	908	1,007	***909***	**0.2507**	260	438	738	***461***	19	34	64	***37***
AU4 – Eastern Expulsion Threshold	3,033.09	864	971	1,091	***973***	**0.3228**	278	469	791	***493***	20	37	68	***39***
AU5 – Northwest Expulsion Threshold	5,495.37	613	851	1,182	***868***	**0.1580**	224	411	754	***440***	16	32	0	***0***
AU6 (Conventional) – Middle Sandstone Member	13,454.07				***4***	**0.0003**				***2***				***0***

Source: Table Created by the Author; Assessment Results Available in Pollastro et al. (2008).

Table 9. The Effects of the 2008 USGS Assessment on the *Number of New Leases* and the *Average Number of Acres Per Lease.* Note: Observations are at the Area-Month Level Over January 2000 - March 2013, and the Coefficient for Each Treatment Term is Interpreted as Relative to New Lease Acquisitions in the Unassessed Area of the Williston Basin. The Amount of Oil Resources Estimated for Each Area is Shown in Parentheses Next to Each Interaction Term (in Million Barrels of Oil Per Square Mile).

	Total #Leases	Total #Leases	Avg. #Acres Per Lease	Avg. #Acres Per Lease
Post x AU1 (0.1512 MMBO/Sq. Mile)	-76.92***	-47.32***	78.33***	31.41***
	(1.341)	(3.026)	(4.948)	(7.555)
Post x AU2 (0.0987 MMBO/Sq. Mile)	107.51***	86.17***	60.71***	24.51***
	(0.556)	(1.330)	(3.448)	(2.265)
Post x AU3 (0.2507 MMBO/Sq. Mile)	220.44***	173.40***	87.81***	36.55**
	(1.250)	(2.752)	(12.019)	(9.893)
Post x AU4 (0.3328 MMBO/Sq. Mile)	40.46**	-22.20**	214.78***	161.85***
	(11.007)	(6.044)	(23.403)	(28.631)
Post x AU5 (0.1580 MMBO/Sq. Mile)	105.22***	52.24***	99.43***	50.99**
	(2.652)	(4.391)	(14.988)	(17.061)
Post x AU6 (0.0003 MMBO/Sq. Mile)	-64.01***	-77.61***	93.77***	-6.89
	(7.364)	(4.647)	(13.665)	(14.373)
#Acres Largest 10 Operators (-1)	No	Yes	No	Yes
#Acres Largest 10 Operators (-2)	No	Yes	No	Yes
#Acres Largest 10 Operators (-3)	No	Yes	No	Yes
#Acres Largest 10 Operators (-4)	No	Yes	No	Yes
#Acres Largest 10 Operators (-5)	No	Yes	No	Yes
Constant	Yes	Yes	Yes	Yes
Assessment Unit FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
N	1013	977	1013	977
R-Squared	0.632	0.842	0.378	0.398

Standard errors in parentheses. Clustered on area. * p<0.10, ** p<0.05, *** p<0.01

Table 10. The Effects of the 2008 USGS Assessment on the *Average Primary Term Length* and *Average Royalty Rate*. The Royalty Rate Can be Interpreted as the Percent of Production to be Paid to Lessor. Note: Observations are at the Area-Month Level Over January 2000 - March 2013, and the Coefficient for Each Treatment Term is Interpreted as Relative to New Lease Acquisitions in the Unassessed Area of the Williston Basin. The Amount of Oil Resources Estimated for Each Area is Shown in Parentheses Next to Each Interaction Term (in Million Barrels of Oil Per Square Mile).

	Avg. Primary Term Length	Avg. Primary Term Length	Avg. Royalty Rate	Avg. Royalty Rate
Post x AU1 (0.1512 MMBO/Sq. Mile)	5.55620***	4.41278***	-0.00130**	-0.00133**
	(0.28096)	(0.19882)	(0.00036)	(0.00048)
Post x AU2 (0.0987 MMBO/Sq. Mile)	3.99154***	3.34556***	0.00315***	0.00290***
	(0.04962)	(0.03990)	(0.00026)	(0.00028)
Post x AU3 (0.2507 MMBO/Sq. Mile)	2.15689***	1.17454***	0.00548***	0.00415***
	(0.38699)	(0.28516)	(0.00019)	(0.00040)
Post x AU4 (0.3328 MMBO/Sq. Mile)	0.99210***	-0.80939	0.00499***	0.00384***
	(0.24786)	(0.57437)	(0.00040)	(0.00055)
Post x AU5 (0.1580 MMBO/Sq. Mile)	2.85750***	1.64118**	0.00087**	0.00005
	(0.48085)	(0.50835)	(0.00033)	(0.00060)
Post x AU6 (0.0003 MMBO/Sq. Mile)	1.0173**	-1.9547**	0.0299***	0.0306***
	(0.35450)	(0.64020)	(0.00018)	(0.00036)
#Acres Largest 10 Operators (-1)	No	Yes	No	Yes
#Acres Largest 10 Operators (-2)	No	Yes	No	Yes
#Acres Largest 10 Operators (-3)	No	Yes	No	Yes
#Acres Largest 10 Operators (-4)	No	Yes	No	Yes
#Acres Largest 10 Operators (-5)	No	Yes	No	Yes
Constant	Yes	Yes	Yes	Yes
Assessment Unit FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
N	1013	977	1013	977
R-Squared	0.768	0.786	0.964	0.968

Standard errors in parentheses. Clustered on area. * p<0.10, ** p<0.05, *** p<0.01

Table 11. The Effects of the 2008 USGS Assessment on the *Average Bonus Payment* and the *Proportion of New Leases with an Option to Extend the Primary Term.* The Bonus Payment is in Dollars Per Acre Paid to the Lessor. Note: Observations are at the Area-Month Level Over January 2000 - March 2013, and the Coefficient for Each Treatment Term is Interpreted as Relative to New Lease Acquisitions in the Unassessed Area of the Williston Basin. The Amount of Oil Resources Estimated for Each Area is Shown in Parentheses Next to Each Interaction Term (in Million Barrels of Oil Per Square Mile).

	Avg. Bonus Payment	Avg. Bonus Payment	Prop. Leases with Ext. Option	Prop. Leases with Ext. Option
Post x AU1 (0.1512 MMBO/Sq. Mile)	-183.976***	-233.213**	0.017***	0.008
	(2.6415)	(93.1015)	(0.0028)	(0.0049)
Post x AU2 (0.0987 MMBO/Sq. Mile)	2407.700***	2453.843***	-0.055***	-0.039***
	(3.4257)	(25.2861)	(0.0008)	(0.0017)
Post x AU3 (0.2507 MMBO/Sq. Mile)	3319.452***	3420.013***	-0.095***	-0.080***
	(6.4661)	(84.9498)	(0.0056)	(0.0052)
Post x AU4 (0.3328 MMBO/Sq. Mile)	5682.054***	5828.016***	-0.080***	-0.062***
	(38.1557)	(247.0504)	(0.0086)	(0.0099)
Post x AU5 (0.1580 MMBO/Sq. Mile)	392.616***	562.888***	0.063***	0.081***
	(8.4641)	(110.7056)	(0.0067)	(0.0054)
Post x AU6 (0.0003 MMBO/Sq. Mile)	-839.690***	-711.197***	0.167***	0.194***
	(30.2286)	(93.5781)	(0.0039)	(0.0089)
#Acres Largest 10 Operators (-1)	No	Yes	No	Yes
#Acres Largest 10 Operators (-2)	No	Yes	No	Yes
#Acres Largest 10 Operators (-3)	No	Yes	No	Yes
#Acres Largest 10 Operators (-4)	No	Yes	No	Yes
#Acres Largest 10 Operators (-5)	No	Yes	No	Yes
Constant	Yes	Yes	Yes	Yes
Assessment Unit FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
N	1013	977	1013	977
R-Squared	0.178	0.179	0.424	0.441

Standard errors in parentheses. Clustered on area. * p<0.10, ** p<0.05, *** p<0.01

Table 12. The Effects of the 2008 USGS Assessment on the *Average Extension Term Length* and *Average Extension Bonus Payment.* The Extension Term Length is in Months and the Extension Bonus Payment is in Dollars Per Acre Paid to the Lessor. Note: Observations are at the Area-Month Level Over January 2000 - March 2013, and the Coefficient for Each Treatment Term is Interpreted as Relative to New Lease Acquisitions in the Unassessed Area of the Williston Basin. The Amount of Oil Resources Estimated for Each Area is Shown in Parentheses Next to Each Interaction Term (in Million Barrels of Oil Per Square Mile).

	Avg. Ext. Term Length	Avg. Ext. Term Length	Avg. Ext. Bonus Payment	Avg. Ext. Bonus Payment
Post x AU1 (0.1512 MMBO/Sq. Mile)	-1.82***	-1.99***	22.73***	23.87***
	(0.0344)	(0.1726)	(0.0848)	(2.3638)
Post x AU2 (0.0987 MMBO/Sq. Mile)	-4.49***	-4.75***	46.42***	46.97***
	(0.0486)	(0.0975)	(0.0610)	(0.6896)
Post x AU3 (0.2507 MMBO/Sq. Mile)	-3.71***	-3.96***	122.76***	122.39***
	(0.1180)	(0.1688)	(0.5982)	(2.1339)
Post x AU4 (0.3328 MMBO/Sq. Mile)	-1.35**	-1.49*	141.30***	138.83***
	(0.3802)	(0.6652)	(0.6204)	(4.9146)
Post x AU5 (0.1580 MMBO/Sq. Mile)	2.36***	2.28***	22.92***	25.85***
	(0.0635)	(0.3101)	(0.1107)	(2.9450)
Post x AU6 (0.0003 MMBO/Sq. Mile)	6.53***	6.02***	-49.63***	-53.31***
	(0.1133)	(0.3635)	(0.7188)	(3.8138)
#Acres Largest 10 Operators (-1)	No	Yes	No	Yes
#Acres Largest 10 Operators (-2)	No	Yes	No	Yes
#Acres Largest 10 Operators (-3)	No	Yes	No	Yes
#Acres Largest 10 Operators (-4)	No	Yes	No	Yes
#Acres Largest 10 Operators (-5)	No	Yes	No	Yes
Constant	Yes	Yes	Yes	Yes
Assessment Unit FEs	Yes	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes	Yes
N	1013	977	1013	977
R-Squared	0.424	0.395	0.565	0.562

Standard errors in parentheses. Clustered on area. * p<0.10, ** p<0.05, *** p<0.01

Figure 22. Example of MRIP Site Register Page for a Fishing Site in Florida.

Site Information

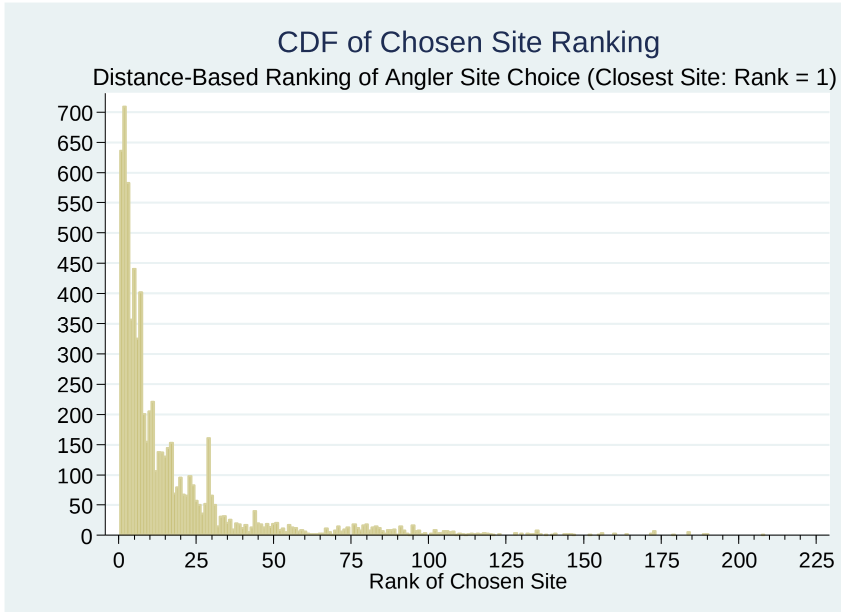
Site | Pressure | Map

Site ID:	0338
Site Name:	JUPITER SEASPORT
Site Status:	Retired
State:	FLORIDA
County:	PALM BEACH
County Code:	099
Latitude:	26:56.761
Longitude:	80:5.047
New?:	N
Address:	1095 N HWY A1A JUPITER, FLORIDA 33477
Contact Name:	TERRY WHITE
Contact Phone:	5615750006

Tackle Shops:	NO	**Headboat Only:**	NO
Fish Cleaning Stations:	YES	**Lighting At Night:**	NO
Retail Bait:	YES	**Can We Interview?:**	YES
Boat Storage:	NO	**Is Site Safe For 2 Samplers At Night:**	YES
Boat Maintenance/Repair:	YES	**Is Fee Charged To The Public For Use Of Site:**	NO
Fuel Dock:	YES	**Fishing Activity Affected by Tide:**	NO
Restaurant (Onsite/Immediate Vicinity):	NO	**# of Boat Slips:**	21
Lodging (Onsite/Immediate Vicinity):	NO	**# of Car Parking Spaces:**	114
Major Tournaments:	NO	**# of Trailer Parking Spaces:**	0
Private Access?:	NO	**# of Ramps:**	0
Shore Area:	NA	**# HB Using Site:**	0
Shore Mode:	N/A	**#CB Using Site:**	7

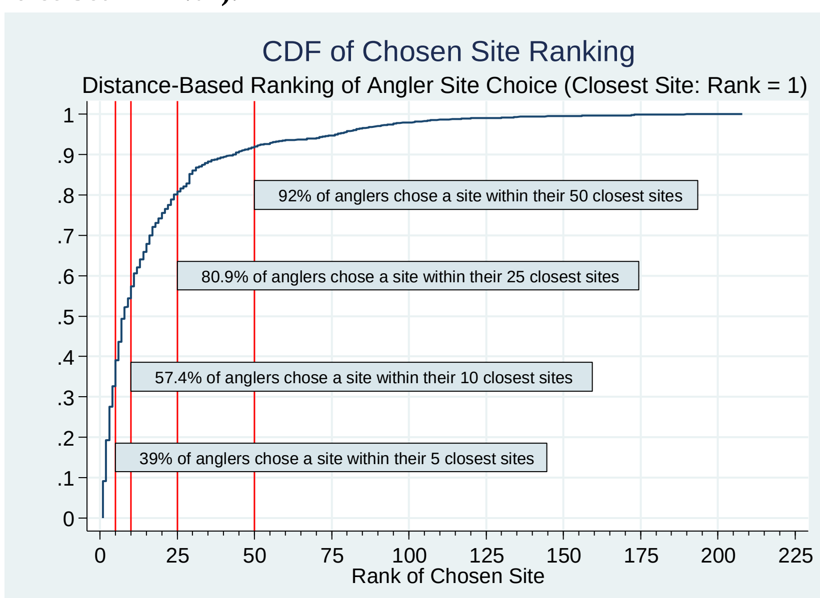
Source: Characteristics for Individual Fishing Sites Can be Found by Following https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp, and Choosing Appropriate State, County, and Locational Filters.

Figure 23. Histogram of Ranked Angler Site Choices.



Source: Data from NOAA's Marine Recreational Information Program.

Figure 24. CDF of Ranked Angler Site Choices. The Distance-Based Ranking Shows That the Majority of Day-Trip Anglers Choose Within their 20 Nearest Sites (Mean Distance Traveled to Chosen Site = 30.98 Miles; Mean Ranking of Chosen Site = 17.28; Mean Number of Sites in Choice Set = 144.07).



Source: Data from NOAA's Marine Recreational Information Program.

Table 13. Summary Statistics Over 2013-15 by Angler State of Residence.

	Alabama	Florida	Georgia	Louisiana	Mississippi	Other
#Weekday Trips	161	1018	24	4	280	0
#Weekend Trips	523	3977	109	4	890	6
Total #Trips	684	4995	133	8	1170	6
Choice Set Size (#Sites)	120.6	163.2	34.18	67.75	89.19	153
	(11.84)	(44.66)	(18.14)	(60.36)	(14.34)	(45.81)
One-Way Dist. Traveled (Miles)	33.41	31.90	105.2	65.78	16.99	21.98
	(30.42)	(29.75)	(26.22)	(28.08)	(23.97)	(23.34)
One-Way Time Traveled (Hours)	3.995	0.876	2.746	1.177	0.456	0.535
	(20.29)	(0.719)	(0.625)	(0.404)	(0.501)	(0.365)
#Fishermen in Party	2.225	2.362	2.579	3.375	2.286	2
	(1.064)	(0.934)	(0.889)	(1.598)	(0.940)	(0.632)
#Hours Fished	4.136	4.327	4.380	5.438	4.120	3.417
	(2.125)	(1.817)	(1.787)	(1.400)	(1.941)	(1.429)
#Days Fished Last 12 Months	62.66	46.64	43.89	8.875	56.80	77
	(160.2)	(108.6)	(169.9)	(9.523)	(94.36)	(52.84)
#Days Fished Last 2 Months	7.417	6.329	4.571	2.625	8.704	11.83
	(10.81)	(8.931)	(14.63)	(3.292)	(9.802)	(8.495)

For All but the First Three Statistics, Averages are Presented with Standard Deviations in Parentheses.
Source: Data from NOAA's Marine Recreational Information Program.

Table 14. Summary Statistics of Fishing Activity in States with Fishing Site Destinations in this Study.

	Alabama	Florida	Mississippi
#Fishing Sites	47	333	52
#Weekday Trips	158	1042	287
#Weekend Trips	480	4132	897
Total #Trips	638	5174	1184
Choice Set Size (#Sites)	121.2	151.3	90.47
	(8.022)	(51.74)	(13.54)
One-Way Dist. Traveled (Miles)	27.78	36.23	16.04
	(22.71)	(33.93)	(21.41)
One-Way Time Traveled (Hours)	4.229	1.009	0.438
	(21.91)	(0.842)	(0.452)
#Fishermen in Party	2.286	2.255	2.263
	(1.163)	(0.893)	(0.948)
#Hours Fished	4.515	4.613	4.577
	(2.093)	(1.669)	(1.960)
#Days Fished Last 12 Months	54.67	49.55	59.01
	(117.3)	(109.7)	(77.76)
#Days Fished Last 2 Months	8.571	6.696	9.427
	(10.98)	(8.820)	(9.167)
#Fish Available for Inspection	2.944	1.938	3.933
During Interview	(4.324)	(2.776)	(5.687)

For All but the First Four Statistics, Averages Per Trip Taken are Presented with Standard Deviations in Parentheses.
Source: Data from NOAA's Marine Recreational Information Program.

Table 15. Summary Statistics of Site Characteristics for All Sites within Each State.

	Alabama	Florida	Mississippi
#Fishing Sites	47	327	58
Exp. Catch - Spotted Seatrout	3.228	4.325	3.407
	(0.882)	(1.216)	(0.674)
Fishing Pressure (#Anglers per Day)	5.762	7.096	6.206
	(5.787)	(9.159)	(6.264)
#Charter Boats	2.426	1.835	2.385
	(6.125)	(3.241)	(7.113)
#Head Boats	0.106	0.0240	0.0962
	(0.598)	(0.189)	(0.358)
#Car Parking Spaces	61.11	40.48	56.38
	(114.7)	(48.28)	(78.06)
#Boat Ramps	1.702	1.180	1.577
	(1.667)	(1.524)	(1.775)
#Boat Trailer Spaces	19.96	22.72	23.10
	(22.63)	(32.73)	(24.93)
#Boat Slips	69.21	30.62	33.10
	(150.3)	(92.91)	(66.96)
Private or Restricted Public Access (0 or 1)	0.0638	0.249	0
Access Fee Charged (0 or 1)	0.298	0.393	0.212
Tide Affects Fishing (0 or 1)	0.191	0.387	0.308
Fuel Dock On Site (0 or 1)	0.319	0.246	0.212
Boat Storage On Site (0 or 1)	0.255	0.192	0.173
Boat Maintenance/Repair On Site (0 or 1)	0.149	0.144	0.0192
Fish Cleaning Stations On Site (0 or 1)	0.447	0.411	0.423
Night Lighting Available (0 or 1)	0.638	0.517	0.712
Is the Fishing Site Safe? (0 or 1)	0.745	0.360	0.673
Bait Shop On Site (0 or 1)	0.404	0.348	0.308
Tackle Shop On Site (0 or 1)	0.255	0.330	0.212
Tournaments Held On Site (0 or 1)	0.149	0.102	0.154
Restaurant Nearby (0 or 1)	0.319	0.315	0.250
Lodging Nearby (0 or 1)	0.0426	0.213	0.0962

For the first statistic and all non-binary statistics, averages per trip taken are presented with standard deviations in parentheses. Source: data from NOAA's Marine Recreational Information Program.

Table 16. Conditional Logit Model Results Using *Complete* Choice Sets Including All Sites within 150 Miles of An Angler's Home Zip Code are Considered Feasible. Column Headings Indicate Whether Specified Site Characteristics are Time Invariant or Time Varying.

	Benchmark		Time Invariant	Time Varying		
Outcome: fishing site choice	(1)	(2)	(3)	(4)	(5)	(6)
Benchmark Characteristics						
Distance (Miles)	-0.081***	-0.1089***	-0.0822***	-0.109***	-0.109***	-0.109***
	(0.00124)	(0.00205)	(0.00129)	(0.00205)	(0.00205)	(0.00205)
Exp. Catch - Spotted Seatrout	0.2870***	0.4166***	0.4201***	0.4186***	0.4153***	0.4083***
	(0.02671)	(0.07967)	(0.02973)	(0.07972)	(0.07973)	(0.07964)
Site Characteristics						
#Boat Trailer Spaces			0.0170***	-0.0011	-0.0013	-0.0013
			(0.00039)	(0.00097)	(0.00098)	(0.00099)
#Car Parking Spaces			0.0008***		0.0037**	0.0036**
			(0.00020)		(0.00150)	(0.00151)
#Boat Slips						-0.0011
						(0.00123)
Night Lighting (0 or 1)			-0.2394***			-0.2456
			(0.03391)			(0.17618)
#Boat Ramps			0.1759***			-0.0076
			(0.00716)			(0.03880)
Private Access (0 or 1)			-0.6071***			
			(0.07099)			
Access Fee (0 or 1)			-0.4716***			
			(0.03648)			
Tide Affects Fishing (0 or 1)			-0.0596*			
			(0.03045)			
Bait Shop (0 or 1)			0.2874***			
			(0.04473)			
Fish Cleaning Station (0 or 1)			-0.2642***			
			(0.03746)			
Fuel Dock (0 or 1)			-0.0320			
			(0.05097)			
Tournaments (0 or 1)			0.8624***			
			(0.03728)			
Restaurant (0 or 1)			-0.5048***			
			(0.03781)			
Lodging (0 or 1)			-0.9208***			
			(0.05331)			
#Charter Boats			0.0534***			
			(0.00472)			
#Head Boats			-1.0996***			
			(0.14962)			
Alternative Specific Constants	No	Yes	No	Yes	Yes	Yes
Number of Anglers	6,996	6,996	6,996	6,996	6,996	6,996
Observations	1,007,925	1,007,925	1,007,925	1,007,925	1,007,925	1,007,925
Total #Sites	432	432	432	432	432	432
Avg. #Site Choices	144.07	144.07	144.07	144.07	144.07	144.07
Min. #Site Choices	5	5	5	5	5	5
Max. #Site Choices	224	224	224	224	224	224
Log Likelihood	-25,491.80	-16,166.96	-21,027.71	-16166.40	-16,163.70	-16,161.96
AIC	50,987.60	33,199.91	42,089.42	33,200.80	33,197.39	33,199.93
BIC	51,011.25	38,319.45	42,290.42	38,332.15	38,340.57	38,378.58
Run Time (Seconds)	414.79	9,628.69	335.34	9,581.28	5,814.20	6,004.59

Robust standard errors in parentheses, adjusted for clustering on angler. * p<0.10, ** p<0.05, *** p<0.01

Table 17. Conditional Logit Model Results Using *Partially Aggregated* Choice Sets. All Sites with a Ranking Worse than 50 are Lumped Into a Single *Outside Option* for Each Angler. Column Headings Indicate Whether Specified Site Characteristics are Time Invariant or Time Varying.

	Benchmark		Time Invariant		Time Varying	
Outcome: fishing site choice	(1)	(2)	(3)	(4)	(5)	(6)
Benchmark Characteristics						
Distance (Miles)	-0.0691***	-0.0917***	-0.0715***	-0.0918***	-0.0917***	-0.0917***
	(0.00124)	(0.00194)	(0.00131)	(0.00194)	(0.00194)	(0.00194)
Exp. Catch - Spotted Seatrout	0.4030***	0.3691***	0.3220***	0.3711***	0.3654***	0.3606***
	(0.01812)	(0.07853)	(0.02035)	(0.07860)	(0.07871)	(0.07872)
Site Characteristics						
#Boat Trailer Spaces			0.0168***	-0.0011	-0.0015	-0.0015
			(0.00040)	(0.00116)	(0.00119)	(0.00119)
#Car Parking Spaces			0.0005***		0.0041**	0.0040**
			(0.00021)		(0.00163)	(0.00164)
#Boat Slips						-0.0008
						(0.00117)
Night Lighting (0 or 1)			-0.3063***			-0.2103
			(0.03449)			(0.18258)
#Boat Ramps			0.1824***			0.0073
			(0.00757)			(0.05726)
Private Access (0 or 1)			-0.8325***			
			(0.07543)			
Access Fee (0 or 1)			-0.5092***			
			(0.03835)			
Tide Affects Fishing (0 or 1)			-0.0977***			
			(0.03027)			
Bait Shop (0 or 1)			0.2941***			
			(0.04672)			
Fish Cleaning Station (0 or 1)			-0.2181***			
			(0.03739)			
Fuel Dock (0 or 1)			0.0996*			
			(0.05319)			
Tournaments (0 or 1)			0.8815***			
			(0.03772)			
Restaurant (0 or 1)			-0.5300***			
			(0.03952)			
Lodging (0 or 1)			-0.6701***			
			(0.06396)			
#Charter Boats			0.0289***			
			(0.00502)			
#Head Boats			-0.9114***			
			(0.15249)			
Alternative Specific Constants	No	Yes	No	Yes	Yes	Yes
Number of Anglers	6,996	6,996	6,996	6,996	6,996	6,996
Observations	353,379	353,379	353,379	353,379	353,379	353,379
Total #Sites	433	433	433	433	433	433
Avg. #Site Choices	50.51	50.51	50.51	50.51	50.51	50.51
Min. #Site Choices	5	5	5	5	5	5
Max. #Site Choices	51	51	51	51	51	51
Log Likelihood	-23,562.49	-14,791.01	-19,661.08	-14,790.48	-14,787.55	-14,786.53
AIC	47,128.99	30,450.02	39,352.16	30,450.95	30,447.10	30,451.06
BIC	47,150.54	35,126.49	39,513.79	35,138.21	35,145.13	35,181.42
Run Time (Seconds)	103.15	2,651.64	148.58	2,978.38	2,835.84	3,322.95

Robust standard errors in parentheses, adjusted for clustering on angler. * p<0.10, ** p<0.05, *** p<0.01

Table 18. Fishing Site Pressure Descriptions from the MRIP Public Access Fishing Site Register.

	Site Pressure Item	Description
1	STATE_CODE	Two-digit numeric codes for representing the 50 states, the District of Columbia, outlying areas of the U.S., the freely associated states, trust territory, and individual minor outlying island territories. See FIPS Pub 5-2 for more details.
2	STATE	Two letter USPS abbreviation
3	COUNTY_CODE	Codes representing the counties and other entities treated as the equivalents of counties for legal and/or statistical purposes in the 50 States, the District of Columbia, the possessions, and freely associated areas of the United States.
4	COUNTY	Name of the counties and other entities treated as the equivalents of counties for legal and/or statistical purposes in the 50 States, the District of Columbia, the possessions, and freely associated areas of the United States.
5	SITE_EXTERNAL_ID	Identifier used in MRIP Site Register.
6	SITE_NAME	Name of the fishing site.
7	MONTH	Month for which pressure is valid for a given fishing site. e. g. JUN, JAN etc.
8	KOD	Kind-of-day for which pressure is valid for a given fishing site. e.g. weekday, weekend.
9	INTERVAL	Time interval for which pressure is valid for a given fishing site. e. g. 0200-0800, 0800-1400 etc.
10	SITE_GROUP	Indicates the primary mode for the pressure as determined by an internal formula.
11	BEACH/BANK	Indicates that the pressure value applies to BEACH/BANK fishing mode.
12	MAN-MADE	Indicates that the pressure value applies to MAN-MADE fishing mode.
13	CHARTER BOAT	Indicates that the pressure value applies to CHARTER BOAT fishing mode.
14	PRIVATE/RENTAL BOAT	Indicates that the pressure value applies to PRIVATE/RENTAL BOAT fishing mode.
15	SHORE	Indicates that the pressure value applies to SHORE fishing mode.
16	OFFSHORE	Indicates that the pressure value applies to OFFSHORE fishing mode.

Source: https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp. From the Help tab, click “Export Guide.”

Table 19. Full List of Site Characteristics Available on the MRIP Public Access Fishing Site Register.

	Characteristic Name	Description
1	SITE_EXTERNAL_ID	State and county specific identifier for the Site; aka Site ID in the Site Register web site.
2	SITE_NEW_FLAG	A 'Y' indicates whether the fishing site is new.
3	SITE_NAME	Name of the fishing site.
4	SITE_ADDRESS	Street address of a fishing site.
5	SITE_CITY	Name of the city where fishing site is located.
6	SITE_ZIP	Five-digit postal code base.
7	SITE_ZIP_ADDON	Four-digit postal code add-on.
8	COUNTY_CODE	Codes representing the counties and other entities treated as the equivalents of counties for legal and/or statistical purposes in the 50 States, the District of Columbia, the possessions, and freely associated areas of the United States.
9	COUNTY	Name of the counties and other entities treated as the equivalents of counties for legal and/or statistical purposes in the 50 States, the District of Columbia, the possessions, and freely associated areas of the United States.
10	STATE_CODE	Two-digit numeric codes for representing the 50 states, the District of Columbia, outlying areas of the U.S., the freely associated states, trust territory, and individual minor outlying island territories. See FIPS Pub 5-2 for more details.
11	STATE	Two letter USPS abbreviation.
12	SITE_LAT	Latitude of the fishing site.
13	SITE_LONG	Longitude of the fishing site.
14	SITE_DIRECTION	Directions to the fishing site from a major highway.
15	SITE_NOTE (Only admin)	Notes on the fishing site.
16	SITE_CONTACT_NAME	Name of the contact person.
17	SITE_CONTACT_PHONE	Phone number of the contact person.
18	MODIFIED_BY (Only admin)	The last user who modified this record.
19	SITE_REVIEWER_NOTE (Only admin)	Reviewer's notes on the fishing site.
20	STATUS	Status of Site such as Submitted, Draft, etc.
21	SITE_TEMP_RETIRE_FLAG	A 'Y' indicates that a site is marked for temporary retirement.
22	# OF BOAT SLIPS	Total number of boat slips at site (included permanent, transient/guest slips for any boat type).
23	# OF CAR PARKING SPACES	Total number of spots for cars.
24	# OF CB USING SITE	Total number of charter boats using the site. Include posted Charter boat slips, other charter boats that regularly use the site, and guide boats that regularly use the site, regardless of whether or not the boats have posted/assigned slips.
25	# OF HB USING SITE	Total number of head boats using the site. Included are posted Head boat slips and other head boats that regularly use the site, regardless of whether or not the boats have posted/assigned slips.
26	# OF RAMPS	Total number of Ramps or number trailered boats that can be simultaneous launched from the site.

Table 19 Continued.

	Characteristic Name	Description
27	# OF TRAILER PARKING SPACES	Total number of spots for trailers.
28	BOAT MAINTENANCE/REPAIR	Boatyard facility onsite for general boat or engine repairs or maintenance work.
29	BOAT STORAGE	Boat storage onsite. Land boat storage only (shed, barn, botel, open lot, etc.).
30	CAN WE INTERVIEW	Answer NO if we have been refused ongoing access to interview anglers; answer YES if the refusal was only on specific occasion(s).
31	FISH CLEANING STATIONS	Designated fish cleaning table/stand with or without running water (either DIY or pro).
32	FISHING ACTIVITY AFFECTED BY TIDE	Is recreational fishing activity at this site (any mode present) affected by the tide level, tidal flow, or seasonal tide conditions?
33	FUEL DOCK	Fuel Dock onsite.
34	HEADBOAT ONLY	Is this site a head boat docking or pick-up site only, i.e., no other recreational fishing modes present?
35	IS FEE CHARGED TO THE PUBLIC FOR USE OF SITE	Specific to recreational fishing access point; does not include general park entrance fee or general parking fee. Parking fee may apply if site is solely a fishing access point and not for other recreational use.
36	IS SITE SAFE FOR 2 SAMPLERS AT NIGHT	Is the site considered safe enough to send two samplers to conduct night sampling Well-lit, not isolated, not known crime hotspot, no gang activity?
37	LIGHTING AT NIGHT	Fishing site is well lit at night and throughout the night.
38	LODGING	Visitor lodging establishment (hotel, motel, guesthouse, bed & breakfast, etc.) on the site or in on an adjacent property (onsite/immediate vicinity).
39	MAJOR TOURNAMENTS	Based from this site at any time during the year (not wave specific). A major fishing tournament is defined as an organized competitive fishing event and involves the award of trophies, prizes or other recognition. Major tournaments are generally publicized in advance and typically require that participants pre-register to participate.
40	PRIVATE ACCESS	Restricted or no public access, regardless of whether or not we are allowed to interview.
41	RESTAURANT	Restaurant onsite or on an adjacent property (onsite/immediate vicinity).
42	RETAIL BAIT	Bait for sale to the public (not commercial fishing) onsite; may include live, fresh, frozen, preserved, or vending machine bait sales.
43	SHORE AREA	What type of water body is accessed when fishing from shore at this site (N/A if no shore fishing present at this site)?
44	SHORE MODE	What type of shore access is available at the site, man-made or beach/bank (i.e. natural) (N/A if no shore fishing present at the site)?
45	TACKLE SHOPS	Shop, store, or other building onsite that sells fishing tackle with or without bait.

Source: https://www.st.nmfs.noaa.gov/msd/html/siteRegister.jsp. From the Help Tab, Click “Export Guide.”

APPENDIX B

SUPPLEMENTARY MATERIAL FROM SECTION 2

B.A Additional Background Information

B.A.1 More on the Rule of Capture

The Texas Supreme Court ultimately chose the rule of capture based on two public policy considerations. First: “Because the existence, origin, movement and course of such waters, and the causes which govern and direct their movements, are so secret, occult and concealed that an attempt to administer any set of legal rules in respect to them would be involved in hopeless uncertainty, and would therefore be practically impossible.” Second: “Because any such recognition of correlative rights would interfere, to the material detriment of the commonwealth, with drainage of agriculture, mining, the construction of highways and railroads, with sanitary regulations, building, and the general progress of improvement in works of embellishment and utility” (Potter 2004). However, for more than a century the Texas Supreme Court had not made an official decision on whether a landowner owns not only the water that emerges from the ground, but the water in place underground as well (i.e. ownership before the water is produced). Finally, on February 24, 2012 in *Edwards Aquifer Authority v. Day*, the Supreme Court announced for the first time that under Texas law the ownership of the groundwater in place also belongs to the owner of the property and is subject to takings (when property owners require compensation for having their withdrawals capped or reduced). This is similar to mineral rights associated with oil and gas resources, yet it is still unclear what is considered effective groundwater management and regulatory takings (McCarthy and Jackson, Sjoberg, McCarthy and Townsend LLP 2012 and Texas Water Code Section 36.002).

B.A.2 Priority Groundwater Management Areas

SB1 moved to treat the state as a whole by setting up regional planning groups and providing data collection to close data gaps. Priority Groundwater Management Areas (PGMAs) are identified by the TCEQ with assistance from the TWDB as areas that currently have no GCD and will potentially have "critical problems" within the next fifty years.[177] They were created to enable effective management of groundwater resources in areas of the state where critical groundwater problems exist or may exist in the future.[177] As of January 2017, seven PGMAs have been designated in Texas and cover all or part of 35 counties (TCEQ and TWDB 2017).[178] Once a decision to designate an area as a PGMA has been made, the affected counties must take one of several actions within two years: (1) join an existing GCD, (2) create one or more GCDs, (3) or a combination of (1) and (2) depending on the hydrogeology. If affected counties do not take steps in creating a GCD, the TCEQ will step in and create one or more districts under Chapter 36 of the Texas Water Code.

B.A.3 Water Life Cycle in Hydraulic Fracturing

Freshwater consumption is water that, following its use, is removed from the local hydrologic cycle and is therefore unavailable to other potential users (U.S. EPA 2016b).

---

[177] Source: https://www.tceq.texas.gov/groundwater/pgma.html.

[178] A map of Texas PGMAs is available at http://hayscountyroundup.blogspot.com/2009/11/tceq-report-looks-at-options-to-plug.html, and an (outdated) shape file is available at http://www.twdb.texas.gov/mapping/gisdata.asp.

Hydraulic fracturing operations can consume water in a variety of ways, such as through evaporation from storage ponds (used to store water near the well pad before stimulation occurs), retention of water in the geologic formation, or disposal of wastewater in Underground Injection Control (UIC) Class II injection wells (U.S. EPA 2016b). Although the successful stimulation of wells has become more resilient to the use of various water types,[179] historically, the majority of hydraulic fracturing operations have used freshwater because it requires minimal testing and treatment (U.S. EPA 2016b), and therefore is usually the least cost water option. The U.S. EPA (2016b) outlines five stages in the hydraulic fracturing water cycle, where each stage is defined by an activity involving water that supports hydraulic fracturing (Table 20).

Table 20. The Stages and Activities in the Hydraulic Fracturing Water Cycle.

Stage	**Activity**
Water Acquisition	The withdrawal of groundwater or surface water to make hydraulic fracturing fluids.
Chemical Mixing	The mixing of a base fluid, sand or proppant, and additives at the well site to create hydraulic fracturing fluids.
Well Injection	The injection and movement of hydraulic fracturing fluids through the oil and gas production well and in the targeted rock formation.
Wastewater Handling	The on-site collection and handling of water that returns to the surface after hydraulic fracturing stimulation and the transportation of that water for disposal or reuse.
Wastewater Disposal and Reuse	The disposal and reuse of hydraulic fracturing wastewater.

Source: U.S. EPA (2016a).

B.A.4 Water Sourcing and Disposal

The amount of water needed for a hydraulic fracturing stimulation depends on the region and many other factors (see Appendix B.A.5), and is needed within a short period of time to ensure sufficient pressure can be applied to stimulate the well and meet

---

[179] Mentioned in a phone conversation with Gabriel Collins, an attorney in Houston, Texas. https://www.bakerinstitute.org/experts/gabe-collins/.

production expectations, as designed by the completion engineer. Although the Texas Railroad Commission is the primary authority regulating the oil and gas industry, it has no statutory authority to regulate water use in the industry,[180] and operators can currently use any amount of water in development activity. Since water is over allocated in Texas, operators obtain water by purchasing water from owners of water rights, or land or water rights themselves. Due to industry water needs, lucrative markets for water have developed in regions with hydraulic fracturing activity. In fact, since the revenue from selling water is so large, some landowners will not sign an oil and gas lease unless the terms specify that the operator must purchase its water from a supply well located on their property (Goldenberg 2013; Hiller 2018). Further, most ranchers would rather have an operator drill a new freshwater well on their land because after the oil and gas well is completed, they have a useful new freshwater well.

In areas with relatively low water availability, large water withdrawals occurring over a short period can lead to a decline in water availability. Anecdotal evidence of such impacts has been in the form of drying domestic wells; cattle wells running dry on the Fasken Oil and Ranch, located in Midland, Texas (Dallas News 2014); and stream capture decline, which has caused private stock dams to run dry in western North Dakota (Kusnetz 2012). Aside from the direct impacts on groundwater availability, there is a fear of displacement of local homeowners as water scarcity increases. Kusnetz (2012) documents this fear amongst residents in western North Dakota, where several cases of fruitless

[180] Source: http://www.rrc.state.tx.us/about-us/resource-center/faqs/oil-gas-faqs/faq-water-use-in-association-with-oil-and-gas-activities/.

drilling of new water wells occurred, as well as reports of residents that had to haul water for domestic and livestock use from out of town.

Further, in Barnhart, Texas, a small town in the western part of the state located on the eastern fringe of the Permian Basin, Goldenberg (2013) reported that the "town well ran dry." However, precluding this were "warning signs," where residents reported seeing sand in toilet bowls, sputters of air in faucets, and water pumps that worked overtime but produced no water. While much of the town's water supply was being used for oil and gas development (one rancher reported 104 water supply wells were drilled on his leased land), residents complained of water rationing restrictions. The article also reported that many local ranchers sold off much of their herds, and cotton farmers lost significant yields, as it became increasingly difficult (or prohibitively expensive—new wells can cost tens of thousands of dollars) to provide feed and water under conditions of drought combined with the new water demands of the industry.[181]

The collective depletion of aquifers therefore can necessitate water users to invest in new (and larger) water wells or pumps, which effectively must be drilled deeper in order to access the available groundwater. Since these wells are also pumping water from deeper depths, the costs associated with pumping water became more expensive.[182] In this case, large spatial externalities exist since the pumping of water by one or many users affects other nearby water users, and these externalities are important causes of welfare losses

---

[181] This is an example of the "stock" externality (Provencher and Burt 1993), where water used today is not available tomorrow, and pumping with existing provides no water.
[182] This is an example of the "pumping cost" externality (Provencher and Burt 1993).

(Pfeiffer and Lin 2012). Careful management of water can therefore be needed at the local level, such as in regions with significant hydraulic fracturing activity, as it is local water availability that is the most sensitive in terms of social welfare.

Wastewater– both the initial flowback, and the produced water –is pumped throughout the life of a well along with oil and gas, and contains many of the salts, minerals, and other petroleum residues that naturally exist in the formation. It poses an expensive logistical challenge for operators. For example, in the Permian Basin, 252 to 336 gallons (6 to 8 barrels) of water are produced per barrel (42 gallons) of oil (Carr 2017), and the volumes of both produced water and oil decline at relatively the same rate as the well ages (Kondash and Vengosh 2015). Operators must dispose, treat, or reuse this wastewater in a safe and responsible manner. In Texas, operators have typically opted to dispose wastewater via injection into UIC Class II injection wells since it is less expensive to dispose of the wastewater and purchase new freshwater, as opposed to treating and reusing recycling wastewater (Texas Railroad Commission Undated and Collins 2017). Disposal of produced water in injection wells however, has been connected to seismic activity in certain areas, particularly in north Texas and Oklahoma (Ellsworth 2013; Walsh and Zoback 2015). If total wastewater volumes continue to increase, as is projected to happen due to increasing hydraulic fracturing activity and longer horizontal wellbores that use more water, the wastewater disposal problem may become even more pronounced due to limited disposal well capacities. In attempt to circumvent both the water supply and disposal issues, several companies have recently started developing cheaper ways to recycle wastewater.[183] However, efforts to recycle can be limited if landowner agreements require operators to utilize their water resources, a common obstacle to recycling in Texas (Goldenberg 2013; Hiller 2018).

B.A.5 Factors Affecting Total Water Use in a Well

There are many factors and decisions that operators make that affect the volume of water needed to stimulate a well, including the measured depth, or the total length of pipe used in a well, which is a function of the true vertical depth and the lateral (or horizontal) length. The measured depth of a well relates to total water use, as some shale formations lie deeper than others and therefore require more water to fill the vertical space in the wellbore. Similarly, the horizontal length component of measured depth has a direct correlation to the amount of water used as well. As the "lateral" increases, more water is needed to fill space in the horizontal portion of the wellbore, maintain pressure, and carry water and sand into more fractures.

After the well is drilled and the casing and cement are in place (sealing the wellbore from the hydrocarbon-producing formation), perforating guns containing explosive charges are pulled into the wellbore and detonate to "punch" perforations (or holes) through the casing and cement and into the formation.[184] The holes reconnect the formation to the wellbore and provide a path for the fracturing fluid to be forced into the formation. The choice of perforation design has an immediate influence on the well's total

[183] Apache Corporation has built five water-recycling facilities in Balmorhea (West Texas) that can store around 126 million gallons of wastewater (Hiller 2018).

[184] An example of a perforating gun is available at http://www.halliburton.com/en-US/ps/wireline-perforating/wireline-and-perforating/perforating-services/high-pressure-gun-systems/default.page.

water use and productivity as it dictates the number of holes punched into the target formation (more holes mean more water is needed). Accordingly, the charges used in perforation treatments can be of different sizes, although they are constrained by the amount of space in the wellbore. Some are designed to create longer or shorter perforation lengths extending into the formation (ranging from 6” to 48”); and others are designed to create perforation holes of various sizes (ranging from .23” to .72” by diameter).[185,186,187]

Learning in the industry has also occurred on other dimensions that contribute to the amount of water used per well. As exploration and drilling have generated significant amounts of new information on local geology, operators have become more adept at choosing optimal completion inputs and stimulation techniques to free more of the oil and gas trapped within, such as by drilling wells with longer horizontal lengths and stimulating them in multiple discrete stages (or intervals).[188] The latter has significantly increased production per well as it creates a larger fracture network that reaches more of the

---

[185] More in-depth discussions and visuals of the perforation process are available at https://www.epa.gov/sites/production/files/documents/casingperforatedoverview.pdf and https://info.drillinginfo.com/well-completion-101-part-2-well-perforation/.

[186] It is true that significant heterogeneity exists in firm beliefs over optimal perforation lengths and hole-size, but after conversations with Bob Kleinberg, a former employee at Schlumberger (https://www.bu.edu/ise/profile/robert-kleinberg/), it appears that the geology of the formation usually dictates these choices.

[187] DrillingInfo also describes how too large of a perforation hole or too long of a perforation length can lead to excess debris from the explosions, which can cause blockage in the wellbore and therefore reduce well productivity. Similarly, too small of a perforation hole diameter and or too short of a perforation length can affect well productivity since less “damage” to the formation is created. To circumvent these concerns, various perforation patterns have been used in attempt to maximize wellbore exposure to the producing formation without jeopardizing well productivity, and there are usually 4 to 8 holes perforated per foot, where the most common patterns create holes in 3, 4, or 6 directions across a given perforated interval or stage.
Source: https://info.drillinginfo.com/well-completion-101-part-2-well-perforation/.

[188] A more in-depth discussion and a visual for the concept of multi-stage stimulation is available at https://www.researchgate.net/publication/270340648_Integrated_Shale_Gas_Reservoir_Modeling/figures?lo=1.

producing formation around each interval, and enabling the entire length of the wellbore to be stimulated more thoroughly. Over time, operators have also increased the number of stages used to stimulate a wellbore, which directly affects total water use.[189] With more stages, and possibly a larger number of perforations in each stage, more water is needed to stimulate and enter a larger number of fractures.

In addition to its contribution to well productivity, the stimulation of multiple stages can take place iteratively or all at once before production, potentially offering operators more freedom to pace extraction with other decisions or market conditions (Vissing 2018). Similarly, refracturing a well one or more times over its life is another important determinant of total water use. Refracturing has become more common in the industry as it provides a relatively low-cost means of maintaining total output by increasing production rates from older wells as opposed to drilling new ones.[190] It can also be an effective way to revive production from wells where the initial stimulation was poor and did not offer good returns.[191] The caveat is that each refracture requires an additional volume of water, but usually less than the initial stimulation.

Lastly, the composition of the hydraulic fracturing fluid used and geology of the formation are other important factors affecting total water use. The hydraulic fracturing fluid composition is an especially important determinant as the industry has begun to use

---

[189] Mentioned in conversations with Bob Kleinberg, formerly of Schlumberger.
[190] Source: https://info.drillinginfo.com/makes-successful-refrac/.
[191] Conversations with DrillingInfo indicate that, for a variety of reasons, operators are not usually able to stimulate fractures in all perforations along a horizontal wellbore, meaning that reserves are commonly left behind during the initial stimulation. This feature has created the refracture market, and firms commonly use downhole technology such as a flowmeter to detect intervals along the wellbore that were not fractured, and provide potential refracture targets.

more sand per lateral foot, which requires additional water to carry sand particles deeper into the fractures. The proportions of each water type (e.g. freshwater, brackish or saltwater, and recycled wastewater[192]) in a fracture fluid can affect water needs due to their respective densities (denser fluids take up more space), and the use of alternative water types reduces freshwater needs. Other fracture fluids can contain non-aqueous substances such as liquid-gas mixtures of nitrogen or carbon dioxide, both of which reduce the amount of water needed to stimulate a well. Geologic characteristics, such as shales, tight sands, and coalbeds, affect the perforation choices mentioned above, but they also influence the amount of water used per well as some formations are harder and require a more water pressure to stimulate fractures than softer formations. Some formations also have more cracks and associated natural leakage, and are therefore more conducive to other unintended losses, which increases water use (U.S. EPA 2015).

B.A.6 Impacts of Water Use

In a county-level analysis of water use (or consumption) and availability, U.S. EPA (2016b) found that large volumes of water used in hydraulic fracturing alone do not necessarily result in impacts to drinking water resources. Where water availability is low, compared to use, withdrawals for hydraulic fracturing are more likely to affect drinking water resources or require curtailments. For example, in Pennsylvania, a water rich state, water withdrawals have been restricted during summer and drought conditions in the

[192] I use the colloquial term "wastewater" to refer to both flowback and produced water that may be reused in hydraulic fracturing; I do not distinguish between flowback and produced water except when specifically reported in the literature.

Susquehanna River Basin (SRBC 2015). Furthermore, groundwater withdrawals exceeding natural recharge rates may lower the water level in aquifers (particularly for unconfined aquifers, i.e. those with no connection to surface recharge), potentially mobilizing contaminants or increasing their concentration. These results suggest that the potential for impacts exists, and that more studies are needed to understand where impacts will occur at the local scale.

B.B Background on Hydraulic Fracturing Regions in Texas

B.B.1 Major Unconventional Oil and Gas Formations and Water Sources

There are five major unconventional oil and gas formations in Texas.[193] Each is located in an area with different geological characteristics, implying that water use for an average well is likely to be different across space (Nicot et al. 2012). Similarly, these formations are located in areas with different levels of water availability, implying that water supplies for hydraulic fracturing operations come from a variety sources. Nicot et al. (2012) provide an outline of these and estimate the majority are surface and groundwater (including fresh and brackish or salt water). Other sources include recycled or reused wastewater from previous completions or from other industries or municipalities. In some areas, they also report that some operators have experimented with gel-based fracturing fluids, which reduce water needs.

[193] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/.

Barnett Shale

The Barnett Shale is one of the largest onshore natural gas fields in the U.S., and is of great economic significance to Texas as it is home to the original shale gas boom, where operators began using hydraulic fracturing techniques to enable production form unconventional sources (Nicot et al. 2012). The productive part of the formation is estimated to cover 5,000 square miles and at least 18 counties in the Dallas area, and contributes significantly to total U.S. natural gas production.[194] Water supply for hydraulic fracturing in this region comes from both surface and ground sources, and is estimated to be in equal 50% and 50% proportions (Nicot et al. 2014).

Eagle Ford Shale

The Eagle Ford Shale is of significant importance due to its capability of producing both natural gas and more oil than other traditional shale plays.[195] It sits below all or part of 27 counties and trends across Texas from the Mexican border into East Texas. It is roughly 50 miles wide and 400 miles long with an average thickness of 250 feet, and lies at a depth of between 4,000 and 12,000 feet. The play has been important to Texas, as it was where the majority of unconventional oil production first occurred in the state over 2010-2011, following the unconventional oil boom in North Dakota in 2008. Water supply for hydraulic fracturing in this region predominantly comes from groundwater sources (90%), although a small portion (10%) comes from surface sources (Nicot et al. 2012).

[194] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/barnett-shale-information/.
[195] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/eagle-ford-shale-information/.

Granite Wash

The Granite Wash is a tight sand play within the Anadarko Basin, encompassing a number of oil and gas producing formations, and lying below 26 counties in the panhandle of Texas and others in western Oklahoma.[196] Its spatial extent is approximately 160 miles long and 30 miles wide, and varies in depth from 11,000 to 15,400 feet and is 3000 feet thick on average. It is significant for both oil and gas production and, although the formation is predominantly composed of sand, has been a beneficiary of horizontal drilling methods developed for shale plays. Water supply for hydraulic fracturing in this region comes from groundwater sources (80%), and a small portion (20%) from surface sources (Nicot et al. 2012).

Haynesville Shale

The Haynesville shale is a gas-producing formation, which lies below 10 counties in East Texas and others in Western Louisiana.[197] The productive interval of the shale lies greater than 10,000 feet below the land surface. Nicot et al. (2012) jointly estimate portions of water supply for hydraulic fracturing in the Haynesville shale, and what they refer to as the East Texas Basin, which includes other smaller plays in the area. They estimate that 70% comes from groundwater sources and 30% from surface sources in this region.

[196] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/granite-wash-information/.
[197] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/haynesvillebossier-shale-information/.

Permian Basin

The Permian Basin in west Texas has become the one of the largest and most important hydrocarbon-producing regions in the U.S. It covers an area approximately 250 miles wide and 300 miles long and is composed of more than 7,000 fields.[198] Much of the area sits over cake-layered formations, where large amounts of oil and natural gas are produced from depths ranging from a few hundred feet to five miles below the surface. Importantly, it is located in a primarily semi-arid to arid environment (making it prone to drought), and sits under the Ogallala Aquifer in the northern part of the basin and under the Edwards-Trinity Aquifer in the southern part of the basin. Nicot et al. (2012) estimate that 100% of water use in hydraulic fracturing comes from groundwater sources in this region, and Cook and Webber (2016) note that landowners selling freshwater to operators typically pump it from the two aforementioned aquifers.

B.B.2 Understanding the Potential for Water Scarcity

Although part of this paper concerns the question of whether water use in hydraulic fracturing is large enough to affect local water availability, it is important to understand where the largest effects are most likely to occur. The regions in Texas with the most hydraulic fracturing activity are the Permian and Eagle Ford Basins, which are located in areas with low and relatively low rainfall and groundwater recharge. Given their water-scarce nature, it is reasonable to assume that water withdrawals are more likely to have a larger effect local on water availability in these areas, especially during times of drought.

---

[198] Source: http://www.rrc.state.tx.us/oil-gas/major-oil-and-gas-formations/permian-basin-information/.

In the Permian Basin, water use in roughly ten counties with significant hydraulic fracturing activity is not currently managed by a GCD; but Midland, Reagan, and Upton counties are identified as PGMAs. Similarly, in the Eagle Ford Basin, water use in roughly eight counties is not managed by a GCD. As shown in Nicot et al. (2012), water use in hydraulic fracturing in these areas increased between 2008 and 2012. Since then, many new wells have been completed and water use per well has continued to rise, in GCD and non-GCD areas.

Although the topic has been studied previously, and no study has shown a causal link between water of the industry and local availability,[199,200,201] water availability can be important in order for operators to keep completion costs low, particularly during times of low oil prices when profit margins are smaller. For example, during times of low water availability, operators may have to obtain water from more distant areas where it is plentiful, which can increase water-related expenditures in several ways. First, water prices should theoretically rise with scarcity, along with increasing competition over limited water supplies. Second, since operator demand for water may be inelastic (at least after a well is drilled),[202] water might be sourced from greater-than-average distances. In

---

[199] Scanlon et al. (2014) study whether water scarcity in will affect hydraulic fracturing activity in the Eagle Ford Basin. They find that with appropriate management, such as by increasing the use of brackish groundwater and produced water, water availability should not physically limit future shale energy production.

[200] Stevens and Torell (2018) posit that water availability may affect the amount of water used in hydraulic fracturing stimulations, among other drilling decisions and development outcomes. During times of drought, they show evidence that smaller wells are completed, which could collectively be due to a number of reasons.

[201] Freyman (2014), as part of a Ceres report, analyzes increasing water demand in hydraulic fracturing in water-stressed regions, and provides recommendations to stakeholders and operators for mitigating exposure to water sourcing risks.

[202] A completion engineer designs wells, but I am not sure how much cohesion there is between completion engineers and water managers (those responsible for ensuring sufficient water supplies are gathered for

these cases, water expenditures can be higher since transportation distance increases (transportation is the largest water-related cost) or if are export fees are associated with the area the water is obtained.[203]

B.B.3 Sourcewater

Due to plausible future water constraints such as having to haul water greater distances and therefore facing higher water costs, *sourcewater.com* was developed to be the largest online water source, reuse, and disposal database in the upstream energy industry. The platform aims to complement (or even replace) traditional sourcing methods used in the industry, which often involve operators sourcing water from the same water supplier or 'friend' in a particular area, or cold calling, and not finding the nearest water source. Not obtaining water from the nearest source means operators are not minimizing total water costs, which makes it more difficult for freshwater alternatives to be competitive when purchasing water. Hence, by using its online platform with over 100,000 water sources, Sourcewater advertises that operators can significantly cut down on total water costs by obtaining water from nearer sources.[204]

---

completions). Hence, after a horizontal well is drilled, I have reason to believe that few onsite decisions can be made to change the amount of water specified by the completion engineer. However, this warrants further study.

[203] One company was fined for using water from a source that was banned for us in hydraulic fracturing. Source: https://fuelfix.com/blog/2011/10/06/parched-texans-impose-water-use-limits-for-fracking-gas-wells/.

[204] Although this seems like an opportunistic way to obtain water prices, phone conversations with Benjamin Reed, COO, and Josh Adler, CEO, of Sourcewater, have indicated a reluctance of operators to transact on their platform, who instead opt instead for phoning the suppliers found in Sourcewater search results.

B.C Data

B.C.1 Additional Oil and Gas Well Statistics

An interesting trend in Table 21 is that the number of wells in the Permian basin in non-GCD areas is significantly greater than the number stimulated in GCD areas until 2015. This trend also holds across both drilling orientations. Although this could be due to better geology or other factors making non-GCD areas preferred in this sample period, it also potentially suggests that an easier access to water, and therefore plausibly lower water costs, could play a role in operator's decision on where to drill.

Table 21. Additional Descriptive Statistics for Oil and Gas Wells. Number of Wells in GCD and Non-GCD Areas in the Permian Basin (27,978 Observations).

		2012	**2013**	**2014**	**2015**	**2016**	**Through May 2017**
Number of Horizontal and Directionally Drilled Wells	GCD Areas	358	678	1,067	1,027	891	380
	Non-GCD Areas	650	803	1,297	1,309	992	317
Number of Vertically Drilled Wells	GCD Areas	2,567	2,495	2,019	596	234	82
	Non-GCD Areas	3,101	3,144	2,611	859	342	159

Source: Data from Primary Vision.

Table 22. Additional Summary Statistics for Oil and Gas Wells. Number of Wells Completed in Texas over February 2012 – May 2017.

	Unique Operators	# Wells	Mean # Wells Per Operator	HFFM?	Any Info. on Water Type?	Hor/Dir Wells	Mean Water Use Per Well
Operators in GCD and Non-GCD Areas	255	47,521	186	95.29%	78.84%	58.65%	4,130,954
Operators in GCD Areas Only	253	4,348	17	97.31%	73.71%	66.7%	4,719,932
Operators in Non-GCD Areas Only	148	1,313	9	99.92%	90.48%	19.04%	1,089,263

Source: Data from Primary Vision.

B.C.2 Additional Data

There are several other variables I would like to consider in future research, including an indicator for whether a well was completed near the expiration date of the primary term on its associated lease. The primary term specifies the maximum number of years within which an operator must drill and produce from at least one well, otherwise it will lose the lease. Herrnstadt et al. (2019) show that these expiration dates have a significant impact on drilling decisions, and a large share of wells are completed just prior to expiration, and controlling for this characteristic would be important if reporting for these wells is systematically worse. Additionally, given variation in the time taken to submit a completion report to FracFocus, this variable could be important if it is correlated with reporting less detailed information.

B.D Robustness Checks

B.D.1 Analysis of Reporting

Table 23. Logit Model Results for Reporting Metric 1. Outcome: Hydraulic Fracturing Fluid Mass Calculated (0 or 1)? Each Model Was Estimated Using Wells in Texas Over February 2012 Through May 2017, but O*mitting* Observations for Operators Who Did Not Have Wells in Both GCD and Non-GCD areas.

	(1)	(2)	(3)	(4)
GCD (0 or 1)	-0.6521***	-0.4201**	-0.3920**	-0.3883**
	(0.23310)	(0.17543)	(0.19775)	(0.19773)
Well Orientation (0 or 1)		-1.1573***	-1.0736***	-1.0626***
		(0.18880)	(0.19280)	(0.18665)
TWV (One Unit = 100k Gallons)		- 0.00547***	-0.00551***	-0.00544***
		(0.001092)	(0.001175)	(0.001153)
Refrac (0 or 1)		0.3354	0.3951	0.3885
		(0.32464)	(0.32987)	(0.32606)
#Wells by Operator in County-Month			-0.0078	-0.0064
			(0.01228)	(0.01248)
Operator's 1st Well			-0.2534	-0.2989
			(0.70965)	(0.70212)
Operator's 1st Well in County			0.2981	0.2985
			(0.20153)	(0.19807)
Cumulative #Wells in County	No	No	Yes	Yes
Cumulative #Wells by Largest 5 in County	No	No	Yes	Yes
Cumulative #Wells by Operator in County	No	No	Yes	Yes
#Wells by Largest 5 in County-Month	No	No	Yes	Yes
Total #Wells by Operator in County	No	No	Yes	Yes
Drought Controls	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes
N	37455	37455	37455	37455
Pseudo-R-Squared	0.374	0.395	0.398	0.400
Operator FEs	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on county. * p<0.10, ** p<0.05, *** p<0.01

Table 24. Logit Model Results for Reporting Metric 2. Outcome: *Any* Information on Water Type in Completion Report (0 or 1)? Each Model Was Estimated Using Wells in Texas Over February 2012 Through May 2017, but *Omitting* Observations for Operators Who Did Not Have Wells in Both GCD and Non-GCD Areas.

	(1)	(2)	(3)	(4)
GCD (0 or 1)	-0.3629***	-0.3761***	-0.4244***	-0.3986**
	(0.12950)	(0.13561)	(0.15918)	(0.15815)
Well Orientation (0 or 1)	-0.5892***	-0.7775***	-0.8116***	-0.8173***
	(0.15272)	(0.18209)	(0.14608)	(0.14439)
TWV (One Unit = 100k Gallons)		0.00554***	0.00511***	0.00513***
		(0.001621)	(0.001820)	(0.001809)
Refrac (0 or 1)		0.0223	0.0038	0.0047
		(0.16734)	(0.16721)	(0.16663)
#Wells by Operator in County-Month			0.0472	0.0472
			(0.03085)	(0.03082)
Operator's 1st Well			-0.5514**	-0.5530**
			(0.24142)	(0.24266)
Operator's 1st Well in County			0.2158*	0.2162*
			(0.12421)	(0.12493)
Cumulative #Wells in County	No	No	Yes	Yes
Cumulative #Wells by Largest 5 in County	No	No	Yes	Yes
Cumulative #Wells by Operator in County	No	No	Yes	Yes
#Wells by Largest 5 in County-Month	No	No	Yes	Yes
Total #Wells by Operator in County	No	No	Yes	Yes
Drought Controls	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes
N	45502	45502	45502	45502
Pseudo-R-Squared	0.367	0.369	0.380	0.380
Operator FEs	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on county. * p<0.10, ** p<0.05, *** p<0.01

B.D.2 Analysis of Hydraulic Fracturing and Groundwater Levels

B.D.2.A Various Specifications of Equation (2.2)

Table 25. Fixed Effects Model Results for Water Use in Hydraulic Fracturing Across Space. Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)
TWV10 (Unit = 100BBLs)	0.00027	0.00027	0.00100***
	(0.000219)	(0.000220)	(0.000138)
TWV (10-15 Miles)	0.00010	0.00009	0.00064***
	(0.000127)	(0.000129)	(0.000082)
TWV (15-20 Miles)	0.00001	0.00001	0.00007
	(0.000056)	(0.000053)	(0.000107)
TWV (20-25 Miles)	0.00007	0.00007	0.00024
	(0.000052)	(0.000053)	(0.000142)
TWV (25-30 Miles)	0.00003	0.00003	0.00015
	(0.000034)	(0.000033)	(0.000180)
TWV (30-35 Miles)	0.00011*	0.00010	0.00029
	(0.000060)	(0.000060)	(0.000205)
TWV (35-40 Miles)	0.00000	-0.00000	0.00007
	(0.000024)	(0.000024)	(0.000050)
TWV (40-45 Miles)	0.00007*	0.00007*	0.00006
	(0.000040)	(0.000040)	(0.000055)
TWV (45-50 Miles)	0.00008	0.00008	0.00016
	(0.000050)	(0.000049)	(0.000100)
Drought Index	Yes	Yes	Yes
Rain	Yes	Yes	Yes
Temp	No	No	Yes
Wind	No	No	Yes
Population	No	Yes	Yes
Total Corn Acres	No	Yes	Yes
Irrigated Corn Acres	No	Yes	Yes
Total Cotton Acres	No	Yes	Yes
Cotton Acres Irrigated	No	Yes	Yes
Total Sorghum Acres	No	Yes	Yes
Sorghum Acres Irrigated	No	Yes	Yes
Total Wheat Acres	No	Yes	Yes
Wheat Acres Irrigated	No	Yes	Yes
Total Rice Acres	No	Yes	Yes
Constant	Yes	Yes	Yes
N	15014	15014	5805
# Stations	267	267	106
R-Squared	0.135	0.144	0.224
Station FEs	Yes	Yes	Yes
Year-Month FEs	Yes	Yes	Yes

Standard errors in parentheses. Clustered on monitoring station. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

B.D.2.B Dynamic Treatment Effects

Table 26. Fixed Effects Model Results with Lags and Leads. Lags indicated by (-) and Leads Indicated by (+). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)
TWV10 (-6)	0.000189	0.000189*	0.000184	0.000179
	(0.0001141)	(0.0001128)	(0.0001133)	(0.0001144)
TWV10 (-5)	0.000137***	0.000142***	0.000137***	0.000134***
	(0.0000463)	(0.0000489)	(0.0000483)	(0.0000496)
TWV10 (-4)	0.000119***	0.000124***	0.000117***	0.000113***
	(0.0000349)	(0.0000364)	(0.0000349)	(0.0000351)
TWV10 (-3)	0.000088***	0.000091***	0.000086***	0.000085***
	(0.0000311)	(0.0000306)	(0.0000284)	(0.0000281)
TWV10 (-2)	0.000079*	0.000085*	0.000080*	0.000079*
	(0.0000462)	(0.0000453)	(0.0000455)	(0.0000454)
TWV10 (-1)	0.000067*	0.000075*	0.000068*	0.000070*
	(0.0000404)	(0.0000404)	(0.0000405)	(0.0000407)
TWV10	0.000062	0.000066	0.000062	0.000063
	(0.0000433)	(0.0000451)	(0.0000434)	(0.0000435)
TWV10 (+1)	0.000020	0.000025	0.000018	0.000018
	(0.0000181)	(0.0000184)	(0.0000182)	(0.0000189)
TWV10 (+2)	0.000001	-0.000005	-0.000009	-0.000009
	(0.0000150)	(0.0000196)	(0.0000180)	(0.0000177)
TWV10 (+3)	-0.000009	-0.000013	-0.000015	-0.000015
	(0.0000144)	(0.0000174)	(0.0000147)	(0.0000145)
TWV10 (+4)	-0.000028	-0.000029	-0.000028	-0.000029
	(0.0000262)	(0.0000288)	(0.0000257)	(0.0000249)
TWV10 (+5)	-0.000015	-0.000017	-0.000016	-0.000019
	(0.0000299)	(0.0000301)	(0.0000277)	(0.0000274)
TWV10 (+6)	-0.000025	-0.000024	-0.000020	-0.000023
	(0.0000363)	(0.0000357)	(0.0000287)	(0.0000270)
Drought Index	No	Yes	Yes	Yes
Drought Index Lags - 6	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes
Rain Lags - 6	No	No	No	Yes
Population	No	No	Yes	Yes
Total Corn Acres	No	No	Yes	Yes
Irrigated Corn Acres	No	No	Yes	Yes
Total Cotton Acres	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	Yes	Yes
Total Sorghum Acres	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	Yes	Yes
Total Wheat Acres	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	Yes	Yes
Total Rice Acres	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes
N	13437	13176	13176	13082
# Stations	254	254	254	250
R-Squared	0.122	0.136	0.160	0.162
Station and Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

B.D.2.C Leave-One-Out Tests

Table 27. Fixed Effects Model Results Leaving Out Monitoring Stations in the *Permian Basin*. Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00080*	0.00081*	0.00157***	0.00081*	
	(0.000469)	(0.000469)	(0.000134)	(0.000467)	
TWV10 x Eagle Ford					0.00110**
					(0.000486)
Drought Index	No	Yes	Yes	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	12904	12617	4371	12617	12617
# Stations	219	219	73	219	219
R-Squared	0.133	0.146	0.213	0.155	0.159
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 28. Fixed Effects Model Results Leaving Out Monitoring Stations in the *Eagle Ford Shale*. Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00009*	0.00011**	0.00062**	0.00009**	
	(0.000048)	(0.000050)	(0.000294)	(0.000045)	
TWV10 x Permian					0.00013***
					(0.000040)
Drought Index	No	Yes	Yes	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	14565	14278	5732	14278	14278
# Stations	256	256	105	256	256
R-Squared	0.116	0.128	0.169	0.138	0.138
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

Table 29. Fixed Effects Model Results with Lags and Leaving Out Monitoring Stations in the *Permian Basin*. Lags Indicated by (-). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00009	0.00009	0.00026**	0.00009	0.00009
	(0.000128)	(0.000138)	(0.000094)	(0.000129)	(0.000127)
TWV10 (-1)	0.00010	0.00012	0.00027***	0.00011	0.00011
	(0.000107)	(0.000101)	(0.000041)	(0.000102)	(0.000102)
TWV10 (-2)	0.00014	0.00015	0.00035***	0.00014	0.00014
	(0.000117)	(0.000114)	(0.000056)	(0.000112)	(0.000113)
TWV10 (-3)	0.00016**	0.00016***	0.00019**	0.00016***	0.00016***
	(0.000065)	(0.000061)	(0.000070)	(0.000055)	(0.000056)
TWV10 (-4)	0.00019***	0.00020***	0.00026***	0.00020***	0.00019***
	(0.000050)	(0.000048)	(0.000037)	(0.000041)	(0.000041)
TWV10 (-5)	0.00021***	0.00022***	0.00036***	0.00022***	0.00022***
	(0.000070)	(0.000080)	(0.000043)	(0.000070)	(0.000070)
TWV10 (-6)	0.00024	0.00024	0.00054***	0.00024	0.00024
	(0.000153)	(0.000153)	(0.000089)	(0.000150)	(0.000150)
Drought Index	No	Yes	Yes	Yes	Yes
Drought Index Lags - 6	No	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Rain Lags - 6	No	No	No	No	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	11592	11331	3933	11331	11235
# Stations	212	212	72	212	208
R-Squared	0.136	0.152	0.230	0.179	0.180
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

Table 30. Fixed Effects Model Results with Lags and Leaving Out Monitoring Stations in the *Eagle Ford Shale*. Lags Indicated by (-). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00002	0.00002	0.00048**	0.00002	0.00002
	(0.000024)	(0.000024)	(0.000230)	(0.000021)	(0.000020)
TWV10 (-1)	0.00003	0.00004*	0.00048**	0.00003	0.00003
	(0.000022)	(0.000023)	(0.000228)	(0.000020)	(0.000020)
TWV10 (-2)	0.00004	0.00005**	0.00052**	0.00004*	0.00004*
	(0.000022)	(0.000022)	(0.000216)	(0.000021)	(0.000021)
TWV10 (-3)	0.00005***	0.00005***	0.00030***	0.00005***	0.00005***
	(0.000018)	(0.000015)	(0.000103)	(0.000015)	(0.000016)
TWV10 (-4)	0.00005***	0.00006***	0.00026**	0.00005***	0.00005***
	(0.000017)	(0.000017)	(0.000118)	(0.000015)	(0.000016)
TWV10 (-5)	0.00006**	0.00006**	0.00020	0.00006***	0.00005***
	(0.000023)	(0.000026)	(0.000164)	(0.000021)	(0.000020)
TWV10 (-6)	0.00003	0.00003	0.00029	0.00003	0.00002
	(0.000034)	(0.000037)	(0.000287)	(0.000034)	(0.000035)
Drought Index	No	Yes	Yes	Yes	Yes
Drought Index Lags - 6	No	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Rain Lags - 6	No	No	No	No	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	13036	12775	5107	12775	12679
# Stations	247	247	102	247	243
R-Squared	0.117	0.131	0.181	0.155	0.157
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

B.D.2.D Leave-Many-Out Tests

Table 31. Fixed Effects Model Results Leaving Out Monitoring Stations in the *Permian Basin* and *Eagle Ford Shale*. Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)
TWV10 (Unit = 100BBLs)	-0.00028	-0.00026	-0.00021	-0.00021
	(0.000184)	(0.000175)	(0.000144)	(0.000144)
Drought Index	No	Yes	Yes	Yes
Rain	No	Yes	Yes	Yes
Population	No	No	Yes	Yes
Total Corn Acres	No	No	Yes	Yes
Irrigated Corn Acres	No	No	Yes	Yes
Total Cotton Acres	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	Yes	Yes
Total Sorghum Acres	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	Yes	Yes
Total Wheat Acres	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	Yes	Yes
Total Rice Acres	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes
N	12168	11881	11881	11881
# Stations	208	208	208	208
R-Squared	0.129	0.144	0.152	0.152
Station FEs	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

Table 32. Fixed Effects Model Results with Lags and Leaving Out Monitoring Stations in the *Permian Basin* and *Eagle Ford Shale*. Lags Indicated by (-). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	-0.00008	-0.00007	0.00085**	-0.00005	-0.00003
	(0.000088)	(0.000092)	(0.000315)	(0.000087)	(0.000088)
TWV10 (-1)	-0.00012	-0.00006	0.00089***	-0.00007	-0.00005
	(0.000086)	(0.000071)	(0.000248)	(0.000076)	(0.000077)
TWV10 (-2)	-0.00009	-0.00003	0.00098***	-0.00002	-0.00003
	(0.000082)	(0.000070)	(0.000190)	(0.000069)	(0.000069)
TWV10 (-3)	-0.00005	-0.00003	0.00048	-0.00001	-0.00001
	(0.000066)	(0.000052)	(0.000280)	(0.000051)	(0.000052)
TWV10 (-4)	-0.00005	-0.00005	0.00053*	-0.00002	-0.00003
	(0.000065)	(0.000061)	(0.000284)	(0.000053)	(0.000051)
TWV10 (-5)	-0.00007	-0.00011	0.00046	-0.00007	-0.00007
	(0.000085)	(0.000104)	(0.000275)	(0.000079)	(0.000075)
TWV10 (-6)	-0.00010	-0.00013	0.00034	-0.00009	-0.00010
	(0.000102)	(0.000112)	(0.000312)	(0.000095)	(0.000097)
Drought Index	No	Yes	Yes	Yes	Yes
Drought Index Lags - 6	No	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Rain Lags - 6	No	No	No	No	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	10922	10661	3866	10661	10565
# Stations	201	201	71	201	197
R-Squared	0.132	0.149	0.213	0.173	0.174
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

B.D.2.E Leave-2011-and-2012-Out Tests

Table 33. Fixed Effects Model Results with Lags and Leaving Out Observations from *2011* and *2012* During the Major Texas Drought. Lags Indicated by (-). Outcome: Distance to Groundwater Level (Feet).

	(1)	(2)	(3)	(4)	(5)
TWV10 (Unit = 100BBLs)	0.00006	0.00007	0.00025***	0.00006	0.00007
	(0.000046)	(0.000049)	(0.000068)	(0.000049)	(0.000049)
TWV10 (-1)	0.00007*	0.00008**	0.00022***	0.00007*	0.00007*
	(0.000038)	(0.000039)	(0.000024)	(0.000042)	(0.000043)
TWV10 (-2)	0.00006**	0.00007**	0.00019***	0.00007**	0.00007**
	(0.000029)	(0.000029)	(0.000040)	(0.000032)	(0.000032)
TWV10 (-3)	0.00008***	0.00008***	0.00010**	0.00008***	0.00008***
	(0.000027)	(0.000027)	(0.000038)	(0.000026)	(0.000026)
TWV10 (-4)	0.00009***	0.00010***	0.00011***	0.00009***	0.00009***
	(0.000034)	(0.000035)	(0.000029)	(0.000035)	(0.000034)
TWV10 (-5)	0.00009**	0.00010***	0.00015***	0.00010***	0.00010***
	(0.000034)	(0.000037)	(0.000036)	(0.000035)	(0.000034)
TWV10 (-6)	0.00009	0.00009	0.00021***	0.00011	0.00010
	(0.000073)	(0.000073)	(0.000057)	(0.000074)	(0.000073)
Drought Index	No	Yes	Yes	Yes	Yes
Drought Index Lags - 6	No	No	No	Yes	Yes
Rain	No	Yes	Yes	Yes	Yes
Rain Lags - 6	No	No	No	No	Yes
Temp	No	No	Yes	No	No
Wind	No	No	Yes	No	No
Population	No	No	No	Yes	Yes
Total Corn Acres	No	No	No	Yes	Yes
Irrigated Corn Acres	No	No	No	Yes	Yes
Total Cotton Acres	No	No	No	Yes	Yes
Cotton Acres Irrigated	No	No	No	Yes	Yes
Total Sorghum Acres	No	No	No	Yes	Yes
Sorghum Acres Irrigated	No	No	No	Yes	Yes
Total Wheat Acres	No	No	No	Yes	Yes
Wheat Acres Irrigated	No	No	No	Yes	Yes
Total Rice Acres	No	No	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
N	10454	10248	3811	10248	10170
# Stations	248	248	99	248	244
R-Squared	0.127	0.148	0.216	0.187	0.188
Station FEs	YES	YES	YES	YES	YES
Year-Month FEs	YES	YES	YES	YES	YES

Standard errors in parentheses. Clustered on monitoring station. * p<0.10, ** p<0.05, *** p<0.01

Appendix References

Carr, H. 2017. 'Wipe Out! - How Will Permian E&Ps Dispose Of All That Produced Water?' RBN Energy LLC Daily Blog. Available at: https://rbnenergy.com/wipe-out-how-will-permian-eandps-dispose-of-all-that-produced-water. [Accessed 7/18/17].

Collins, G. 2017. 'Oilfield Produced Water Ownership in Texas: Balancing Surface Owners' Rights and Mineral Owners' Commercial Objectives.' Center for Energy Studies, The Baker Institute - Rice University. Available at: https://www.bakerinstitute.org/media/files/files/23bd889f/CES-pub-ProdWaterTX-020817.pdf. [Accessed 6/10/17].

Cook, M., and M. Webber. 2016. "Food, Fracking, and Freshwater: The Potential for Markets and Cross-Sectoral Investments to Enable Water Conservation." *Water,* 8(45): 1-21.

Dallas News. 2014. 'Fracking Companies Begin Slow Shift to Recycling Wastewater.' *The Dallas Morning News.* Available at: https://www.dallasnews.com/business/energy/2014/08/09/fracking-companies-begin-slow-shift-to-recycling-wastewater. [Accessed 7/17/17].

Ellsworth, W.L. 2013. "Injection-Induced Earthquakes." *Science,* 341(6142): 1225942.

Freyman, M. 2014. "Hydraulic Fracturing & Water Stress: Water Demand by the Numbers." Shareholder, Lender, & Operator Guide to Water Sourcing – A Ceres Report. Available at: https://www.ceres.org/resources/reports/hydraulic-fracturing-water-stress-water-demand-numbers. [Accessed 6/15/17].

Goldenberg, S. 2013. "A Texas Tragedy: Ample Oil, No Water." The Guardian – Environment. Available at: https://www.expressnews.com/business/eagle-ford-energy/article/Bigger-wells-but-more-water-12512262.php. [Accessed 8/9/18].

Herrnstadt, E., R. Kellogg, and E. Lewis. 2019. "Information Asymmetry, Drilling Distortions, and Oil and Gas Leases." Working Paper. Available at: https://home.uchicago.edu/~kelloggr/Papers/HKL_primaryterms.pdf. [Accessed 1/22/19].

Hiller, J. 2018. 'Bigger Wells, But More Water.' San Antonio Express News. Available at: https://www.expressnews.com/business/eagle-ford-energy/article/Bigger-wells-but-more-water-12512262.php. [Accessed 8/8/18].

Kondash, A., and A. Vengosh. 2015. "Water Footprint of Hydraulic Fracturing." *Environmental Science & Technology,* 2: 276-280.

Nicot, J.-P., R.C. Reedy, R.A. Costley, and Y. Huang. 2012. 'Oil & Gas Water Use in Texas: Update to the 2011 Mining Water Use Report.' Austin, TX: Bureau of Economic Geology, University of Texas at Austin. Available at: http://www.twdb.texas.gov/publications/reports/contracted_reports/doc/0904830939_2012Update_MiningWaterUse.pdf. [Accessed 6/21/17].

Kusnetz, N. 2012. 'The Bakken Oil Play Spurs a Booming Business – In Water.' High Country News. Available at: http://www.hcn.org/issues/44.13/the-bakken-oil-play-spurs-a-booming-business-in-water. [Accessed 6/4/17].

McCarthy, E.R., and Jackson, Sjoberg, McCarthy and Townsend LLP. 2012. "Texas Groundwater Ruling: Edwards Aquifer Authority v. Day." The Water Report – Texas Supreme Court Confirms Landowner's Ownership of Groundwater in Place. Available at: http://www.thewaterreport.com/NewFiles/Page%2019%20TWR%2099.pdf. [Accessed 9/5/17].

Nicot, J.-P., B.R. Scanlon, R.C. Reedy, and C.A. Costley. 2014. "Source and Fate of Hydraulic Fracturing Water in the Barnett Shale: A Historical Perspective." *Environmental Science & Technology,* 48(4): 2464-2471.

Pfeiffer, L., and C.-Y.C. Lin. 2012. "Groundwater Pumping and Spatial Externalities in Agriculture." *Journal of Environmental Economics and Management,* 64: 16-30.

Potter, H. 2004. "History and Evolution of the Rule of Capture." Conference Proceedings, TWDB Report 361, June 15, 2004. (quoting from East, 81 S.W., at 280).

Provencher, B., and O. Burt. 1993. "The Externalities with the Common Property Exploitation of Groundwater." *Journal of Environmental Economics and Management,* 24: 139-158.

Scanlon, B.R., R.C. Reedy, and J.-P. Nicot. 2014b. "Will Water Scarcity in Semiarid Regions Limit Hydraulic Fracturing of Shale Plays?" *Environmental Research Letters,* 9(12): 1-14.

Stevens, R.B., and G.L. Torell. 2018. "Exceptional Drought and Unconventional Energy Production." *Sustainability,* 10(4): 1-16.

Susquehanna River Basin Commission (SRBC). 2015. 'Water Resources Program – Fiscal Years 2016-2017'. Available at: http://www.srbc.net/planning/assets/documents/2016-2017_WRP_20150612.pdf. [Accessed 10/5/17].

TCEQ and TWDB. 2017. 'Priority Groundwater Management Areas and Groundwater Conservation Districts.' Report to the 85th Texas Legislature. Available at:

https://www.tceq.texas.gov/assets/public/comm_exec/pubs/sfr/053-10.pdf. [Accessed 6/20/17].

Texas Railroad Commission. Undated. 'Is Hydraulic Fracturing Flowback Fluid Recycled?' Frequently Asked Questions - Texas Railroad Commission. Available at: http://www.rrc.state.tx.us/about-us/resource-center/faqs/oil-gas-faqs/faq-hydraulic-fracturing/. [Accessed: 7/17/2017].

U.S. EPA. 2015. 'Case Study Analysis of the Impacts of Water Acquisition for Hydraulic Fracturing on Local Water Availability.' Available at: https://www.epa.gov/sites/production/files/2015-07/documents/hf_water_acquisition_report_final_6-3-15_508_km.pdf. [Accessed 6/6/17].

U.S. EPA. 2016a. 'The Hydraulic Fracturing Water Cycle.' Available at: https://www.epa.gov/hfstudy/hydraulic-fracturing-water-cycle. [Accessed 6/6/17].

U.S. EPA. 2016b. 'Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States.' Final Report. Available at: https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=332990. [Accessed 6/6/17].

Vissing, A. 2018. "One-to-Many Matching with Complementary Preferences: An Empirical Study of Market Power in Natural Gas Leasing." Working Paper. Available at: https://pdfs.semanticscholar.org/5014/ca081949118b77a24e5c0b75f050280ce9c6.pdf. [Accessed 4/2/19].

Walsh, R.W., and M.D. Zoback. 2015. "Oklahoma's Recent Earthquakes and Saltwater Disposal." *Science Advances,* 1(5): 1-9.